Mamoru Mitsuishi
Kanji Ueda
Fumihiko Kimura
Editors

Manufacturing Systems and Technologies for the New Frontier

The 41st CIRP Conference on Manufacturing Systems May 26–28, 2008, Tokyo, Japan

Springer

CD-ROM INCLUDED

CIRP

JSPE

Manufacturing Systems and Technologies for the New Frontier

Mamoru Mitsuishi • Kanji Ueda • Fumihiko Kimura
Editors

Manufacturing Systems and Technologies for the New Frontier

The 41st CIRP Conference on Manufacturing Systems
May 26–28, 2008, Tokyo, Japan

Springer

Editors

Mamoru Mitsuishi, Dr. Eng.
Department of Engineering Synthesis
The University of Tokyo
Tokyo 113-8656
Japan

Kanji Ueda, Dr. Eng.
Research into Artifacts
Center for Engineering
The University of Tokyo
Chiba 277-8568
Japan

Fumihiko Kimura, Dr. Eng.
Department of Precision Machinery
Engineering
The University of Tokyo
Tokyo 113-8656
Japan

ISBN 978-1-84800-266-1 e-ISBN 978-1-84800-267-8

DOI 10.1007/978-1-84800-267-8

British Library Cataloguing in Publication Data
CIRP Conference on Manufacturing Systems (41st : 2008 :
Tokyo, Japan)
Manufacturing systems and technologies for the new frontier
: the 41st CIRP Conference on Manufacturing Systems, May
26-28, 2008, Tokyo, Japan
1. Manufacturing processes - Congresses
I. Title II. Mitsuishi, Mamoru III. Ueda, Kanji IV. Kimura,
F. (Fumihiko), 1945-
670
ISBN-13: 9781848002661

Library of Congress Control Number: 2008924671

Cover design: eStudio Calamar S.L., Girona, Spain

Printed on acid-free paper

9 8 7 6 5 4 3 2 1

springer.com

Preface

The 2008 CIRP Conference on Manufacturing Systems, held at the Hongo Main Campus of the University of Tokyo, Japan, marks the 41st event of the conference series since its inception in 1969. The purpose of this conference is to promote research activities in various areas of manufacturing systems through the offering of a forum for the exchange of concepts, dissemination of technological breakthroughs, and discussion of future directions.

This year, the conference program covers a wide array of topical areas including manufacturing system evaluation, manufacturing system organization, implementation and design of systems, planning, human aspects of manufacturing systems, scheduling, manufacturing system design, service engineering, and novel processes in the area of electro-physical and chemical processes, cutting, machines and forming, micro-nano technology and surfaces, and grinding. The conference includes a total of 106 technical papers that have been accepted for presentation after a rigorous peer review and revision process handled by the Program Committee. These papers were submitted by corresponding authors from 15 countries including (in descending order of paper quantity): Japan, Germany, Sweden, China and Taiwan (R.O.C.), Italy, Netherlands, Norway, Slovenia, Brazil, Denmark, Finland, France, the Philippines, Thailand, and the USA. The papers are grouped into 32 sessions. The conference is grateful for the invited speeches given by Mr. H. Yamamoto, "Driving Innovation, An Industry Case to Enhance Manufacturing Competitiveness," and by Professor T. Fujimoto, "Architecture-based Comparative Advantage in Japan and Asia." The keynote speeches on "Manufacture and Sustainable Manufacturing" by Professor E. Westkaemper, on "Challenges for the Manufacturing Enterprise to Achieve Sustainable Development" by Professor J. W. Sutherland, on “Complex Adaptive Systems (CAS) Approach to Production Systems and Organizations” by Professor L. Monostori and on “Scientific Approach to Services: What is the Design of Services?” by Professor T. Arai will be of great interest to the conference participants. The program will also be enriched by the addition of the special talk by the Past President, Professor H. Yoshikawa, which is expected at the banquet.

Our sincere appreciation goes to the International Program Committee and Local Organizing Committee members for their wonderful efforts in reviewing papers, handling papers, and preparing the technical program. We would like to thank Associate Professor N. Sugita and Ms. A. Ishihama for their technical support and secretarial assistance in making this event possible. We would also like to extend our deep appreciation to the paper authors for their excellent contributions to the conference. The authors who are willing to share their most recent and critical findings with the research community represent the dominant factor in the success of this conference. We anticipate that the conference will be an exciting event and sincerely wish that all the participants enjoy and benefit from this meeting. Finally, the organizations and companies who contributed to the financial support of the conference also deserve our great respect for this successful event.

Mamoru MITSUISHI Kanji UEDA Fumihiko KIMURA

Co-organizers

The 41st CIRP Conference on Manufacturing Systems

Organization

Seminar founders

Colding, B. (USA)
Peklenik, J. (Slovenia)
Sata, T. (Japan)
Spur, G. (Germany)

Honorary chairman

Yoshikawa, H. AIST, Japan

Co-organizer

Mitsuishi, M. The University of Tokyo
Ueda, K. The University of Tokyo
Kimura, F. The University of Tokyo

International program committee

Alting, L. (Denmark)
Byrne, G. (Ireland)
Dornfeld, D.A. (USA)
Galantucci, L.M. (Italy)
Kals, H.J.J. (Netherlands)
Kjellberg, T. (Sweden)
Krause, F-L (Germany)
Lee, J.M. (Korea)
McKeown, P.A. (UK)
Monostori, L. (Hungary)
Rasch, F.O. (Norway)
Shpitalni, M. (Israel)
Teti, R. (Italy)
Tseng, M.M. (Hong Kong)
Van Brussel, H. (Belgium)
Weck, M. (Germany)
Westkaemper, E. (Germany)
Bouzakis, C.D. (Greece)
Chryssolouris, G. (Greece)
Duffie, N.A. (USA)
Jovane, F. (Italy)
Kimura, F. (Japan)
Klocke, F. (Germany)
Levy, G.N. (Switzerland)
Lee, K. (Korea)
Mitsuishi, M. (Japan)
Pritschow, G. (Germany)
Santochi, M. (Italy)
Suh, N.P. (USA)
Toenshoff, H.K. (Germany)
Ueda, K. (Japan)
Van Houten, F.J.A.M. (Netherlands)
Wertheim, R. (Israel)

Local organizing committee

Aoyama, T.	Arai, E.	Arai, T.	Fujii, N.
Fujishima,M.	Fukuda, Y.	Kaihara, T.	Kanda, Y.
Kimura, F.	Kunieda, M.	Mitsuishi, M.	Mohri, N.
Nakao, M.	Ohmori, H.	Onosato, M.	Shamoto, E.
Shibukawa, T.	Shirase, K.	Sudo, M.	Sugimura, N.
Sugita, N.	Suzuki, H.	Takata, S.	Takeuchi, Y.
Tani, Y.	Ueda, K.	Umeda, Y.	Yanagimoto, J.
Yoshimura,H.			

Contents

Plenary and Keynote Papers

Manufacturing System Evaluation

Manufacturing System Organization

Implementation and Design System

Planning

Human Aspects in Manufacturing Systems

Scheduling

Manufacturing System Design

Electro-Physical and Chemical Processes

Cutting

Machines and Forming

Micro-Nano Technology and Surfaces

Grinding

Service Engineering

Plenary and Keynote Papers

Driving Innovations, an industry case to enhance manufacturing competitiveness

Hironori Yamamoto

CANON ELECTRONICS INC.

Abstract

This paper gives an overview of the international position in the information and communication industry from a competitive standpoint. Canon has strength in the information technology, particularly in the field of imaging. Imaging technology is Canon's core competency. Digital technology does not replace imaging technology, rather, digital technology supplements imaging technology. With this background, this paper reviews activities at Canon to strengthen its competitive edge, mainly from technology development angle. Although the reality has multiple factors, this paper focuses on product innovation in new product development, aimed to maximize customer value. It also provides outline of processes innovation.

Keywords:

Innovative Target; Product Innovation; Process innovation

1 INTRODUCTION

An Index to measure the competitiveness of a nation is published by the Institute for Management Development (IMD), Switzerland in "World Competitiveness Yearbook" (WCY) since 1989. Between 1989 and 1993, Japan was ranked number one. Since then, Japan's competitiveness declined. In 2007, Japan was ranked number 24 in the Overall Performance, down from 16th the previous year. Among the four Competitiveness Factors, Japan maintains relatively high position in Infrastructure thanks to good accumulation from the past. Low Government Efficiency and Business Efficiency contributed to lower Overall performance. This proves insufficient efficiency of the Japanese operations in the times of highly-sophisticated information society. When one reviews the transition of the Japanese Gross Domestic Product (GDP), in 1995, the Japanese production was 71% of the United States, while in 2005, it dropped to 37% of USA. This is a reduction of 14%. This illustrates that while the Japanese domestic economy was growing mildly, Japan is falling behind relative to other economies. Globalization is in progress. Such time requires social systems that create new set of values. Innovation drives creation of new values. This paper attempts to narrow the focus to a challenge of an enterprise in driving innovation.

2 THE IT INDUSTRY AND ISSUES FOR CANON

Originally, Canon was a manufacturer of imaging products, such as cameras and copiers. However, with the development of information technology, as more and more equipment became digital and network-capable, Canon was forced to adopt to the networked environment. Canon today manufactures and sells equipment that serves as man-machine interface in the highly developed imaging information. In other words, Canon entered the information technology industry. The structure of the IT industry is described in Figure 1.

Figure 1:Existing state of the competitive edge in IT industry.

The most basic layer is science, contents form the top layer.

Network system, assembled product, device, material, and manufacturing equipment all form layers below the contents layer. USA has advantage in contents and network systems, leading introduction of new concept products and standardization activities Assembled products and devices include digital home appliance, office equipment such as copiers, semiconductors, and electronic parts where Japan maintains high competitive edge. However, Japan's strength in these layers is diminishing. Flat panel displays is an area where high growth is expected. In 2000, almost 50% was manufactured in Japan. Today, Korea and Taiwan together have grown to have 75% manufacturing share. This example illustrates the competitiveness in assembled products and devices for Japan, area that Japan used to have competitive advantage enjoy, is diminishing. It is particularly true in price competitiveness. Figure 2 summarizes the issues Japanese IT industry face. It is viewed from two angles, management/corporate environment and technology. Most issues listed in management/corporate environment have been solved over the last several years, by rigorous selection and concentration of resources to specific business area and restructuring.

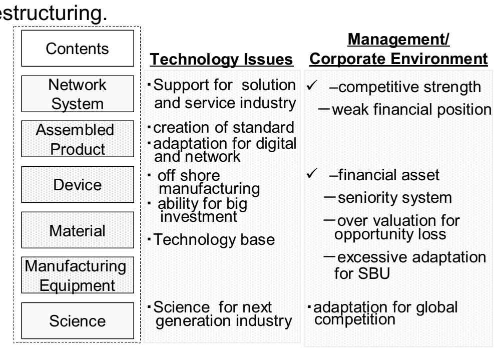


Figure 2: Issues for Japanese IT industry.

However, measures to address global competition accelerating in the IT industry are still insufficient. Among technology issues, lack of progress in technology supporting solution and service industries is a concern. Weak generation of technology coming from scientific field must be addressed. Through Management Reform since 1996, Canon removed these concerns greatly, but some issue remains.

3 CHALLENGES AT CANON

Activities at Canon to solve these management issues are described in this chapter. One of the characteristics of Canon management since its foundation is seen in the strong ties between long-term vision and management policies, embodied in its technology development. In 1996, Canon launched the “Excellent Global Corporation Plan” which promoted change from a venture company tendency placing higher priority on research and development, to enterprise-oriented tendency. The philosophy that forms the back bone of this conversion can be seen in shift from partial optimization to total optimization, shift from degree of completion to speed and quality, shift from “engineering is almighty” mentality to goal-orientated approach, and shift from revenue first to profit most. With the view of total optimization, the operation of research, manufacturing, and sales reflected lack of productivity at the manufacturing, typical of R and D oriented corporation leaving the manufacturing behind. Therefore, with the implementation of the “Excellent Global Corporation Plan,” manufacturing reform was driven under top management policy, which brought organized innovation on the manufacturing floor. Plant managers took initiative to remove ineffective practices and processes. Belt conveyor was removed, replaced by cell manufacturing system, so that the proficiency, originality and ingenuity of the workers may be brought out. Furthermore, this allowed ingenious attempt to improve productivity, a shift to more human-centric manufacturing. Factories reviewed its operation from logistics point of view. Inventory was reduced greatly. All these efforts add up to improved profitability, and made the plants more capable to cope with fluctuation. Processes went through great innovation on the manufacturing floor.

3.1 Approach towards Product Innovation

This chapter reviews technology development that took place at Canon, and describe approaches taken to gain competitive edge. As a result, Canon earned competitive advantage by possessing original technology. There are four points to consider.

(1)Processes leading to competitive advantage

(2)Important factors in setting “Innovative Target”

(3)Innovation by ongoing research and development

(4)R and D activity integrated and supported by Management mid and long-term direction

(4) is an important condition to concentrate resources driving innovation. This corresponds to technology management, linking R and D with long-term management vision. Here, (1) through (3) is explained in more detail. Business diversification that took place at Canon over the years, viewed from revenue size is described in Figure 3.

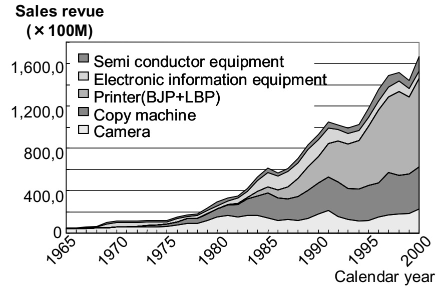


Figure 3:Concentric diversification in canon.

Canon business expanded from the single business of Camera, adding Copiers, Semiconductor equipment, and other products. Re-organizing this chart to show composition trends is shown in Figure 4, illustrating changes in the business structure.

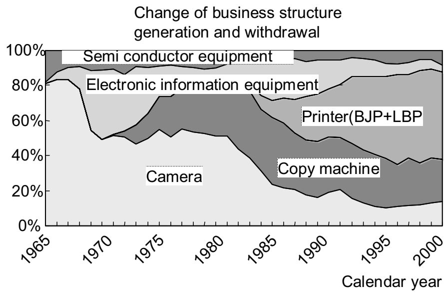


Figure 4:Concentric diversification in canon composition of each sales revenue.

One can find that it was not an easy expansion, rather, the reduction of sales, withdrawal, and creation of new business took place. Figure 5 describes processes taken to develop new product. During these processes, complex and compounding issues are broken down to individual task, which becomes a scientific or technology target. Innovation happens to solve such target. Depending on the magnitude of the target, there may be discontinuous innovation that accompanies paradigm shift, or there may be continuous innovation that may cause important differentiation to the product. Such processes allow enterprise to acquire core competency that are the source of corporate competitiveness.

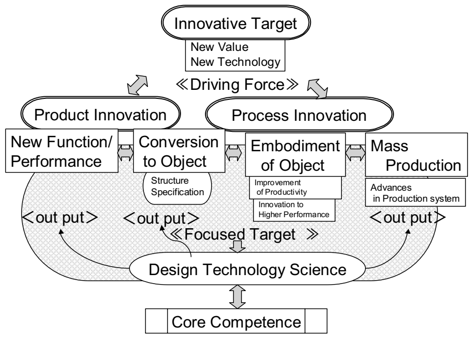


Figure 5:Innovation process and structure.

One can point out particular features of innovative target set by Canon. The first is what is called "Global niche." Market size of a product in one country is not big, but the world wide market is large. From its inception, Canon addressed the global market. Canon businesses in the Americas, Europe and Asia each accounts for about one third. The second feature is rather than conducting business within the mainstream of a technology, create an original product with original technology which is aligned to the technology trend. More specifically, Canon did not attempt to create a family of products based of digital technology. Rather, having imaging technology as Canon's core competency, Canon developed imaging products that enjoy the benefit of digital. Together with such product, Canon worked to strengthen competitiveness of the system. . The third feature is to create new value from new technology. Such new technology indicates solution to the innovative targets in the area of the first and second feature. Because they are new technology, they bring new values. New values are also the source of break through to win competitiveness. The fourth feature is to experience multiple trials. Retreat from business at an early stage minimizes loss, as the R and D investment is relatively small, and the technology itself can be saved. Data from the real market gained during such early business can be reflected to future products. Specific example at Canon includes entry to Electronic calculator and Personal Computer businesses, both closed later. However, technology acquired through such entry, and people who gained experience in developing these technology helped Canon in coping and adopting the new trend of Digital and Network in the businesses of Camera, Copier, and Printers. Figure 6 is a portfolio chart, showing value on the vertical axis, different technology on the horizontal axis.

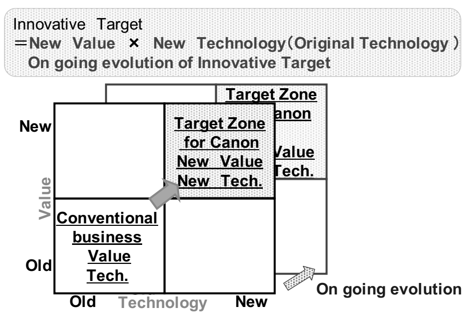


Figure 6: Innovative target.

This chart helps clarify our objective, and set the Innovative Target. Even though new value are created by existing technology, because it is based on current technology, others can easily catch up. Meanwhile, if new technology serves to create existing value, it is unlikely that the market will welcome this. Therefore, Canon created new products with values existing products could not deliver, with new technology. Innovative targets leads to competitive advantage because original value is created by original technology. Furthermore, by continuous R and D effort, evolution of original technology takes place, which elevates the strength of each original technology. This is how innovation targets are set at Canon. The structure to drive product innovation aimed for diversification is shown on Figure 7.

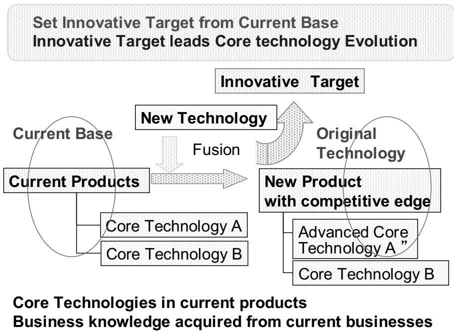


Figure 7: Processes to strengthen product competitive edge.

Core technology embodied in current product and the knowledge of the business domain form the basis to set innovative target. New technology and knowledge of the new business domain is added to consider the innovative target for diversification. Through this process, core technology the company possessed is being brushed up to become a stronger technology. Newly acquired technology and knowledge of the new domain becomes part of the core technology of the new domain.

3.2 Approach towards Process Innovation

Maintaining and enhancing competitive edge of a product is the driving force for process innovation. From technology stand point of view, the driving force is pursuit for product differentiation. Elements of differentiation can be found in improving product performance, and in reduction of cost.

When one looks at manufacturing technology from product life-cycle and its phase point of view, one can understand that different quality of technology needs to be emphasized depending on the stage.

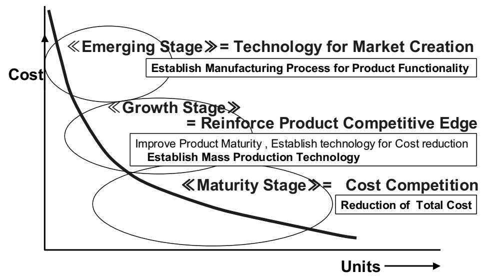


Figure 8: Correspondence between product life cycle and manufacturing technology.

During the emerging stage of new business development, as introducing the new product itself is difficult, establishment of technology elements to make this developing less difficult is given priority (establish manufacturing process.) In the growth stage, reinforcing the product competitive edge by improving maturity of the product and to reduce manufacturing cost is more important (Establish mass production technology.) In the maturity stage, in order to win the cost competition, reduction in total cost becomes important. So, at each stage of product life cycle, different driving forces seek diverse product differentiation. Thus, process innovation takes place involving most advanced knowledge and technology of the times, acquiring them interdisciplinary and concentrating them.

More specifically, here is the discussion on the case of optical industry. One example of process conversion in functionality realization and performance improvement can be found in the development of laser scanning optical system found in laser beam printers. The model is the copier, which established electro-photography process. The goal of this development is to print images from a computer, by emitting laser beam on the photoconductor drum to get dot patterns exposed.

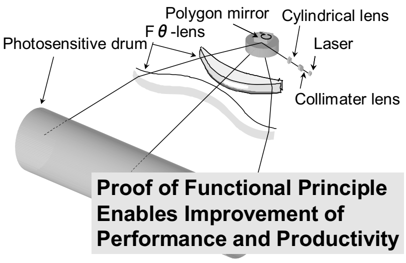


Figure 9: Optical scanning system in the LBP.

In the polygonal mirror scanning system (fig9), a toric lens was designed which dramatically relaxes accuracy requirement on the mirror rotation. In early days of development, rotation mechanism was ball bearing, material of polygon mirror was glass, and the toric lens was also made of glass. Inclusion of the toric lens from manufacturing process point of view is an improvement of the process, but it was important as it proved the functional principle of the scanning optics incorporating this special optical parts. Later, demands to improve performance, such as increase print speed, reduction of dot size, demands to reduce package size, request to lower cost were met by employing state of the art technology of the time, such as semiconductor laser, develop hydrodynamic bearing, and utilize new optical element.

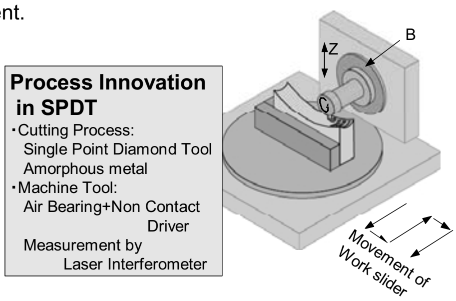


Figure 10: An example of metal molding, Toric lens application.

In particular, among the optical element, the polygon mirror, which was glass grinding and polishing process, was replaced by metal cutting work. Toric lens made by glass was replaced by plastic molding process, evolved to odd-shaped free-form aspheric lens. (fig.10) Early effort in learning single point diamond turning (SPDT) technology, which originated in the United States, helped refine manufacturing processes towards target performances in both cases of polygon mirror and metal mold of toric lens. Process innovation driven by need to introduce new functionality in a product places higher priority on realizing product functionality. In such case, applying specific technology to new products happens more predominantly, than large-scale process innovation to take place. At a later stage, when performance enhancement or cost reduction drives process innovation, transformation of material, or transformation of manufacturing technology takes place in a larger scale. Such cycle enhances competitive edge of the products, and as a result, build core competency.

4 SUMMARY

The structure of Canon product technology which has grown over the years is shown on Figure 11.

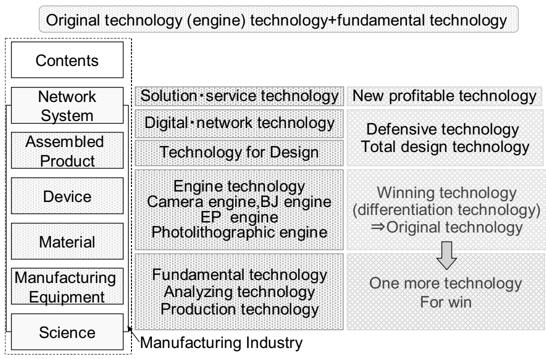


Figure 11:Technological structure of Canon.

Engine technology is positioned to win the competition. Digital and network technology, together with technology for design are positioned as defensive technology, and is systematically reinforced. More recently, the importance of fundamental technology is being recognized as technology to win. Experiments and tests of material element and other analytic approach accelerates conversion of desired function to physical state. Such approach reduces the chance of designing a product using parts with unknown character. This helps to speed up the design process and improve the design quality. Optical technology, material and process technology in the electrophographic products, inkjet process technology can only enjoy the benefit of advanced visualization and measurement techniques using super computers. In manufacturing technology too, technology regarding manufacturing equipment benefit from fundamental technology. Growth in digital technology brought effective tools. Proper use of these tools are another means to competitiveness

Architecture-based Comparative Advantage in Japan and Asia

Takahiro Fujimoto

Manufacturing Management Research Center, Graduate School of Economics, The University of Tokyo

1 INTRODUCTION

Asia has become a global center of manufacturing during the last quarter of the 20th century. First, Japan was the only major exporter of manufacturing goods from Asia. Then, yen was rapidly appreciated after The Plaza Accord in 1985, and newly industrialized economies (NIES) such as Korea, Taiwan, Hong Kong and Singapore emerged as exporters of relatively standardized goods. Japanese manufacturing firms also started to shift their production facilities mainly in ASEAN countries.

In the 1990s, China emerged as major exporters of certain labor intensive goods. NIES also continued to expand their manufacturing bases. Japanese economy stumbled, but its trade surplus continued to be significant. America came back as a center of digital network goods and softwares. How can we explain these dynamics of manufacturing competitiveness? In this situation, after all, we may better go back to the basics of comparative advantage theory.

Generally speaking, when there is good fit between a nation's characteristics and an industry's characteristics, the industry tends to enjoy competitive advantages in that country. Ricardo's Theory of Comparative Advantage implied that "good fit" is translated into relatively high labor productivity vis-a-vis other countries (Ricardo, 1971). Neoclassicists such as Heckscher, Ohlin and Samuelson advocated that countries having larger endowment of a certain productive resource (for example, labor-rich countries) will have better fit with industries that heavily use this particular resource (for example, labor-intensive industries), assuming that productivity is identical across the countries (Hecksher, 1949; Samuelson, 1948). More recent version of competitive advantages (e.g., Porter, 1990; Cho and Moon, 2000) also follows this tradition of fit between industry and country characteristics.

In more recent years, however, various phenomena that are difficult to explain using existing theoretical frameworks alone have been emerging. These phenomena include the recent fact that Japan has been apparently surpassed by China, Korea and Taiwan in some technology-intensive products (e.g., DRAM, CD media, DVD recorder), which were assumed to be Japan's stronghold for many years.

2 EXPORT COMPETITIVENESS OF JAPAN'S INTEGRAL ARCHITECTURE PRODUCTS

Against this background, the author advocated that we need an additional framework that focuses on "fit between organizational capacity and architecture" – a version of the comparative advantage theory seen from our observations of manufacturing activities on the shop floor.

Specifically, this framework argues that Japanese manufacturing firms, facing high economic growth amid shortages of work force, materials and money, tended to engage in economically rational long-term transaction/long-term employment. As a result, they built organizational capability that emphasizes teamwork among multi-skilled workforce, or "*integrative organizational capability of manufacturing*," which raised their productivity and quality simultaneously. Toyota Production System is a typical example of such a capability (Monden, 1993; Fujimoto, 1999).

On the other hand, it was thought that there are two basic types of product-process architecture: (1) "*Integral architecture*" with complex interdependence between product functions and product structures (such as automobiles, etc.); (2) "*Modular architecture*" in which the relationship between a product's functional and structural elements have a simple and clear one-to-one correspondence (such as personal computers, etc.) (Ulrich, 1995).

It was also thought that Japan, which is a country with a high endowment of "*integrative organizational capability*" stemming from its long-term employment and long-term transaction practices, tends to have a competitive advantage in "integral architecture" products – a prediction based on our "*architecture-based comparative*" hypothesis. In other words, Japan, where coordination-oriented organizational capability has been concentrated due to its historical trajectory in the late 20th century, tends to export *coordination-intensive goods*, or products with integral architecture.

3 PRELIMINARY EMPIRICAL RESULTS

With this framework of capability-architecture matching, can this new approach to industrial competitiveness demonstrate additional explanatory power for the reality of Japan's industrial competitiveness? Although the research is still at the exploratory stage, The Manufacturing Management Research Center (MMRC) at the University of Tokyo conducted a survey analysis of selected Japanese manufacturing firms in cooperation with the Ministry of Economy, Trade and Industry (METI). The survey targeted both assembled products and processed products (chemicals, etc.), including automobiles, household appliances, electronics, parts, industrial machines, chemicals, iron and steel, fibers, and food and drink (Fujimoto and Oshika, 2006).

As the results indicate, our "integral architecture index," constructed from about a dozen of questions regarding architectural characteristics of each product surveyed, and export ratio of the product (export value/domestic production value) in question generally statistically significant positive correlations (**Figure 1**). The positive correlations were

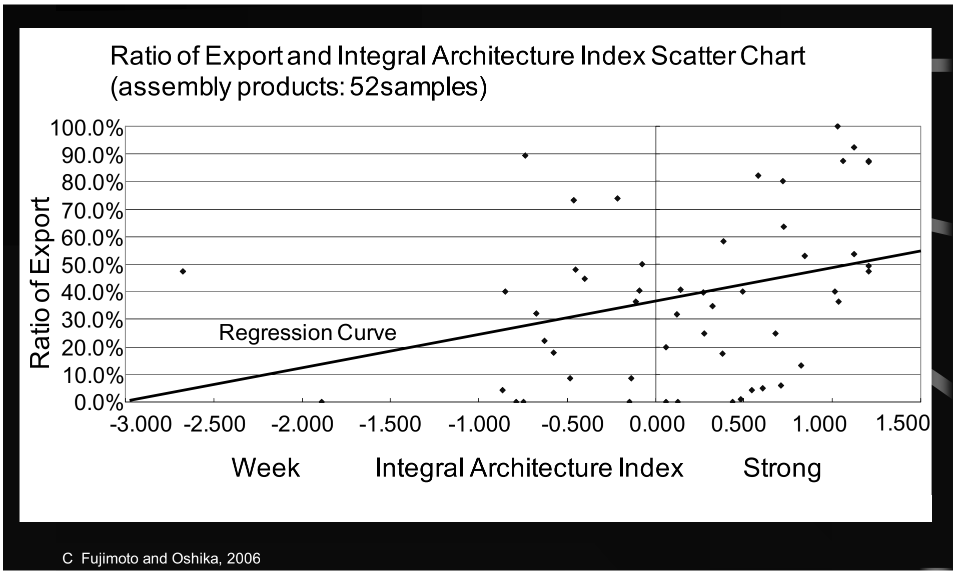


Figure 1: Ratio of export and integral architecture index (assembly products : 52 samles)

observed in both fabrication- assembly goods (e.g., machinery) and process goods (i.e., chemical). Also, the integral architecture index was positively correlated with not only export ratio, but also foreign activity ratio (export plus overseas production/domestic production), indicating that Japanese multinational firms tend to be good at integral architecture products wherever they are produced.

4 HYPOTHESES ON ARCHITECTURAL ADVANTAGE IN ASIA-PACIFIC AREA

Let's turn to architecture-based comparative advantage outside Japan. The following hypotheses are very preliminary and impressionistic ones, which are based mostly on ad-hoc empirical and historical observations of each geographical area (**Figure 2**).

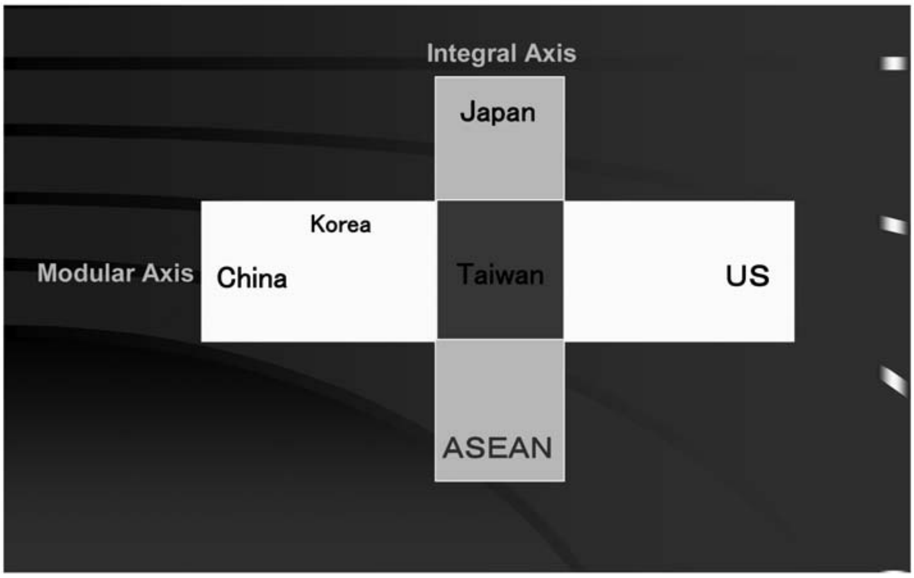


Figure 2 : Architectural gepolitics : a prediction in the pacific region

The basic logic is the same across the regions, however: Each region has its own historical path; A certain type of organizational capability tends to become concentrated in a certain region as a result of emergent capability-building process, which causes concentration of region-specific capability; Products with a certain type of product-process architecture and other characteristics tend to match better with a certain type of organizational capability, that results in relatively high productive performance (e.g., productivity, lead time, and quality).

Hypothesis on America: In a sense, America has been a country of immigrants in the past few centuries. In other words, it continued to attract human resources with industrial and technical knowledge and skills. For a society of this dynamism, it made sense to minimize coordination in order to make use of newcomers' capability as quick as possible.

As a result, American industries tended to emphasize division of labor, specialization, standardization of work, clear job demarcation, and use of market mechanism, while minimizing coordination efforts. Thus, the American System of Manufacture, throughout the 19th century, emphasized interchangeable parts and specialized equipment while minimizing coordination on the shop floor (e.g., fitter). American Mass Production System perfected this idea in the early 20th century. In the last decades of the 20th century, America rediscovered the power of a manufacturing system that economizes coordination cost—the Silicon Valley model of designing and producing digital network goods.

With this social and historical background, the framework of the architecture-based comparative advantage predicts that America-based firms tends to show comparative advantage in certain *technology-intensive modular architecture goods*.

Hypothesis on China: In the late 20th century, China, under the Communist Party regime, adopted Soviet-style national innovation system, in which industrial R&D activities were highly concentrated at the nation state level. Manufacturing firms in China were virtually equal to factories without R&D functions. The design of Chinese products tended also to lag behind that of advanced countries. Thus, when China chose an open economy path in the 1970s, many of its manufacturing firms, those in Southern coastal provinces in particular, had to acquire design information for their new products by licensing foreign technologies or copying foreign products.

For rapid catch-up of product design, many of the Chinese firms, state-owned or private, went for buying licensed or copied parts as generic modules and quickly started up new manufacturing businesses by mix-and-match of such de-facto generic components. The author calls this type of products "*quasi-open architecture*." Many of the machinery industries, such as motorcycle, truck, air conditioner, TV, and other digital consumer goods, were occupied by more than one hundred assembly makers. Copy parts themselves were also produced by hundreds of local suppliers. These firms also tended to rely on mix-and-match of standard equipment and low-wage temporary workers from low-income regions of inland China.

As a result, by the end of the 20th century, China became a major exporter of *labor-intensive modular architecture goods*. Thus, through a very different historical path, America and China became two major producers of relatively modular goods in the Pacific-rim side of the globe. This sharply contrasts with postwar Japan, which became a major exporter of integral architecture products.

Hypothesis on Korea: The most distinctive feature of the postwar Korean economy is a small number of large conglomerates, called Chaebols (e.g., Hyundai and Samsung), which somewhat resemble prewar Zaibatsu in Japan (The two share the same Chinese characters). Each Chaebol was controlled by its founder-owner and family. Because of strong top-down control by the founder-owners, Korean Chaebols tended to have strength in quick decision-making and investment on capital-intensive processes.

Thus, Korean large firms tended to have advantages in standard capital intensive goods, where mix-and-match of the latest production equipment results in competitive products, such as general purpose steels, DRAM, and CLD. In other words, Korean export power is highly concentrated in *capital-intensive modular architecture goods* produced by large firms, many of which stem from Chaebols.

Hypothesis on Taiwan: Taiwan is another significant exporter of manufacturing goods. Taiwanese economy maybe characterized as that of "competitive small country" (e.g., The Netherlands). Taiwan, because of its complicated history in the 20th century, and because of its geographical location (the intersection of America-China-Japan-ASEAN axes) has had strong economic links with the U.S., Japan, and mainland China. Taiwanese export-oriented firms tend to be good at making the most of their overseas linkages in building their organizational capabilities.

Where the products are modular and technology intensive (e.g., digital network goods), Taiwanese specialist producers tended to create networks with American firms. Where the products are integral (e.g., the automobile), Taiwanese firms tend to link themselves to the Japanese production networks. Thus, their strength resides in *versatility* of quickly moving between modular and integral architectures.

Hypothesis on ASEAN countries: As far as manufacturing competitiveness is concerned, ASEAN countries (except Singapore) have not demonstrated concentration of distinctive organizational capability. Although there is a significant degree of variety among ASEAN countries, none of them has industrial agglomeration of local firms that are technologically competitive. ASEAN countries have long functioned as production bases of the Japanese and Western multinational firms.

As such, ASEAN's manufacturing firms were mostly dependent on product designs originating from the multinational firms. Certainly, it is not realistic to foresee emergence of a cluster of ASEAN local firms with distinctive design capability in the near future. However, some of ASEAN countries, such as Thailand and Vietnam, may emerge as production bases of *labor-intensive integral architecture goods*. Their potential advantage over typical Chinese factories may be that it is easier for the former to keep multi-skilled workers with relatively low wages. China may possess a huge supply of low-wage single skilled workers, but the wage level tends to be higher and increasing for multi-skilled workers because of the volatile nature of Chinese labor market.

The key for this possible path toward integral goods is training for multi-skilled workers. In order for ASEAN economies to avoid direct competition against China, which is overwhelmingly strong in labor-intensive modular products, the former may find it beneficial to differentiate themselves from China by focusing on low-price, labor-intensive integral architecture goods. In order to produce such products competitively, it is crucial to strengthen teams of multi-skilled workers. The most effective training fields for this type of work force are, obviously, factories of Japanese firms. Thus, ASEAN firms may have a chance to become the export center of *labor-intensive integral architecture goods*, but only potentially at this point.

5 IMPLICATION FOR ODA IN ASEAN COUNTRIES

Japan's ODA to ASEAN nations have been historically significant in terms of its volume. It may need to be more strategic in the future. That is, a significant portion of Japan's ODA to ASEAN firms may be used for training of multi-skilled workers. Large scale systems and high-tech equipment may look spectacular, but it is difficult to differentiate to create distinctive manufacturing competence vis-à-vis China, a giant in modular manufacturing. The main players of such capability-building are Japanese and ASEAN manufacturing firms, but policy makers can assist their strategic activities.

Policy makers of both Japan and ASEAN may need to share a strategic vision and road map regarding manufacturing competitiveness in Asia. High technology and large systems may be favorite items for bureaucrats, but if all the countries go for such technologies, they are not necessarily the strategic solution for sustainable manufacturing competitiveness.

Asia-Pacific Area is probably the most competitive region in manufacturing. And this is the very reason why policy makers and industrialists in this region need to have a keen sense of comparative advantage. Architecture-based framework of comparative advantage may give them some additional insights. As Ricardo advocated, a country cannot be a major exporter of goods with all kinds. This principle holds true in the case of product architectures as well.

REFERENCES

Cho,D.-S. and Moon,H.-C. (2000) From Adam Smith to Michael Porter: Evolution of Competitiveness Theory. Singapore: World Scientific.

Fujimoto,T. (1999) The Evolution of a Manufacturing System at Toyota. New York: Oxford University Press.

Fujimoto,T. and Oshika,T (2006) "A Memo on the Hypothesis of Architecture-based Competitive Advantage" MMRC Discussion Paper --.

Hecksher,E.F. (1949) "The Effect of Foreign Trade on the Distribution of Income." In Howard S. Ellis and Lloyd A. Metzler, eds., Readings in the Theory of International Trade. Homewood: Irwin.

Monden,Y. (1993) Toyota Production System (2nd. ed.), Norcross, GA: Industrial Engineering and Management Press.

Porter,M. (1990) The Competitive Advantage of Nations, New York: Free Press.

Ricardo,D. (1971) On the Principles of Political Economy, and Taxation. Baltimore: Penguin.

Samuelson,P. (1948) "International Trade and the Equalization of Factor Prices." Economic Journal, 58, 165-184.

Ulrich,K. (1995) "The Role of Product Architecture in the Manufacturing Firm," Research Policy, 24, 419-440.

Manufuture and Sustainable Manufacturing

Engelbert Westkämper[1]

Fraunhofer-Institut für Produktionstechnik und Automatisierung (IPA), Stuttgart

Institut für Industrielle Fertigung und Fabrikbetrieb (IFF), University Stuttgart, Germany

Abstract

Manufacturing is permanent on change and research requires orientations to the requirements of the future. Together with stakeholders from Manufacturing in Europe a Technology Platform – Manufuture - has been developed to formulate the strategic orientations towards the Visions of 2020 and objectives of future Research. The basic model of Manufuture has been discussed several times in the CIRC Workgroup "Paradigms" as well as in the European Manufuture Conferences. The known Model has been added by the aspects of sustainability in manufacturing to take into account the pressure for environment protection and climate. This paper gives a short summary about the way if integration of sustainability in manufacturing systems.

Keywords:

Manufacturing Systems, Sustainability, Strategies

1 INTRODUCTION

The environmental problems, caused by the consumption of natural resources and pollution in the life of technical products leed to increasing political pressure and stronger regulations for manufacturing and usage of products. Manufacturing industries are additionally under the economic pressure to compensate increasing cost and create adding value. Under these conditions it is necessary to change the paradigms from former cost orientation to competition, adding value and sustainability. Manufacturing research has the potential to develop technologies for high competitive manufacturing, adding value and sustainability by changing the orientation and the criteria of optimisation to support the structural change of Manufacturing.

2 OBJECTIVES OF SUSTAINABLE MANUFACTURING

The European Union defined the following objectives in the so called Lisbon Agenda:

- Creating more value for more (growth) and better jobs,

- Increasing the competition of European industries and and the communities in the knowledge century,

- Sustainable development of economies.

Now new scientific based studies about the climate und consequences show a dangerous development: global warming etc.

The European Technology Platform Manufuture – established to formulated the goals and strategies for the European research followed the Lisbon Agenda [2]. Like 25 others, they became the base of the European research programme (FP7). Now we have to change the objectives gain to give the environmental aspects higher priorities and reduce the consumption of energy and material under the influence of increasing costs of input for manufacturing [6].

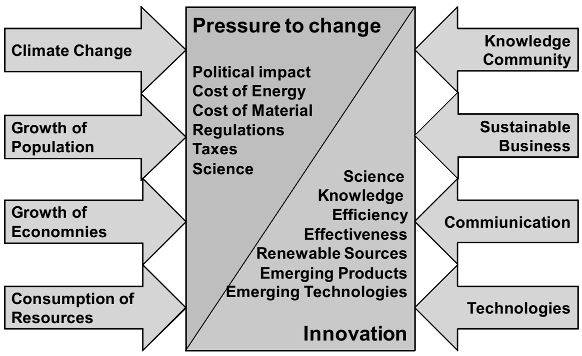


Figure 1: Pressure to change the paradigms of industrial manufacturing

Growth of population and growth of economies, which of course create more value and better life, make material and energy costly. The climate change (CO2) brings us political pressure by stronger regulations for emissions and pollution. On the other side we find the fields of innovation in material and process-technologies, the global information and communication with the electronics as drivers, the economic models and the knowledge of people which is expanding worldwide. Emerging technologies and products, renewable sources effectiveness and the efficiency of processes are needed as answers for the future of sustainable economies [3].

Manufacturing is the key area which has the potential for change [4]. But it is not only a question of research and engineering, it is even a question of business systems and activation of the human resources to implement innovation as fast as possible to change the practice from cost and profit optimisation to competition, innovation and sustainability.

Manufuture put the following objectives in the centre of the and strategic development in Europe and offers global cooperation:

Competitiveness of manufacturing industries

- to survive in the turbulent economic environment
- to compensate migration and consumption of technologies
- to have more and better jobs
- to stabilise economic results (groth)
- to ensure wellfare and social standards of living

Leadership in manufacturing technologies

- to support innovative products and platforms
- to lead manufacturing with global standards
- to garantee human and social standards of work

Environmental friendly products and manufacturing

- to reduce the environmental losses
- to change the consumption of limited resources
- to maximise the benefits of each product in the life cycle

Figure 2: Strategic Objectives of Manufuture

The holistic approach of the generic Manufuture model brings together the forces for change: Management, Products and Processes. A revolutionary change from tayloristic to sustainable manufacturing is required to solve the coming problems in the industrial world.

- sustainability of enterprises and business

- sustainable technologies in products and processes

- global envirionmental and social standards of work.

This can be summarised as the “New Taylor” for industrial manufacturing.

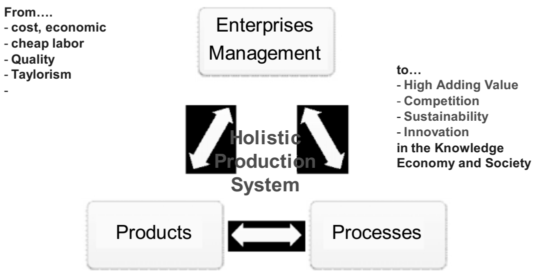


Figure 3 : The Manufure generic model – developed by CIRP

The CIRP generic model of manufacturing is a holistic view on the players of industrial development and their relations in the complex social and technical system of each enterprise. All of the players should follow a common paradigma of the future sustainable manufacturing.

3 INTEGRATION OF SUSTAINABILITY

Sustainability has many definitions and various dimensions. In the field of manufacturing, sustainability is a part of the optimisation of the overall efficiency of enterprises, products and processes. In this area Efficiency has the dimensions of economy, ecology and socials. Costs of energy or materials have an impact on the economic effectiveness. The reduction of ressources is a contribution to the economic and ecologic effectiveness.

Factories have a social dimesion. Humans influence the policy of enterprises develop products and processes and are a factor of cost. The holistic view includes the social effectiveness as part of the system.

The social effectiveness includes the skill and education and the conditions of work like ergonomic layout of workplaces, wagesystems, regulations of labor management, the cultural aspects of work and others.

The implementation of the dimensions of sustainability is possible by integration of the new paradigms in the so called holistic production systems. Production systems include the criteria of optimisation (cost, time, quality), the set of methodologies (lean, JIT, TQM,TPM etc.). Increasing ecologic or social effectiveness is a contribution to the economic effectiveness.

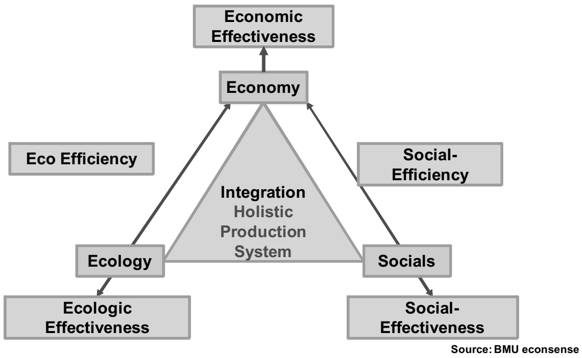


Figure 4 : Sustainability and Efficiency

The development of the basics and methodologies of holistic production systems, which have been born by Taylor [6] and which are mainly oriented to rationalisation and ecconomic efficiency have to be adapted to the requirements of future products, process - technologies and sustainability of the manufacturing system. Global standards are required in the global network of manufacturing.

4 IMPLEMENTATION OF MANUFUTURE

The Stakeholders of the Manufuture Platform defined five Pillars for the orientation of manufacturing. Taking into account the potential of new technologies and the variety of information and communication technologies in relation to industrial sectors and knowledge areas, the shown fields are parts of an overall strategy towards the development of next generation of factories.

The figure 5 shows the main fields of actions towards the European Production System for the sectors of capital intensive goods and consumer goods. Innovations in all of these action fields are contributions to European Manufacturing.

4.1 Driving force ICT for Manufacturing

We can state, that the diffusion of ICT has reached nearly all workplaces in Europe. It becomes an enabler technology for manufacturing. But ICT technologies influence even the private way of life. Basic innovations and basic research for ICT will find its way into the area of manufacturing the industrial market and are even drivers of innovation in manufacturing.

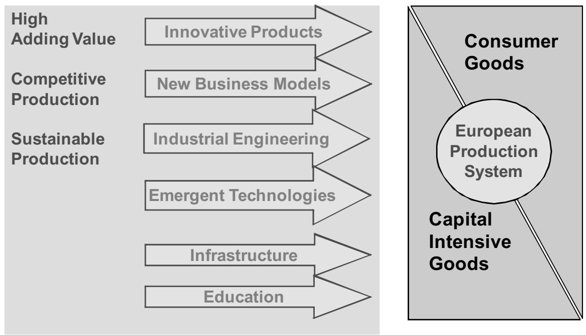


Figure 5 : Driving areas for Manufacturing development

The objectives follow the generic model of ManuFuture and objectives of High-Adding-Value and Sustainable Manufacturing. Strategic goals, targets and main visions elaborated by the ManuFuture actors follow in the next chapters. They can give a vision, which way manufacturing development should follow.

4.2 Emerging technologies

Emergent technologies are the drivers of the economic, ecologic and social effectiveness in manufacturing. The balance sheet classification covers the life cycle of products in the process chains from engineering to endof life and the öife cycle of factories from birth to Recyling. Factories are seen as komplex socio-technical products which have to be adapted to the needs of producst, markets and environment. Products meet factories in the manufacturing processes. Factories are usually elements of production networks. Their overall efficiency depends on the utilisation and the value chains. Under the aspects of sustainability the main criteria of optimisation are not only logistivcs but even the efficiency of the resources.

Manufuture defined the fields of emergent technologies, which are contributions to the overall economic, ecologic and social efficiency of both, the products and factories. The management and optimisation of the benefits (economic, ecologic and social) is one of the core areas of Research for manufacturing. This includes methodologies for engineering evaluation and management.

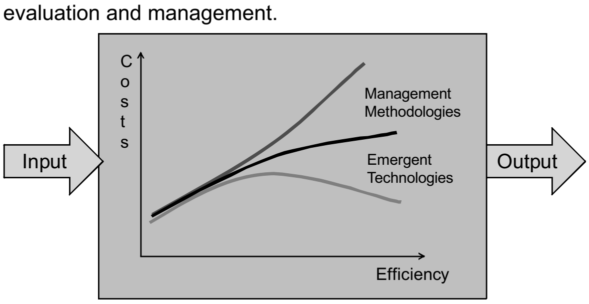


Figure 6 : Costs and Efficiency

Usually cost and efficiency have a linear and progressive relation. Industries need therefore methodologies and technologies to reduce the costs of higher efficiency. There are many examples which show, that emergent technologies have the potential to reach lower costs and at the same time higher efficiency: Dry cutting, integration of process chains, etc. They illustrate that manufacturing engineering is the key-technology of future development.

Manufacturing engineering the understanding of the holistic production systems includes the engineering of management systems, products and processes and has to use methodologies for optimization. The next figure shows the role of manufacturing engineering. More and more it becomes the goal as system-integrator of technologies from different areas: Process-technologies, materials, mechatronics, Nano-technologies or IT-methodologies to overcome existing state of the art in products processes and management.

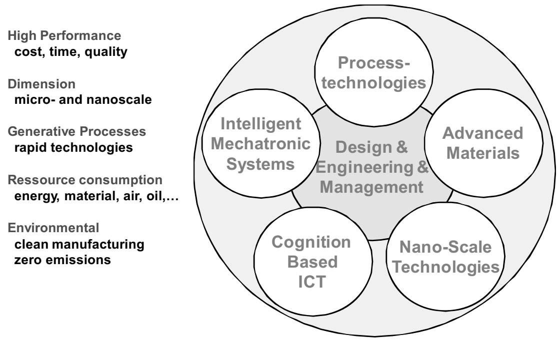


Figure 7: Development of emerging manufacturing systems

Manufacturing is on the way to overcome not only the existing limitation in performance, which was the main objective to increase economic efficiency but even in the dimensions (micro- and nano scale) to reduce the consumption of energy and material and in the consumption of resources or emissions. Following the lean-philosophy the consumption of energy and material is loss.

Technical intelligence is the long technical perspective of development toward disruptive factories

The set of methodologies of lean management defines the management system in series production [6,7]. The transformation and adaptation of these methodologies towards higher complexity, higher variants and changeable manufacturing aim at activating productivity potential in small batch production.

Lean management methodologies, which have been successfully implemented in automotive and supplier industries, are quasi standards of today's efficient manufacturing. The competitiveness under the European conditions of high wages and changing markets and technologies revealed that lean management can not reach the cost advantages in general.

The philosophy of lean mangement can easily take over the orientation of sustainable manufacturing by adding metghodolofgies of increasing the efficiencies (economic, ecologic, social) into the production system.

4.3 Survival in the turbulent environment

The development of the markets depend mainly on economic factors. In many industrial sectors cyclic fluctuations with ups and downs represent main factors of market turbulences. Manufacturing enterprises are influenced by multiple dynamic external factors concerning the products' behaviour in global markets, the strategies of competitors, the regional level of wage and reward systems including management of employee healthcare cost, regional infrastructure, the pace of technical innovations, the financial requirements of the investors and the financial constraints of operations, the robust supply of materials and components. Internal business factors such as qualification and capability of employees and the management, the demands and systems required by different customers, the utilization of resources and the capability of processes represent main influence factors as well. The enterprise environment is tough and turbulent. Only those enterprises can survive and be successful in this turbulent environment, which are robust enough and have the capability to continuous adaptations and transformations.

There is a requirement for sustainable manufacturing management supported by methods for:

- Balancing the load in mid term cyclic markets;
- Overcoming critical short term situations;
- Dynamic forecast;
- Adaptation of fix costs (dynamic systems)
- Dynamic work force models.

The strategies represent the main directions towards which research and technological development activities have to be oriented, with high priority on short term horizons.

.

4.4 Road Maps for Implementation

Investment in Research, Technology and Development for Manufacturing is necessary to achieve the ambitious goals and targets of competition and sustainability in the wide area of industries. Many enterprises are fighting for survival in the turbulent markets; others are leaders in markets and effectiveness. The third area seeks the future in emerging technologies. The European Way to activate the economic potential by research for manufacturing for the wide range takes care of the sector specific situation and transsectorial synergies to push industries and accelerate the diffusion process from basic research and innovation to applicative research.

The Road Map, developed by exterts of the european manufacturing industries and research is driven by industrial/economic requirements and the need of transformation of manufacturing towards Competitive Sustainable Development. More than 80% of the proposed activities to follow visions, strategic objectives and tasks of the ManuFuture Pillars are of common interest for all industrial sectors. The authors summarised them to several transsectorial Roadmaps, unified under a comprehensive approach representing the Manufuture Vison towards the European Industrial Transformation, further on called Manufuture Workprogramme.

The priopities of actions and research have been defined under the influence of reality. In many areas they did not follow actual academic views. Academia is still influenced by the economic orientation or by basics for technologies. The structural change from conventional manufacturing to manufacturing of the future and for sustainability including technical innovations (intelligent manufacturing even requires changes in the orientation of academic research.

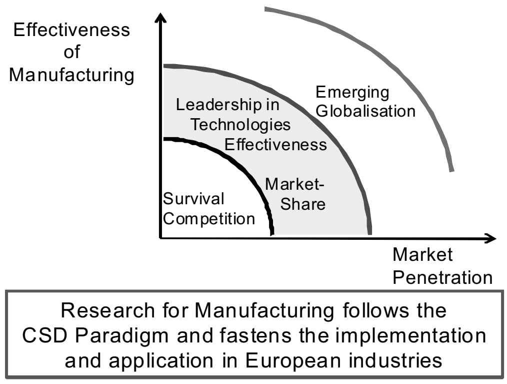


Figure 8: Transformation and Implementation

The figure shows the priorities of research following the paradigms of Manufuture and support the structural change towards the vision on manufuture: Competitive Sustainable Development (CSD).

5 SUMMARY

This paper summarized the visions and objectives of the Manufuture Technology platform in Europe. The Vision and the roadmaps for implementation follow the new paradigm of sustainable manufacturing. Sustainability can be integrated in the overall and holistic production systems. Research in Manufacturing should go ahead with the development of methodologies and intelligent technologies.

6 ACKNOWLEDGMENTS

I extend my sincere thanks to all who contributed to develop the Manufutrure Platform in Europe. Special thanks are given to my friend and CIRP colleague Franco Jovane, who startet in the CIRP Working Group paradigms and the generic model. It added the aspects of Sustainability and the visions of the factory of the future.

7 REFERENCES

[1] EC Conference “Manufuture 2003. European Manufacturing of the Future: role of research and education for European leadership”, Milano, December 1-2, 2003

[2] United Nations – Report of the World Commission on Environment and Development General Assembly Resolution 42/187, 11 December 1987

[3] Zhang, Hong C.; Kuo, Tsai C.; Huitian, Lu; Huang, Samuel H.; Environmentally Conscious Design and Manufacturing: A State-of-the-Art Survey, Journal of Manufacturing Systems Volume: 16, Issue: 5, 1997

[4] The Future of Manufacturing in Europe 2015-2020: The Challenge for Sustainability DG research

[5] Taylor, F.W., The Principles of Scientific Management, New York: Harper & Row 1911

[6]. James P. Womack, Daniel T. Jones, Daniel Roos; The Machine that Changed the World: The Story of Lean Production. New York: Harper Perennial, 1991

[7] Taijchi Ohno; The Toyota Production System: Beyond Large-Scale Production, Portland, OR: Productivity Press, 1998

Challenges for the Manufacturing Enterprise to Achieve Sustainable Development

J. W. Sutherland, J. L. Rivera, K. L. Brown, M. Law, M. J. Hutchins, T. L. Jenkins, and K. R. Haapala
Department of Mechanical Engineering-Engineering Mechanics
Sustainable Futures Institute, Michigan Technological University, USA

Abstract
Manufacturing enterprises are striving to achieve sustainability through changes in products, processes, and systems. Decision-support tools and methods are rooted not only in improving environmental aspects of manufacturing, but also in ensuring long-term productivity and social well-being. Refocused efforts on the development of sustainable technologies can further aid continuous improvement and stimulate revolutionary advancements industry-wide. Current and future challenges facing the manufacturing industry are addressed in terms of manufacturing enterprise, product life cycle design, and manufacturing processes and systems. Opportunities for future research are discussed within each of these areas.

Keywords: Sustainability; Life Cycle; Manufacturing

1 INTRODUCTION

Industry has been forced to evolve in the past in response to new regulations, technologies, and changing customer demands. At the dawn of the 21st century, there are increasing concerns about the sustainability of activities in developing nations and the industrialized world. Sustainable development calls for practices and decisions that will ensure that future generations have access to same opportunities that we presently enjoy [1]. For how long can we continue as a global society to extract natural resources, consume energy, and generate wastes with little thought for future generations? In response to growing concern for the environment, our assistance as technologists is needed to support corporate decisions related to manufacturing enterprises and product/process design all directed at helping to realize a sustainable future.

Sustainability is a globally emerging concept for that recognizes the interdependence of the economy, society, and the environment, frequently referred to as the three pillars of sustainability. Businesses are encountering increasing pressure from consumers, governments, and other organizations to address dimensions of sustainability, and increasingly, this involves more than such traditional measures of performance price and quality. Companies have begun to seriously consider product stewardship, reduction of hazardous substances, carbon footprint, energy and water consumption; and their role in society. Governments are instituting policies and companies are establishing strategies to support progress toward achieving sustainability. These policies and strategies serve as drivers for change, and many corporations have begun to critically examine their traditional practices. Manufacturers who do not evolve in response to these drivers will ultimately fall prey to global competitors who are operating under an evolving rule-set.

In particular, many recent concerns have focused on the continuing reliance on fossil fuel-based energy. Not only is accelerated use of non-renewable fuels unsustainable, but

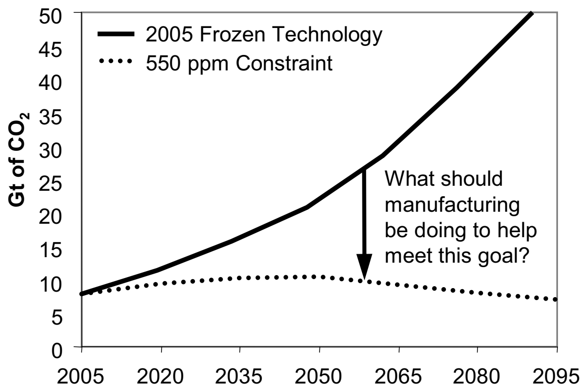


Figure 1: Carbon emissions under two scenarios: with frozen technology and with technology required to achieve 550 ppm concentration of atmospheric CO_2 [2].

combustion results in high levels of carbon dioxide (CO_2) emissions. CO_2 has been linked to climate change, and efforts are underway to reduce its emission to the atmosphere [2]. Figure 1 shows that unless changes are made, carbon emissions will continue to increase, and in less than a century will be nearly ten-fold the current level. Industry remains a large end-use consumer of energy; in the United States it is responsible for about one-third of the total energy demand [3]. The manufacturing sector has a responsibility to reduce energy consumption as well as undertake other efforts to ensure a sustainable future.

This paper begins to explore some of the sustainability-related challenges associated with manufacturing. These challenges will be presented in the context of the following areas: the manufacturing enterprise, product life cycle design, and manufacturing processes and systems. As a starting point for discussion, for each area, some goals for the research community are proposed.

2 MANUFACTURING ENTERPRISES ISSUES

It is increasingly the case that companies are accountable for sustainability-related impacts outside their direct control, including the impacts of supply chain partners. Companies are more and more frequently asking how to select partners that practice sustainability principles or who are socially responsible. What motivates this interest? As one example, it is becoming ever clearer that at some point in the not-to-distant future, embedded energy or carbon of materials and products will be taxed, regulated, or otherwise valued as part of costs. As a guide for supply chain-related decisions, it is believed the following issues must be addressed:

Standard measures for evaluating sustainability performance. As has been noted, corporations have much experience in considering such traditional metrics as cost, quality, and time to delivery in making supplier decisions. There is growing interest in the use of environmental metrics to describe corporate performance, and ISO 14000 certification represents an important first step in standardizing the methods by which all companies approach environmental sustainability. The Global Reporting Initiative has improved sustainability reporting through efforts aimed at standardization and transparency [4]. However, even for companies aware of their corporate-wide environmental footprint, it is very difficult to disaggregate this environmental information based on the materials, components, and products that they may produce. Standard eco-measures for products and processes are needed. In terms of societal sustainability, ISO 26000 is still several years away from being released. Still, it seems likely that formulating socio-measures of performance at the product/process level will be even more challenging than developing eco-measures.

Method for integrating the sustainability-related impacts of all the contributions to a product from a supply chain. Manufacturers create products using components/materials obtained from suppliers and through value-added operations performed on components and materials. Assuming that eco- and socio-measures of performance can be defined and quantified for each component, at issue is how to integrate these measures for the product. Hutchins and Sutherland [5] have discussed this issue and proposed a scheme that weights supplier performance measures by the economic value the supplier-delivered component contributes to the product; the scheme also incorporates the value-weighted performance of the manufacturer itself.

Development of EOUP (end-of-use product) management strategies and associated logistics. As corporations contemplate a future where used products are recovered and returned for recycling and remanufacturing, increased attention must be given to tracking the “inventory” and status of products that are currently being used. Moreover, consideration must be given to the best manner to recover and transport these EOUPs for subsequent processing [6].

The foregoing discussion highlights the fact that there are significant technical challenges that must be addressed to establish meaningful measures of sustainability performance as well as a method for the integration of these measures. An emerging paradigm shift in the notion of how used products are managed will drive many enterprise changes, and will require increased use of life cycle thinking to ensure that any action taken is not counter-productive. It should be evident that the record-keeping associated with such ventures will be significant.

3 PRODUCT DESIGN ISSUES

Products should be designed to respond to the needs of consumers. In the 21st century, the challenge of sustainability suggests that designers must also factor the environment and broader social interests into their decisions. It is now well recognized that designers should consider the entire product life cycle in this decision-making process.

Can we continue to tolerate a situation where waste is “designed into” products? Where manufacturing processes consume vast amounts of energy and materials and produce waste of all forms? Where used products are disposed with little regard to potential environmental impact and value lost? The traditional life cycle includes materials extraction and processing, manufacturing, distribution, use, and end-of-life stages. End-of-life for many products means disposal/ landfilling. Figure 2 suggests a product life cycle to which we should aspire. Green design is employed to reduce environmental impacts across the life cycle, green manufacturing processes are used that require little energy and emit zero wastes, and preventative maintenance is employed to extend product life. The figure shows EOUPs being recovered for reuse, remanufacturing, and recycling.

3.1 Life cycle assessment

Any meaningful progress in terms of product design that promotes sustainability will require a life cycle assessment (LCA). A research challenge associated with this category is

Improvement of LCA methods and software tools. This challenge includes progress on identifying appropriate allocation schemes, guidelines for performing streamlined LCAs, and improved data quality. Gaps are present within existing tools and these must be identified and filled (e.g., not enough manufacturing process-specific information). Not enough is known about the uncertainties and risks embedded within LCA methods and data, and thus on the reliability of the methods themselves. Methods and data are needed for both environmental LCAs and societal LCAs.

As noted previously, weighting of various impacts requires consideration of trade-off analysis methods to simultaneously consider economic, social, and environmental issues.

3.2 Material selection

Material selection is a key element of design, and thus will have a significant effect on sustainability. Guidelines have been developed for the selection of materials [7] [8]. They often are general and do not provide the necessary specificity that designers require.

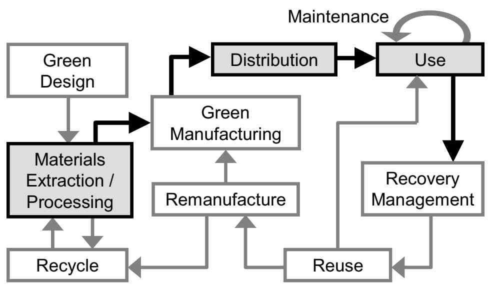


Figure 2: Product life cycle incorporating green design and manufacturing and used product recovery management.

Table 1: Embedded (processing) energy estimates of selected engineering materials (MJ/kg), adapted from [9].

Material	MJ/kg
Aluminum, virgin (recycled)	191 (8.1)
Float glass	15.9
Polypropylene	64.0
Steel, virgin (recycled)	32.0 (10.1)

The following is offered as a research challenge in the area of material selection:

Development of improved materials data for a variety of metrics. Some data are available on the environmental characteristics of materials, but the quality of these data may be suspect, and much property-related data is missing (e.g., energy consumed, waste produced, and water used per unit mass; recyclability; and biodegradability). Considerable progress is needed to tie materials level issues to such societal level issues as education, healthcare, and equity.

Furthermore, the sparse data that are available may be misleading. Table 1 provides estimates of embodied energy for several engineering materials; the values do not include manufacturing energy requirements. The table reveals large energy benefits associated with recycled aluminum and steel and suggests that aluminum is a superior alternative to steel. However, it must be remembered that the embedded energy of a recycled aluminum product must also reflect the significant energy required to process the virgin aluminum.

3.3 Design for the environment

Design for the environment (DFE) encompasses strategies and techniques aimed at reducing environmental impacts through such efforts as extending product life, reducing use-stage environmental burdens, and facilitating post-use recovery, reuse, recycling, and remanufacturing. Research challenges related to DFE include the following:

Increased emphasis on modular design. This philosophy calls for modular, upgradeable platforms that reduce impacts through integration of recovered end-of-use components. Modules should be defined across platforms and product generations, and yet not limit design flexibility. Much work is needed in the general area of design modularity; improved knowledge of returning component availability and condition is also required. Modularity offers a significant manufacturing benefit and supports mass customization. Of course, as in all design-related decisions, modular design should be guided based on life cycle assessment knowledge.

Renewed attention to product dematerialization and service-oriented products. Environmental impact is virtually proportional to the amount of material in many products, e.g., fuel consumption of an automobile is proportional to its mass. Dematerialization calls for the development of products that employ less total material in meeting consumer needs. Methods for dematerialization include more efficient product designs, better materials, and transitioning from a tangible (material-based) product to a service (an intangible product) [10]. While efforts have been underway for decades to develop more material efficient product designs, the concept of integrated product-service systems that have a high service content is more recent phenomenon and is only beginning to receive attention from the research community.

Of course, the research challenges identified represent only two DFE issues, and many others exist. The principle of sustainability embraces societal issues, thus an analogous companion to DFE would be the philosophy of Design for Society. Presumably, DFS would be focused on developing integrated product-service systems that in addition to satisfying consumer needs, also address broader societal concerns. The creation of products for the developing world using appropriate technology is one example of DFS.

4 MANUFACTURING PROCESSES AND SYSTEMS ISSUES

Across their life cycles, products consume large levels of resources and produce substantial wastes. An automobile consumes nearly 1 TJ of energy across its life cycle [11]. While most of this energy is associated with vehicle use, about 125 MJ is expended in materials processing and manufacturing – and this happens over a relatively short time period. Williams [12] reported the majority of the life cycle energy associated with a desktop computer is attributable to manufacturing (81%). Clearly, manufacturing processes have a significant eco-footprint that requires attention.

In an era of increased attention to recycling and remanufacturing, more focus should be directed at the processes employed for these activities. Important challenges related to the manufacturing of virgin and used products exist and these challenges may be classified based on whether they are associated with virgin products (manufacturing processes) or used products (recovery processes). Systems-oriented challenges are also present for both classes. The text below presents some thoughts on the challenges associated with these two classes.

4.1 Manufacturing processes

Our goal for manufacturing processes should be to establish processes that consume very little energy and produce zero or near-zero waste. Moreover, in support of product and process design efforts, information on existing and proposed manufacturing processes needs to be available that relates process inputs to environmental impacts. The following challenges are proposed:

Establish improved information on the environmental impacts of existing manufacturing processes and explore new technological concepts for greener operations. In support of product and process design, and life cycle assessment, the environmental impacts (resources consumed and wastes produced) by manufacturing need to be better quantified. Where good data exists, it is often aggregated and difficult to allocate to a single operation. New processes and technologies need to be established that avoid the traps (and wastes) of traditional operations; for example, we need to identify techniques for eliminating material removal operations, perhaps via additive operations.

Development of systems that can accommodate the use of mixed manufactured and remanufactured components. As we consider manufacturing systems of the future, it may be that both manufacturing and remanufacturing operations are housed in the same facility. Manufacturing processes will produce virgin components and remanufacturing operations will process used components, perhaps with both flows of components flowing into products. Clearly, such a situation will require the development of new approaches to manufacturing system design and operations to accommodate the significant differences between manufacturing and remanufacturing.

4.2 Recovery processes

Recovery processes are essentially those processes used to transform a used product into a component or material that can be re-integrated into a product. In essence, recovery processes include dismantling, sorting, remanufacturing, and recycling operations. Proposed research challenges associated with recovery processes include:

Development of recovery process knowledge and technologies. Improved knowledge is required for all steps in the processing of a used product into either a refurbished component or material that can be formed into a component. This includes inspection, dismantling, sortation, purification, remanufacturing, and recycling processes. At issue is acquiring fundamental insights into the physics of the operations, with particular emphasis placed on the understanding how environmental impacts depend on process inputs. One key challenge is identifying, based upon inspection, whether to recycle or remanufacture a used product, early during the recovery process before additional costly activities are undertaken. A continuing concern will be to develop effective recycling and remanufacturing processes.

Establishment of better manufacturing systems for recycling and remanufacturing. At a systems level, methods are needed that are less labor intensive and more flexible. For example, it would be ideal if dismantling facilities could disassemble a diverse variety of products.

As a general comment, while much work has been performed to characterize manufacturing operations and improve their performance, little of this work has been directed at environment-related issues. Many issues remain to be investigated.

5 SUMMARY AND CONCLUSIONS

Many individuals now believe that we are entering a sustainability revolution, and this revolution will present us with many issues to be addressed. The manufacturing research community needs to recognize that sustainability issues need to be addressed and critically examine all elements of the processes and systems for which we are responsible – what should we be doing to help achieve sustainable development? As a starting point for discussions, this paper has proposed several challenges for the manufacturing enterprise that we should consider as we contemplate the future:

- Standard measures for evaluating sustainability performance.
- Method for integrating the sustainability-related impacts of all the contributions to a product from a supply chain.
- EOUP management strategies and associated logistics.
- LCA methods and software tools.
- Improved materials data for a variety of metrics.
- Emphasis on modular design.
- Product dematerialization and in particular, service-oriented products.
- Information on the environmental impacts of existing manufacturing processes and new technological concepts for greener operations.
- Systems that can accommodate the use of mixed manufactured and remanufactured components.
- Recovery process knowledge and technologies.
- Better manufacturing systems for recycling and remanufacturing.

Additional interaction on these and other issues is welcomed and needed. After all, the stakes are high. Unless appropriate actions are taken, we run the risk of jeopardizing our future.

6 ACKNOWLEDGEMENTS

The authors gratefully acknowledge support from the National Science Foundation through the Sustainable Futures IGERT project (under Grant No. DGE 0333401).

7 REFERENCES

[1] World Commission on Environment and Development, 1987, Our Common Future, Oxford Univ. Press, NY.

[2] IPCC, 2007, Climate Change 2007: Synthesis Report, Summary for Policymakers, Fourth Assessment Report, Intergovernmental Panel on Climate Change.

[3] USEIA, 2006, Annual Energy Review 2006, Report No. DOE/EIA-0384(2006), U.S. Energy Information Admin.

[4] GRI, 2006, Sustainability Reporting Guidelines, Global Reporting Initiative.

[5] Hutchins, M.J., and J.W. Sutherland, 2008, An Exploration of Measures of Social Sustainability and Their Application to Supply Chain Decisions, J. Cleaner Production, to appear.

[6] Sutherland, J.W., K.L. Gunter, K.R. Haapala, K. Khadke, S.J. Skerlos, J.B. Zimmerman, W.W. Olson, and R. Sadasivuni, 2003, Environmentally Benign Manufacturing: Status and Vision for the Future, Transactions of NAMRI/SME, pp. 345-352.

[7] Ljungberg, L., 2007, Materials selection and design for development of sustainable products, Materials & Design, Vol. 28, pp. 466-479.

[8] Ashby, M.F.; and K. Johnson, 2002, Materials and Design: The Art and Science of Material Selection in Product Design, Oxford, 2002.

[9] Victoria University of Wellington, 2008, Embodied Energy Coefficients, Center for Building Research, available at http://www.vuw.ac.nz/cbpr/documents/pdfs/ee-coefficients.pdf, accessed January 2008.

[10] Reiskin, E.D., A.L. White, J.K. Johnson, T.J. Votta, 1999, Servicizing the Chemical Supply Chain, Journal of Industrial Ecology, 3/2-3, pp. 19-31.

[11] Sullivan, J., R. Williams, S. Yester, E. Cobas-Flores, S. Chubbs, S. Hentges, and S. Pomper, 1998, Life Cycle Inventory of a Generic U.S. Family Sedan, SAE Paper No. 98216.

[12] Williams, E., 2004, Energy Intensity of Computer Manufacturing: Hybrid Assessment Combining Process and Economic Input-Output Methods, Environmental Science and Technology, Vol. 38, pp. 6166-74.

Complex Adaptive Systems (CAS) Approach to Production Systems and Organisations

László Monostori[1,2] and Balázs Csanád Csáji[1]

[1]Computer and Automation Research Institute, Hungarian Academy of Sciences, Budapest, Hungary
[2]Department of Manufacturing Science and Technology, Faculty of Mechanical Engineering, Budapest University of Technology and Economics, Budapest, Hungary

Abstract
Theoretical study of complex systems receives more and more attention as most sciences broaden their perspectives. First, the paper overviews a few important complexity approaches, then it argues that complex adaptive systems (CASs) are especially important for production control research. As examples, firstly, a CAS based scheduling mechanism is described in which agents apply reinforcement learning to handle complex production control tasks, secondly, a semi-formal model of production networks is given combining stochastic processes, graph or network theory, and CASs. Further research issues are also highlighted.

Keywords:
Modelling, Production, Complexity, Adaptive Agents

1 INTRODUCTION

Growing complexity is one of the most significant characteristics of today's manufacturing, which is manifested not only in manufacturing systems, but also the products to be manufactured, in the processes, and the company structures [1]. The systems operate in a *changing environment* rife with *uncertainty*.

The need to be able to measure the complexity of a system, structure or problem and to obtain quantitative relations for complexity arises in more and more sciences: besides computer science and engineering, the traditional branches of mathematics, physics, chemistry, biology and social sciences are also confronted more and more frequently with this problem [2].

Complex Adaptive Systems (CASs) represent a relatively new theory, with the goal to study the structures and dynamics of systems and the question, how the adaptability of systems creates complexity [3], [4]. A CAS can be considered as a multi-agent system, where a major part of the environment of any given adaptive agent consists of other adaptive agents.

The main aim of the paper is to illustrate the appropriateness of the CAS approach for modelling different levels of production systems and organisations. Firstly, classical complexity measures are shortly surveyed, followed by enumerating some attempts to characterise the complexity in the manufacturing domain. The CAS approach is described shortly in Chapter 3, followed by chapters focusing on two different levels of the production hierarchy, namely the shop floor level and the level of the production networks. A CAS-based adaptive distributed production control is illustrated in Chapter 4, where the scheduling problem is formulated as a Markov decision process (MDP), and a three-level learning procedure is introduced. Chapter 5 gives a semi-formal model of production networks, combining stochastic processes, graph or network theory, and CASs. Further research issues conclude the paper.

2 COMPLEXITY AND ITS MEASURES

The meaning of the word "complexity" is vague, ambiguous, no universal, precise (e.g., formal) definition exists accordingly. Yet, there are approaches, especially in mathematics and computer science, which aim at defining special forms of complexity. In this section we provide a brief overview of some important classical complexity approaches.

2.1 Classical complexity measures

Since Alan Turing introduced his mathematical machines (viz., the Turing-machines) in the 1930s, they have become a fundamental tool for analyzing algorithms and problems. According to the theory of computational complexity, complexity is measured by the quantity of computational resources made use of by a particular task. Several complexity measures are known which are associated with algorithms [2], e.g., time-complexity, space-complexity and, for distributed systems, communication-complexity. Regarding the complexity of problems, the two most important classes of problems are P and NP. By definition, a P-problem is a decision problem that can be solved by a deterministic Turing machine in polynomial time. An NP-problem can be solved by a nondeterministic Turing machine in polynomial time. Roughly, problems in P are "easy" problems, while problems in NP are "hard" ones. Naturally, this classification is a simplification. We know that any P-problem is also an NP-problem. However, the question, whenever P = NP, is currently the most important problem of computational complexity theory.

Information complexity, viz. entropy, tries to measure the randomness or disorder of objects. This approach was suggested by Claude Shannon who in 1948 introduced entropy to communication theory [5]. Entropy provides a measure of the amount of information associated with the occurrence of given states. This is of key importance in information- and code-theory, however, can also be applied to measure other complex systems, e.g., graphs or networks.

Note that Shannon himself borrowed the concept of entropy from physics (viz. thermodynamics) by using the Boltzmann-Gibbs formulation of entropy.

In the 1960s Solomonoff, Kolmogorov and Chatin (independently) introduced a complexity concept which is often called algorithmic information complexity. Given a universal Turing-machine, the Kolmogorov complexity of a (bit)string (description) is the length of the shortest program that generates the description and halts. In other words, Kolmogorov defined the complexity of a structure as the length of its shortest description (namely, on a universal Turing machine). A structure is "simple", if it can be described by a short program, and "complex", if there is no such a short description. As an example, a random string is complex since its shortest description is specifying it bit-by-bit.

Also in the 1960s Krohn and Rhodes introduced a complexity definition aiming at measuring the complexity of abstract algebraic structures, e.g., groups and semigroups with the concepts of homomorphisms and wreath products. In computer science the Krohn-Rhodes theory gave new methods to build any finite state automaton using series-parallel emulation by simple components.

In the second half of the 20th century it became more and more important to measure the complexity of structures in natural sciences (e.g., in chemistry and biology). The theory of topological or network complexity addresses this problem and applies graph theory as its basis. There are several measures to define the complexity of a graph, some of which will be investigated later.

One of the latest complexity approaches is the theory of Complex Adaptive Systems (CASs). It has deep roots in the interdisciplinary field of multi-agent systems [6]. They will be investigated in Chapter 3.

2.2 Modelling of complexity in manufacturing

The complexity of manufacturing processes and systems, namely, how to deal with it, how to measure it, how to manage or control it, have come in the foreground of research in the past decades (e.g., [1], [6], [7], [8]). Decoupled models of product and process complexities were described in [6] where the quantity of information, diversity of information and the information content were considered as main influencing factors. As a continuation from the same authors, manufacturing operational complexity was also assessed addressing the human element of the production, too [7]. A complexity coding system for the main elements of a manufacturing system, i.e., machines, buffers and material handling equipment, was introduced in [8].

As regard to supply chains or production networks, even bigger complexity is to be faced. In [9] a focal buyer company with its supply base (the part of the supply network which is actively managed by the focal company), was treated. In the supply base complexity was conceptualized in three dimensions: the number of suppliers in the base, the degree of differentiation among them, and the level of inter-relationships among the suppliers. The authors concluded that though a reduction in complexity may lead to lower transaction costs and increased supplier responsiveness, in certain circumstances it may also increase supply risk and reduce supplier innovation. Consequently, reducing supply base complexity, in general, may be cost-effective, but blindly reducing it, may potentially decrease the buyer's overall competitiveness. The operational complexity of supplier-customer systems is modelled from an information-theoretic perspective in [10], capturing, in relative terms, the expected amount of information to describe the state of the system. The operational complexity was found to be associated with the operational costs of running a supply chain, as shown by queuing model- and simulation-based investigations [11]. However, dependency between the complexity index used and the costs was found in case of make-to-stock production, but not in the make-to-order case [11].

3 COMPLEX ADAPTIVE SYSTEMS

The theory of *Complex Adaptive Systems* (CASs) which was put forward by Holland [3], [4] is a new paradigm with the goal to study the structures and dynamics of systems and the question, how the adaptability of systems creates complexity. A CAS can be considered as a multi-agent system with seven basic elements in which "a major part of the environment of any given adaptive agent consists of other adaptive agents, so that a portion of any agent's efforts at adaptation is spent adapting to other adaptive agents". The first four concepts of Holland's seven basic elements, i.e., aggregation, nonlinearity, flow and diversity, represent certain characters of agents, are very important in the adaptation and evolution process, while the other three concepts, i.e., tagging, internal models and building blocks, are mechanisms of agents for communicating with the environment.

Environmental conditions change, due to the agents' interactions as they compete and cooperate for the same resources or for achieving a given goal. This, in turn, changes the behaviour of the agents themselves. The most remarkable phenomenon exhibited by a CAS is the emergence of highly structured collective behaviour over time from the interactions of simple subsystems [12]. The emergence of a complex adaptive behaviour from the local interactions of the agents is demonstrated in Figure 1.

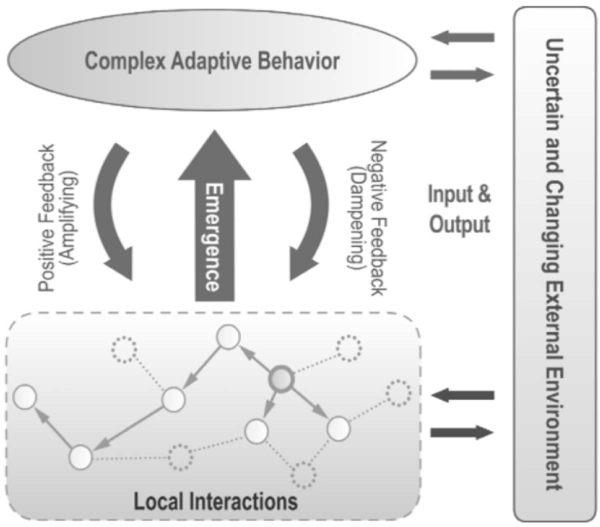


Figure 1: Emergence in Complex Adaptive Systems.

Both the CAS and its environment simultaneously co-evolve in order to maintain themselves in a state of quasi-equilibrium, i.e., on the edge of chaos [13].

In designing CAS, non-linear phenomena, incomplete data and knowledge, a combinatorial explosion of states, dynamic changes in environment and the frame problem are some notable examples of difficulties to be faced. The central question is realising an artifactual system that achieves its purpose in unpredictable conditions. It is difficult to approach problems like this by using only existing principles, such as analysis and determinism [14].

The Webster Dictionary describes *synthesis* as putting parts or elements together so as to form a whole, or the combination of separate elements of thought into a whole, as of simple into complex conceptions, species into genera, individual propositions into systems.

Synthesis is a necessary component of problem solving processes in almost all phases of the artifacts' lifecycle that starts with design, goes through the phases of planning, production, consuming and ends with the disposal of the product. Emergence plays a key role in solving difficult problems arising in synthesis.

The main concern here is whether and when, the completeness of information could be achieved in the description of the environment and in the specification of the purpose of the artifactual system. With respect to the incompleteness of information on the environment and/or the specification, the difficulties in synthesis can be categorised into three classes [15], [12]:

- *Class I*: Problem with complete description: if all the information concerning the environment and specification are given, then the problem is completely described. However, it is often difficult to find an optimal solution.
- *Class II*: Problem with incomplete environment description: the specification is complete, but the information on the environment is incomplete. Since the problem is not wholly described in this case, it is difficult to cope with the dynamic properties of the unknown environment.
- *Class III*: Problem with incomplete specification: not only the environment description but also the specification is incomplete. Problem solving, therefore, has to start with an ambiguous purpose, and the human interaction becomes significant.

3.1 Complex Adaptive Systems as Multi-Agent Systems

Complex Adaptive Systems have deep roots in the interdisciplinary field of *Multi-Agent Systems* (MASs) of artificial intelligence research, [4], [16], [17]. The formers are special cases of the latters which represent a general and flexible framework to describe and model (partially) autonomous systems including their interactions [16]. An agent is basically a self-directed entity, it is an object with its own value system and a means to interact (e.g., communicate) with other objects like this.

A MAS is formed by a network of computational agents that interact and typically communicate with each other. In *hierarchical* (guided) architectures, there are multiple levels of subordination relationships. In a *heterarchical* (self-organized, organic) architecture, agents communicate as peers, no fixed subordination relationships exist and, usually, the global information is eliminated, and, consequently, global optima cannot be guaranteed. On the other hand, advantages of these heterarchical multi-agent systems include: self-configuration, scalability, fault tolerance, emergent behaviour, massive parallelism, reduced complexity, increased flexibility, and reduced cost [18].

What distinguishes a CAS from a MAS is the focus on top-level properties and features like self-similarity, complexity, emergence and self-organization. Basically a CAS consists of large numbers of diverse entities (for example, agents) that are interconnected and have the capacity to change and learn from experience [4], [16]. Many of our most difficult problems centre on CAS.

4 ADAPTIVE AGENTS AT THE SHOP FLOOR LEVEL

4.1 Multi-agent systems in manufacturing

Engineering design; process planning; production planning and resource allocation; production scheduling and control; process control, monitoring and diagnostics; enterprise organisation and integration; production networks; assembly and life-cycle management; were enumerated in a very recent survey on agent-based systems in manufacturing [18]. The application fields cover all important aspects of manufacturing indicating the viability of the MAS-approach.

Agent technology is considered an especially important approach for developing distributed manufacturing systems. Holonic manufacturing systems (HMSs) consist of autonomous, intelligent, flexible, distributed, co-operative agents or holons [19]. One of the most promising features of the holonic approach is that it represents a transition between fully hierarchical and heterarchical systems [20]. HMSs with *adaptive agents* received a great deal of recent attention [21].

4.2 CAS-based adaptive distributed production control

In practice, the agents mostly have only incomplete and uncertain information on the environment (surrounding world) that they have to work in, additionally, this environment could be non-stationary. Moreover, they also have to face algorithmic-complexity problems, viz. even if they deal with static, highly simplified and abstract problems in which the solution is surely exists and can be attained in finitely many steps, they may not have enough computation power to achieve it in practice (as this is the case, for example, with many NP-hard problems).

A promising way to overcome these difficulties is the application of machine-learning techniques. It means designing systems which can adapt their behaviour to the current state of the environment, extrapolate their knowledge to the unknown cases and can learn how to optimize the achieved solutions.

Several complex systems can be effectively dealt with approximation. This means that we use a solution/model which is ”close” to the one that we have originally aimed at, however, it is much more simpler. For example, a lot of very-hard/complex (combinatorial) optimization problems can be efficiently handled, if we satisfy with approximate solutions, viz. solutions which are not optimal (suboptimal) but close to an optimal one. Naturally, the effectiveness of the approximation strongly depends on the used architecture and on the way we measure the distance between the objects.

Now, an adaptive iterative distributed scheduling algorithm is proposed that operates in a market-based production control system. The idea of negotiation based scheduling has emerged long before [22].

First, the basic frame of the approach is informally defined. In a multi-agent based manufacturing system, autonomous agents control different real world entities. In the presented system the two most important types of agents are the *resource agents* and *the order agents*. Resource agents control physical parts (such as machines, furnaces, conveyors, pipelines, material storages, etc.), while order agents control the production of a job. In the presented market-based production control system if a new job arrives at the system, a new order agent is created and associated with that job. An order agent or a group of cooperating order agents announces a sequence of operations and the resource agents can bid for that sequence. Only resource agents being able to do at least the first operation of that job are allowed to bid. Before an agent bids, it gathers information about the possible costs of making that sequence. If the sequence contains only one operation, the agent has all the information it needs, however, if the sequence contains other operations as well, which probably cannot be processed by the machine of the agent, it starts to search for subcontractors. It becomes a partial order agent and announces the remaining part of the sequence. The other resource agents which can do the next operation, may bid for the remaining operation sequence. Consequently, *a recursive announce-bid process* begins. At the end, when all the possible costs of that (partial) job are known, the agent bids. If the order agent which announced that job, is contented with it (it is the best bidder), the agent (and its subcontractors) get the job (award). Therefore, the schedule generation in the suggested agent-based system is a recursive, iterative process with announce-bid-award cycles based on market mechanisms.

The main problem with the mechanism described above is the *combinatorial explosion* of the possible schedules. More precisely, it makes a complete enumeration, in some sense, and thus, its time complexity makes it unusable in practice. The agents should not investigate every potential schedule, because this can be extremely time-consuming. If an agent wants to bid for an operation sequence and it needs information about the production costs of the part of the job, which it cannot do, it should not announce the part to every resource agent. It should make only a restricted tendering among the agents that will give a presumably good bid. They can apply *adaptive sampling* to learn the potentially good partners to cooperate with. These partner-value estimations can be learnt with *neurodynamic programming*, which is the combination of *reinforcement learning* and artificial neural networks, especially *kernel machines*.

Markovian Production Control

It can be shown that this scheduling approach can be formulated as a special *Markov decision process* (MDP) [17]. The aim of learning in an MDP is to find an optimal (or approximately optimal) *policy* that maps the *states* (possible situations) to the *control actions* available in that state. Formally, if we denote the state by x and the set of actions available in the state by $U(x)$ then an action a (form the $U(x)$ set) is executed by $\pi(x,a)$ probability, where π is called a control policy. In an MDP the state transitions are stochastic, however, the *Markov property* is assumed. There is a *reward* $r(x,a)$ associated with each state-action pair and the aim is to find such a policy that optimize the *expected cumulative rewards* over time [23]. The performance of a control policy in the long run is specified by its the *value function*. The value of a state with respect to a given policy is, roughly, the total amount of reward an agent can expect to gather starting from that state and following the policy thereafter. In case of *finite* MDPs, there always exists at least one optimal policy and each such policy shares the *same* optimal value function.

In theory, the optimal control policy of a (finite) MDP can be exactly computed by *dynamic programming* methods, such as value iteration, policy iteration or the Gauss-Seidel method [23]. However, due to the "*curse of dimensionality*" (namely, in practical situations both the required memory and the amount of computation is extremely large) calculating an exact optimal solution by these methods is practically infeasible, except for very small problems. Reinforcement learning techniques often try to use simulation as a sampling technique to overcome the computational demands by advanced statistical approaches, such as MCMC.

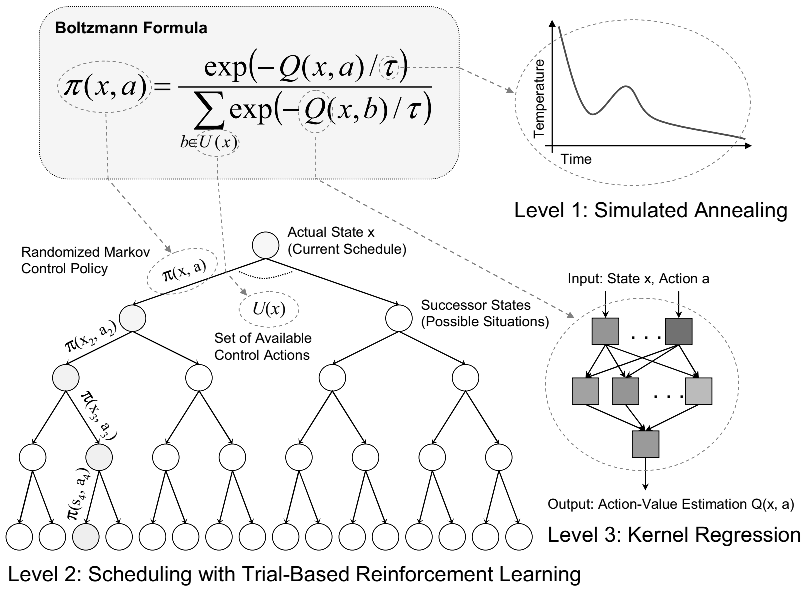


Figure 2: The three main levels of learning combined by the Boltzmann formula.

Three Levels of Learning

The paper suggests using Watkins' Q-learning algorithm [23] to calculate a near optimal policy. The aim of Q-learning is to learn the optimal *action-value function Q** rather than directly learning an optimal policy. The system can search in the space of feasible schedules by *simulating* the possible occurrences of the production process with the model. The trials can be described as state-action pair trajectories. After each sample (trial, iteration) the system makes updates asynchronously on the approximated values of the visited pairs according to the Q-learning rule. Only a subset of all pairs is updated in each trial. The paper suggests a triple-level learning mechanism to achieve effective production control, see Figure 2. The most important learning level is the level of reinforcement learning, since it computes an approximately optimal policy through adaptive sampling. The action-value function is represented as a kernel machine and the temperature of the system is controlled by a Metropolis algorithm. The progress of the system during learning can be described as follows (more details and examples are in [17]):

- Simulate a sate-action trajectory from the starting state using order and resource agents and a policy generated by the Boltzmann formula.
- After a terminal state is reached, back propagate the final performance and update the Q estimates of the visited states according to Q-learning.
- Fit a smooth approximating function to all of the available Q estimates, e.g. by SVR type Kernel Regression (Gaussian kernels).
- Decrease the temperature, except there were changes and disturbances in the system (in that case increase the temperature).
- Increase iteration counter and, unless some terminating conditions are met, go back to the first step.

5 A SEMI-FORMAL MODEL OF PRODUCTION NETWORKS

The conceptual overview of the proposed three-layer model can be found in Figure 3.

5.1 The environment as a stochastic process

We suggest a very abstract environment model: the uncertain behaviour of the environment should be described by a multivariate random variable and, since the environment may change over time, we consider a sequence of variables like this, one for each observable time step. These sequences are called *stochastic processes* [24].

Stochastic processes are standard models used in statistics, signal processing and machine learning. They consist of a sequence of random variables, $X_1, X_2, \ldots, X_{t-1}, X_t, X_{t+1}, \ldots$, where each X_t is a random variable, a measurable function from the sample space of a probability measure space to a measurable space of possible outcomes. They describe an event which is uncertain from the viewpoint of the observer. Multivariate random variables have vector output, namely, they render several values to an element of the sample space. They can be adequately described by their distributions. In the model proposed the environment is considered as a stochastic process, which may be assumed to be stationary or Markovian, i.e., we do not concern with the inner structure and the internal dependencies of the environment. It is treated as a black box, however, we still have a formal statistical model to work with.

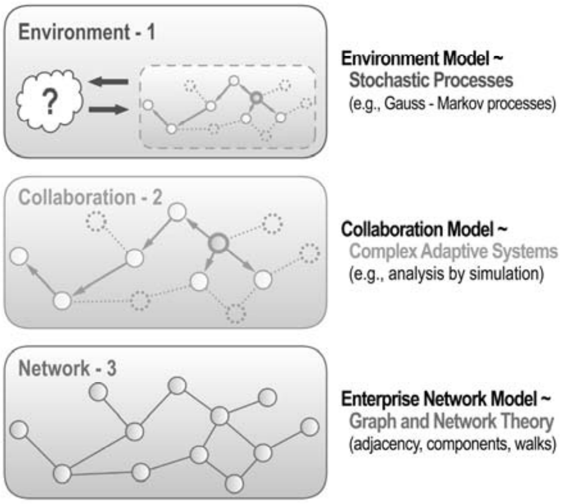


Figure 3: Conceptual overview of the proposed model.

In our model at each time t the state of the environment can be described with a multivariate random variable, X_t, and $X_{t,i}$ is a single valued random variable that describes a particular aspect of the environment we want to take into account. Such aspects are, e.g., the number of requests, the number of products that the customers have ordered, due dates, the trustiness of the customers, external costs, the economic situation (e.g., interest and currency rates, asset prices) or even the social, the cultural and the political situation.

As to measuring the complexity of random variables, the concept of (differential) entropy can be applied [5]. Note that the probability distribution of a random variable can be estimated by using accumulated historical data.

5.2 Modelling structures with network theory

It is assumed that the core topology of an enterprise network is quasi-static or slowly varying, hence it can be adequately modelled with network theory.

The basis of modern network theory [25], [26] and, hence, of network- or topological complexity is graph-theory, which is one of the fundamental theories in discrete mathematics. Its history goes back to Euler's celebrated solution of the Königsberg bridge problem in 1735.

The elements of graphs can be naturally associated with the elements of an enterprise network. An association, for example, could be as follows

- vertices ~ e.g., companies or functionalities,
- edges ~ e.g., connections between companies, functionality associations,
- vertex weights ~ e.g., prod. capabilities,
- vertex labels ~ e.g., competences,
- edge labels ~ e.g., connection type,
- edge weights ~ e.g., collaboration strength,
- colors, labels, weights ~ e.g., roles, types, strengths.

Network theory offers many "off-the-shelf" complexity measures that could also be applied to measure the complexity of the core-structure of enterprise networks [26], e.g., adjacency-related measures, symmetry-based measures, entropy-related measures, component-related measures, total walk count, the A/D Index. By applying results from network theory, measures for the static description of a network can be immediately used.

5.3 Collaborations as Complex Adaptive Systems

Probably the most important elements of an enterprise collaboration are dynamic and, therefore, hard to model and analyze. Here we suggest modelling the dynamic behaviour of an enterprise network as a CAS (Chapter 3).

6 CONCLUSIONS

The paper argued that the CAS approach is viable at different levels of manufacturing. Special emphasis was laid, on the one hand, on CAS-based production scheduling and control, by reinforcement learning, and, on the other hand, a semi-formal model of production networks, combining stochastic processes, graph or network theory, and CASs, illustrating the appropriateness of the CAS-based approach for handling complex, production-related tasks.

Complex adaptive systems, however, exhibit patterns of behaviour that can be considered archetypal or prototypical. One can benefit from the knowledge of these patterns [28]. Managing such systems an appropriate *balance between control and emergence* must be found [28]. The difficulty in understanding the effects of individual characteristics of the agents on their collective behaviour underlines the importance of using simulation as primary tool for designing and optimising such systems. In this respect, the proper *balance between simulation and theory* is to be aimed at [29].

Our further research activities will also go in this direction, i.e., to find the appropriate balances between simulation and theory, on the one hand, and between control and emergence, on the other.

ACKNOWLEDGEMENTS

The work was supported by the Grant “Production structures as Complex Adaptive Systems” of the Hungarian Scientific Research Fund (OTKA) and by the “Coll-Plexity” project belonging to the 6th Framework Programme of the EU.

REFERENCES

[1] Wiendahl, H.P., Scholtissek, P., 1994, Management and Control of Complexity in Manufacturing, Annals of the CIRP, 43/2: 533-540.

[2] Lovász, L., Gács, P., 1999, Complexity of Algorithms. Lecture Notes, Boston Univ., Yale University.

[3] Holland, J.H. (1992). Complex Adaptive Systems, Daedalus, Boston, 121/1: 17-30.

[4] Holland, J.H. (1995). Hidden Order: How Adaptation Builds Complexity, Helix Books, Addison-Wesley, NY.

[5] Shannon, C. E., 1948, A Mathematical Theory of Communication, Bell System Tech. J., 21: 623–656.

[6] ElMaraghy, W.H., Urbancic, R.J., 2003, Modeling of Manufacturing Systems Complexity, Annals of the CIRP, 52/1: 363-366.

[7] ElMaraghy, W.H., Urbancic, R.J., 2004, Assessment of Manufacturing Operational Complexity, Annals of the CIRP, 53/1: 401-406.

[8] ElMaraghy, H.A., Kuzgunkaya, O., Urbancic, R.J., 2005, Manufacturing Systems Configuration Complexity, Annals of the CIRP, 54/1: 445-450.

[9] Choi, T.Y., Krause, D.r., 2006, The Supply Base and its Complexity: Implications for Transaction Costs, Risks, Responsiveness, and Innovation, J. of Operations Management, 24:637-652.

[10] Sivasadan, S., Efstathiou, J., Calinescu, A., Huatuco, L.H., 2006, Advances on Measuring the Operational Complexity of Supplier-Customer Systems, European J. of Operational Research, 171:208-226.

[11] Wu, Y., Frizelle, G., Efstathiou, J., 2007, A Study on the Cost of Operational Complexity in Customer-Supplier Systems, Int. J. of Prod. Economics, 106:217-229.

[12] Ueda, K., Márkus, A., Monostori, L., Kals, H.J.J., Arai, T., 2001, Emergent Synthesis Methodologies for Manufacturing, Annals of the CIRP, 50/2: 535-551.

[13] Waldrop, M. (1992). Complexity, the Emerging Science at the Edge of Order and Chaos, VIKING, Penguin group, London.

[14] Monostori, L., Ueda, K., 2006, Design of Complex Adaptive Systems: Introduction, Advanced Engineering Informatics, Elsevier, 20/5: 223-225.

[15] Ueda, K., Vaario, J., 1998, The Biological Manufacturing System: Adaptation to Growing Complexity and Dynamics in Manufacturing Environment, CIRP J. of Manuf. Systems, 27/1: 41-46.

[16] Holland, J. H., 1998, Emergence: from Chaos to Order, Cambridge (MA), Perseus.

[17] Csáji, B. Cs., Monostori, L., Kádár, B., 2006, Reinforcement Learning in a Distributed Market-Based Production Control System, Advanced Engineering Informatics, Elsevier, 20/5: 279-288

[18] Monostori, L., Váncza, J., Kumara, S.R.T., 2006, Agent-based Systems for Manufacturing, Annals of the CIRP, 55/2: 697-720.

[19] Valckenaers, P., Van Brussel, H., 2005, Holonic Manufacturing Execution Systems, Annals of the CIRP, 54/1: 427-432.

[20] Bongaerts, L., Monostori, L., McFarlane, D., Kádár, B., 2000, Hierarchy in Distributed Shop Floor Control, Computers in Industry, 43/2: 123-137.

[21] Monostori, L., Kádár, B., Csáji, B.Cs., 2002, The Role of Adaptive Agents in Distributed Manufacturing, Proc. of the 4th International Workshop on Emergent Synthesis (IWES’02), Kobe, Japan, 135-142

[22] Márkus, A., Kis, T., Váncza, J., Monostori, L., 1996, A Market Approach to Holonic Manufacturing, Annals of the CIRP, 45/1: 433-436.

[23] Bertsekas, D.P., Tsitsiklis J.N., 1996, Neuro-Dynamic Programming, Athena Scientific

[24] Papoulis, A, Pillai, S.U., 2001, Probability, Random Variables and Stochastic Processes. McGraw-Hill.

[25] Barabási, A.-L.. 2002, Linked: The New Science of Networks. Perseus, Cambridge.

[26] Newman, M. E. J., 2003, The Structure and Function of Complex Networks. SIAM Review, 45/2: 167–256.

[27] Monostori, L., Csáji, B.Cs., 2007, Production Structures as Complex Adaptive systems. Proc. of the 40th CIRP Int. Sem. on Manufacturing Systems, Liverpool, UK

[28] Choi, T.Y., Dooley, K.J.,Rungtusanatham, M., 2001, Supply Networks and Complex Adaptive Systems: Control versus Emergence. J. of Op. Man., 19: 351-366.

[29] Surana, A., Kumara, S., Greaves, M., Raghavan, U.N., 2005, Supply-chain Networks: A Complex Adaptive Systems Perspective. International Journal of Production Research, 43/20: 4235–4265.

Scientific Approach to Services: What is the Design of Services?

Tamio Arai[1], Tatsunori Hara[1], Yoshiki Shimomura[2]
[1] Department of Precision Engineering, School of Engineering, The University of Tokyo, Japan
[2] Department of System Design, Tokyo Metropolitan University, Japan

Abstract
Scientific studies on services are discussed in this paper. After surveying previous literature, a series of studies developed by authors are introduced. Service Engineering has a top-down approach in definition. A service is represented by a set of functions and attributes. The flow of services denotes a graph by a multi-agent system. Customers' satisfaction can be computed by evaluation of each element of services. A CAD system for services, Service Explorer, has been developed to analyse services commonly by many stakeholders. The paper provides many literatures to understand scientific approach to services.

Keywords:
Service Engineering; CAD; Customers' satisfaction;

1 INTRODUCTION

Japanese government realized the innovation of sevice sectors may be the key to improve economic growth for the future inspired by Palmisano report. A new organization SPRING (Service Productivity Innovation for Growth) has made public recognitions of attractive services as Hi-Service 300 selection [1]. Although the government tries to activate research on service science/engineering, either service itself or service activities are still poorly modeled in computer.

Researchers of The Univ. of Tokyo studied service from various viewpoints and has resulted in the following proposal [2]; research should be promoted to have technologies

(1) to measure, visualize, and quantitatively evaluate the quality of services,
(2) to analyze, model, optimize, and implement services, and
(3) to effectively create, diffuse, and establish services.

Approaches to the three purposes may need many fields such as economics, psychology, anthropology, brain science and so forth. Representation and visualization would be the most important and thus the first step in the three.

The purpose of this paper is scientific studies for services by means of representing intangible and subjective objects *"services."* The word "scientific" here means a methodology to handle services objectively, i.e., everybody can describe and evaluate services in a common method. After surveying previous literatures, we introduce a series of concepts on Service Engineering the authors have developed.

2 RELATED LITERATURE

Let us survey the state of the arts of service research from the viewpoint of manufacturing. New studies such as Product-Service Systems (PSS) [3] [4] [5], Functional-Products [6], Function Sales [7], Service-Oriented Products [8] and Service Science [9] have been developed. The engineering targets that need to analyze and design is also shifting from simple products to service offering. The listed studies in this section have different understanding of relationship between products and service. Some believe the two are the same even though conventional marketing differentiates the two. As services are primary to visualize in intelligible way, several steps may be necessary: [definition] *what* a service is; [analysis] *how* a service can be decomposed into *what;* [evaluation] *how* to evaluate elements and the whole service.

Product -Service System(PSS)

PSS is a specific type of value proposition that a business offers to its clients, consisting of a mix of tangible products and intangible services designed and combined so that they are jointly capable of fulfilling final customer needs [3] [4]. Its viewpoints from business models may create much value onto products.

Function Sales (FS)

Products provide some effects onto customers; a refrigerator is necessary because of cooling functions. If foods are supplied directly from shops, a cool storage at home is not indispensable. Instead of selling materialized products, functions or its substitutes can be sold. The concept of function selling started to reduce the environmental load [7].

Service Engineering (SE)

SE here is a series of reseach [10] [11] [12] [13] made by an group in RACE, the Univ. of Tokyo since 2002, including the authors. It is characterized as top-down approach of service definition and representation. It has a great advantage in computer aided design system as we believe that the theory on service must be implemented in computer to prove its effectiveness.

Services Science, Management and Engineering (SSME)

IBM announced the importance of Service Science, and afterwards they use "Services Science, Management and

Engineering". The apprent definition is still ambiguous but aims at multi-disciplinary research and teaching on service activities. It covers the application of scientific, management, and engineering disciplines to tasks that one organization beneficially performs for and with another ('services')[14]

3 APPROACH OF SERVICE ENGINEERING

This section shows the fundamental scheme of Service Engineering, its hypothesis and approach.

3.1 Fundamental scheme of service modeling

Definition of Service

A service is defined as *a contract between a service provider and a service receiver to change the state of the receiver* [10]. This definition is broader than typical definitions in the traditional management and marketing fields starting with obvious difference from products (e.g.[15] [16] [17] [18]). It is quite similar to Hill's [19]: *A change in condition or state of an economic entity (or thing) caused by another*. According to this definition, most business activities are recognized as services, including manufacturing, selling and maintaining physical products. Emerging service research (e.g. [20]) also starts with discussion regarding such broader perspective.

Since our purpose is to describe a service for analysis, we need more precise and simple definition. Thus a service is re-defined in an agent system as illustrated in Fig. 1: *an activity that changes the state of a service receiver by means of a set of contents and channels provided by a service provider, normally with returns in compensation for the change.*

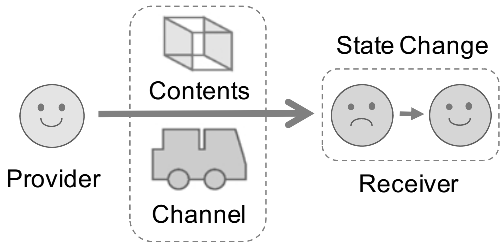


Figure 1: Definition of a service [10]

A service is the delivery of *service content* through *service channels*. Service contents may be materials, energy, or information that influence the receiver, while service channels transfer, amplify and control service contents. Only the contents cause the receiver to change. To define the change of the receiver, we introduced Receiver State Parameter (RSP) [10] [11], which are assumed to be observable and controllable. The total forwarding of a service, as inside of a round square in broken line of Fig. 1, is additionally defined as a set of parameters called *Serviset*: both parameterized contents and channels.

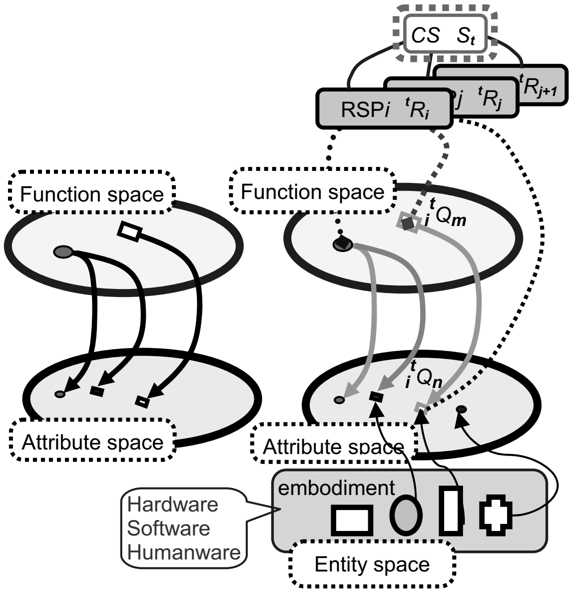
Yoshikawa's General Design Theory

Yoshikawa's General Design Theory (GDT) [21], illustrated in Fig. 2(a), provides a basis for our approach. The theory is discussed in terms of two topologies defined by functions and attributes of artifacts. The projection from functions to attributes can be universally recognized as design of products [22] [23]. Assuming that services can also be designed by the same projection, RSPs may consist of parameters in both function and attrributes. When a parameter is directly evaluated, it is called contents. The fuctions are realized by a set of entities, i.e., an embodiment of a service, whose characteristics are recognized as attributes. All the exiting parameters, sent from the provider to the receiver, construct a *serviset* as illustrated in Fig. 2(b). Therefore RSPs are supported by tree structures of the parameters. Each parameter may have quantitative value including Boolean logic and multi-value logic.

3.2 Customers' satisfaction

In traditional engineering fields, artifacts can be commonly represented by function of a product, and its effects in consuming process are separatedly designed. In service both provision and consumption occurs simultaneously. Thus customers evaluates not only the parameters connected to the products but also those in the service providing process. Service entities in Fig.2 represents both products and all activities.

Customers satrisfaction of a person t, S_t, in Fig. 2(b), is defined as weighted summation of RSPs as $S_t = \Sigma^t\omega_i\,{}^tR_i$, where ${}^t\omega_i$ is a weight of the ith RSP of a person t ; ${}^tR_i$ is denoted as a function of contents parameters ${}_i^tQ_j$ as ${}^tR_i = {}_i^tR({}_i^tQ_j)$. Then each ${}_i^tQ_j$ can be a fuction of other paremeters. These steps tells that customers satisfaction at the top of Fig. 2 (b) is defined as a function of contents parameters and channel parameters both in a tree structure in the graph of all the connected parameters.



(a) parameters in GDT (b) Parameters in SE

Figure 2: Parameters as functions and attributes

Tree representation in a view model

The purpose of the service modelling is to introduce quantitative evaluation on customer satisfaction. When a customer is fixed, its RSPs can be also listed up. Thus let us illustrate the tree of a single RPS, called *a view model* in SE, as shown in Fig.3. In other words, a view model is a tree structure of parameters all of which give direct or indirect effects to an RSP.

Figure 3 depicts a simple example of an RSP “comfortable environment” provided by a coffee shop to customers who do some computer tasks in the coffee shop while drinking coffee. In a view model, the designer starts by describing preliminary functions and/or attributes that give direct effects onto the RSP, i.e., contents parameters. Each of the parameters is deployed as several other parameters, i.e., sub-functions and/or sub-attributes. When constructing a relation trees from RSP to the detailed parameters, the lowest-level functions are associated with actual entities: hardware, software and humanware as realization of the service. These entities may also have innumerable attributes which are not listed in this view model. Some appears as content parameter in this view model, and may be a channel parameter in other view models. The effects from lower parameters to an upper parameter are represented using the functions of the fucntions.

3.3 Multi-agent systems for the flow of services

In Fig. 1, a service is defined as activity from a provider to a receiver, i.e., a single directed flow. A service, however, is commonly recognized as mutual interactions between a provider and a customer, or even among many agents. SE defines a single directional link between two agents as the simplest relationship illustrated in Fig. 4 (a). Let a provider be in green and a receiver in orange. Then a bi-directional relationship can be illustrated by means of a provider coinciding to a receiver as shown in Fig. 4(b), where one agen has two colors. Three agents are connected as Fig. 4 (c), where an intermediate agent may (i)process, (ii) assemble/dis-assemble, and (iii)store the serviset. For instance, a travel agent assembles several components of services such as air ticket, hotel accommodation and collectiong service at airports. The process of assembling components has significant similarity to manufacturing. The intermediate agent may have dual roles: a receiver evaluating services as materials from the upper stream, and a provider manufacturing service with added value to the lower stream. The intermediate agent in Fig. 4 (c) has two colors of both orage and green.

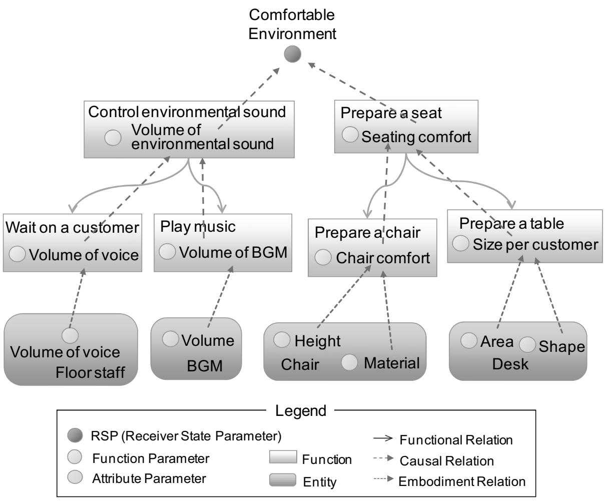


Figure 3: A simple example of a view model

An actual service system involves vairos activities and many stakeholders as illustrated in Fig. 4 (d). Among $2 \times {}_{N}P_{2}$ relationships for N agents, just important flows are analysed. Note that environment is included as agents in Fig. 4 (d), since it is recognized as a final receivers of manufacturing products. Introducing service flow as a multiagent system, we can optimize a system configuration and task allocations. Emprical studies indicates that customers satisfaction may greatly change by the redesign of service flow.

An agnet may be a person or an enterprise. The latter may have an inner strucutre in the agent recursively. An intermediate agent may have also the inner structure expressed by a simple structure as shown in Fig. 4 (e). The structures are expressed by colored PETRI net [24] [25].

3.4 Static representation and Dynamic representation

Aforementioned function-attribute model represents static aspects of a service. As customer’s satisfaction is related to service entities in the view model, designers can estimate the customers’ value as well as the product. Service interactions can be expressed by parameterized characteristics in the static model. However, the representation includes little information with regards to the service delivery process.

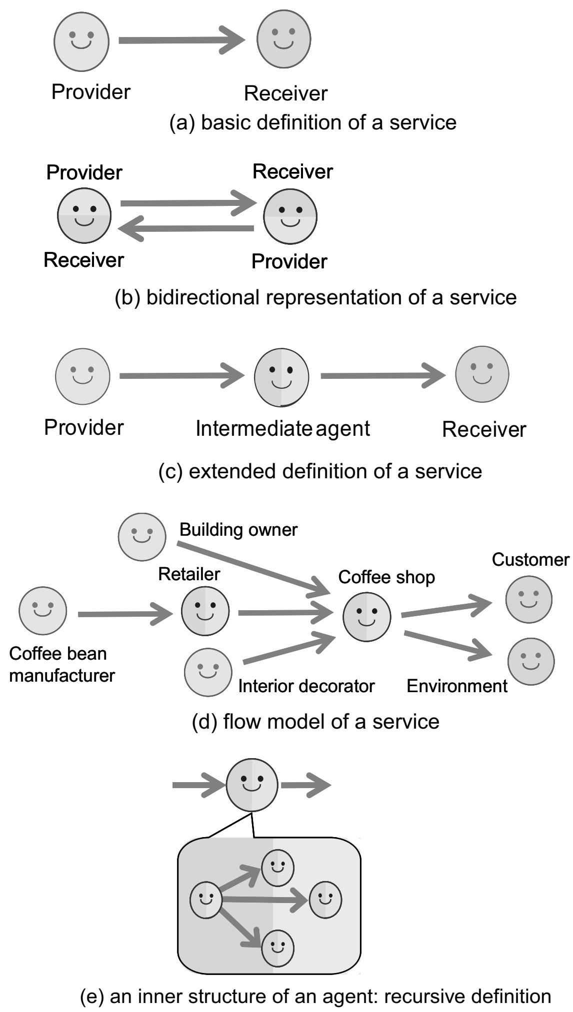


Figure 4: Extended definitions of a service

Conventional research on service emphasizes intangibility and simultaneity of service, which is perceived as an activity. Hence, development of service process representing dynamic aspects was a main topic in the early literature. The service blueprint [26] [27] and the service map [28] are the most famous techniques used by marketers to describe service activities sequentially and visually. The service blueprint is an effective technique to analyze and design the delivery of services on paper prior to their actual delivery.

To simulate service on the computer, we have been developing on integrating the function-attribute model and the service blueprinting scheme [29].

3.5 Service Explorer

A computer-aided design system called Service Explorer [11] [30] [31] has been developed since 2002. Service Explorer can represent the needs of customers and the relationship between those needs. Traditional CAD tools for mechanical product design cannot support to analyze customer needs. The advantages are listed up as follows: (a)common understanding to the process of service by designers, managers, market researchers and even consumers: simultaneous engineering on service can be achieved, (b)evaluation by different segmentation of consumers, (c)simulation of service interactions, and (d)redesign of services.

The concept included in SE may set up a kind of finite element methods in services; RSPs are resolved into small components and each component can be evaluated by simple logics. Service Explorer can visualize, analyze, and evaluate a service. It helps to design a new service. Service Explorer has been verified on several cases by various industries so far; two examples are a warehouse equipment supplier [32] and a hotel in the accommodation industry [33]. Verifications are now being conducted in Sweden, Germany, Denmark, and Japan.

Figure 5 shows the conceptual scheme behind Service Explorer [34]. Service Design Working Space provides functions usable for model building. Service model data is stored in the Service Case Base. Service Design Organizer has two kinds of modules: one involves controlling and supporting design activities; the other involves evaluating services and simulates the behaviours of services. Reasoning Engines help the designer to create new services using reasoning based on analogy. The main functions of Service Explorer are detailed below:

Model building

Designers can model services using Service Explorer's several editors. An instance is stored in Service Case Base. The data can be repeatedly used as basis for later service designs; for example, it is useful to store typical services or product-service cases presented in the literature (e.g. [35] [36] [37]) to use as building blocks for future service models.

Evaluation of services

The main purpose of Service Explorer is to express and to evaluate services. One of the typical characteristics on services is heterogeneity and subjectivity. As satisfaction depends strongly on personal experiences and expectation of each customer, we have introduced the concept of Persona, which is a fictitious and virtual character representing a

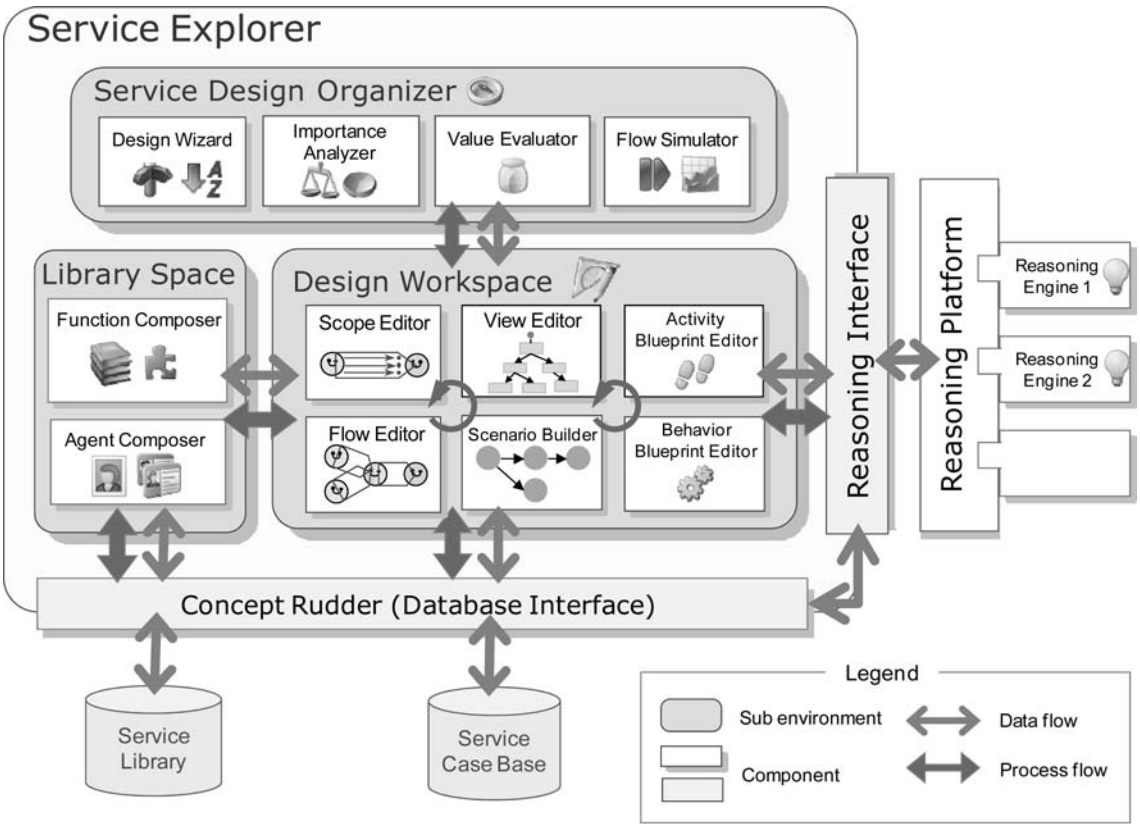


Figure 5: The conceptual scheme underlying Service Explorer

groupe of customers. Assuming the customer to be observable, two mechanisms have been implemented for the current version of Service Explorer:

(1) Importance analyzer: To determine the importance of each parameter, a designer weights each parameter using Quality Function Deployment (QFD) [38], Analytical Hierarchy Process (AHP) [39], and DEMATEL [40]. The latter two methods are empirically introduced because of their powerful quantification of subjective factors. An importance weight for each RSP is computed numerically according to AHP method using bilateral comparisons between the parameters. The importance of each RSP is decomposed into those of individual parameters at lower levels [41].

(2) Value evaluator: Functions of parameters discussed in Sec. 3.1 need to be chosen to compute customers' satisfaction. Several kinds of functions have been introduced: threshold functions and linear functions; then two concepts have been employed to reflect psychological behaviors of consumers: Kano model from product development and prospect theory from behavioral economics. The former defines attractive quality, one-dimensional quality, and must-be quality. The latter defines no-linearity according to expectation. The combined functions among these concepts are proposed as Satisfaction-Attribute functions [42], which obtain satisfaction by function and/or attribute parameters.

Mechanism for generating new ideas

Service Explorer allows users to take advantage of its computing power. The reasoning system in Service Explorer is an effective way for designers to solve service design problems. A reasoning system, Universal Abduction Studio (UAS) [43], is connected to Service Explorer to instantiate new candidates of services.

4 SCIENTIFIC APPROACH TO SERVICES

Service has been studied in comparison with the comparison with product in manufacturing. An assertion "service activities is of different nature from physical products" fascinates researchers to study charactersitics of service: intangibility, simultaneity, heterogeniety, and perishability. The authors, however, take the standpoint where physical products and service activities have almost the same characterstics in design. Therefore services should be describable as objectively as physical products in CAD and CAE systems. The method adopted in this paper is a technique for resolving service to elements as much as possible, namely, reductionism. However, an actual problem lies in what sets of elements can be used for services. Functions and attributes are used in accordance with General Design Theory by Yoshikawa. We proved that the functions and the attibutes of practical service can describe a set of service on the assumption that customer's logics is observable and controllable. Many other hypotheses exist in SE but are not discussed because of shortage of space.

Common language in manufacturing and marketing

A main purpose of the service design is to establish a common language for designers of different disciplines, e.g. market researchers, mechanical engineers and sales advisors, can share their understanding, so that they can accelerate concurrent engineering. Implemented according to engineering principles in consideration of marketing studies, SE can make the best use of concepts in service marketing; for instance, Lovelock's classification of service elements, core product and supplimentary service, is similar to that in our classification, service contents and service channel.

Is SE a scientific approach?

Finally, SE should be evaluated. As a result of case studies, the description in SE may deviate from objective description because of levels of designers, especially personal experiences of the designer and the vocabrary used for expression. Lexicons are standardized to describe Persona and activities in service description, and to select quantitative elements of functions and attributes. It is a newest study supported by METI. Four cases are studied in details and improved as service providers. The vocabrary lists are verified. In this way, authors conclude that SE can be recognized as a scientific approach in description and analysis of services. These lexicons may advance service engineering.

Future works

Our research on service engineering consists of three domains: modeling, analysis and design services as illustrated in Fig.6. Service CAD, which supports designers to design service on computer, has been developed by integrating these research achievements. In Fig. 6, some underlined requirements have been achived already, those in bold are currently in study.

5 CONCLUSION

In this paper, scientific representations on service are studied. The first in the proposals by a research group in The Univ. of Tokyo, *"(1) to measure, visualize, and quantitatively evaluate the quality of services"* seems achieved in representation. Quantitative evaluation will be ready if many cases of services are collected with the evaluation by many different individuals.

The technique taken in SE might be a top-down technique for defining service with many hypotheses. On the other hand, bottom-up approaches may be effective in extracting common characteristics from good services. Even in the latter cases, a novel description method can be implemented on the computer to describe the service activities as well as physical products scientifically.

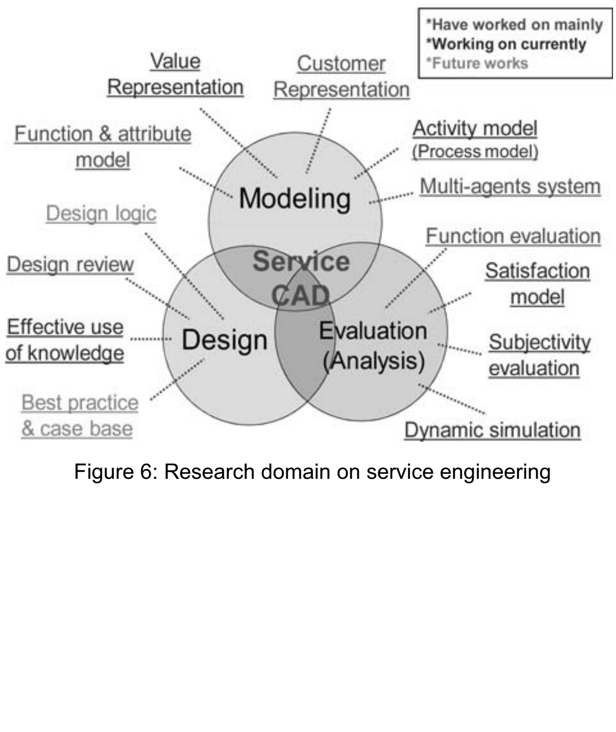


Figure 6: Research domain on service engineering

6 REFERENCES

[1] http://www.service-js.jp/ (in Japanese)
[2] http://www.ducr.u-tokyo.ac.jp/materials/200711service-innovation.pdf (in Japanese)
[3] Tukker, A., Tischner, U., 2006, New Business for Old Europe, Greenleaf Publishing
[4] Tukker, A., Tischner, U., 2006, Product-services as a research field: past, present and future. Reflections from a decade of research, J. of Engineering Design, 14(17), 1552-1556.
[5] McAloone, T.C. and Andreasen, M.M., 2004, Design for utility, sustainability and social virtues, developing product service systems. International Design Conf. - Design 2004, Dubrovnik, Croatia.
[6] Alonso-Rasgado, T.A., Thompson, G. and Elfström B. O., 2004, The Design of Functional (Total Care) Products. J. of Engineering Design, 15(6), 510-540.
[7] Lindahl, M. and Ölundh G., 2001, The Meaning of Functional Sales. Life Cycle Engineering: Challenges and Opportunities, 8th International Seminar on Life Cycle Engineering, Varna, Bulgaria, CIRP.
[8] Fujimoto, J., 2003, Development of Service-Oriented Products Based on the Inverse Manufacturing Concept. Environmental Science and Technology, 37(23), 5398-5406.
[9] Spohrer, J. and Maglio P., 2006, The Emergence of Service Science: Toward systematic service innovations to accelerate co-creation of value [online]. Available from http://www.almaden.ibm.com/asr/SSME/jspm.pdf
[10] Shimomura, Y. and Tomiyama, T., 2002, Service Modeling for Service Engineering. Proc. 5th Intl. Conf. on Design of Information Infrastructure Systems for Manufacturing 2002 -DIISM2002-, 309-316.
[11] Arai, T. and Shimomura, Y., 2004, Proposal of Service CAD System -A Tool for Service/Product Engineering-. Annals of the CIRP, 53(1), 397-400.
[12] Shimomura, Y. and Tomiyama, T., 2005, Service Modeling for Service Engineering. IFIP, Vol. 167, 31-38, Springer.
[13] Shimomura, Y., Sakao, T., Sundin, E. and Lindahl, M., 2007, A Design Process Model and a Computer Tool for Service Design. Proc. 12th Design for Manufacturing and the Life Cycle - DFMLC2007 -, CD-ROM, ASME.
[14] http://www.trl.ibm.com/news/SSS05/Spohrer_presentation.pdf
[15] Sasser, E. and Olsen, R.P. and Wyckoff, D., 1978, Management of service operations, Allyn and Bacon, Boston.
[16] Zeithamal, V.A., Parasuraman, A., and Berry, L., 1985, Problems and Strategies in Services Marketing, J. Marketing, 49 (Spring), 33-46.
[17] Fitzsimmons, J.A. and Fitzsimmons, M.J., 2001, Service management: Operations, strategy, and information technology. Third edition: McGraw-Hill
[18] Lovelock, C.H. and Wright, L.K., 1999, Principle of Service Marketing and Management, Prentice-Hall
[19] Hill, T.P., 1977, On Goods and Services. The Review of Income and Wealth, 23(4), 314-339.
[20] Vargo, S.L. and Lusch, R.F., 2004, Evolving New Dominant Logic for Marketing, J. Marketing, Vol. 68, 1-17.
[21] Yoshikawa, H., 1989, Design philosophy: the state of the art, CIRP Annals 38/2/1989, 579.
[22] Tomiyama, T., Umeda, Y. and Yoshikawa, H., 1993, A CAD for functional design, CIRP Annals 42/1/1993, p.143
[23] Umeda, Y., Ishii, M., Yoshioka, M., Shimomura, Y. and Tomiyama, T., 1996, Supporting Conceptual Design based on the Function-Behaviour-State Modeler. Artificial Intelligence for Engineering Design, Analysis and Manufacturing, 10(4), 275-288.
[24] Jensen, K., 1994, An Introduction to the Theoretical Aspects of Coloured Petri Nets, Lecture Notes in Computer Science Vol. 803, Springer-Verlag, 230-272.
[25] Tian, G., Duan, F. and Arai, T., 2006, Modular Design of Home Service Robot System with Hierarchical Coloured Petri Net, Proc. Intelligent Autonomous Systems 9, 542-549.
[26] Shostack, G. L., 1982, How to Design a Service, European Journal of Marketing, Vol.16, No.1, 49-63.
[27] Kingman-Brundage, J., 1989, The ABCs of Service System Blueprinting, in Designing a Winning Service Strategy, eds. M. Bitner and L. Crosby (AMA, Chicago) .
[28] Kingman-Brundage, J., 1991, Technology, Design and Service Quality. Intl. J. Service Industry Management, 2(3), 47-59.
[29] Hara, T., Arai, T. and Shimomura, Y., 2008, Integrating function model and activity model for design of service. In Proc. of the 41st CIRP Conf. on Manufacturing Systems, Tokyo, Japan, in-printing.
[30] Arai, T. and Shimomura, Y., 2005, Service CAD System -Evaluation and Quantification-. Annals of the CIRP, Vol. 54/1, (ISSN 1660-2773), 463-466.
[31] Hara, T., Arai, T., Shimomura, Y. and Sakao, T., 2007, Service/Product Engineering: a new discipline for value production. In Proc. the 19th Int. Conf. on Production Research (ICPR-19), Valparaiso, Chile, CD-ROM.
[32] Shimomura, Y., Sakao, T., Sundin, E. and Lindahl, M., 2007, A Design Process Model and a Computer Tool for Service Design. Proc. of the 12th Design for Manufacturing and the Life Cycle Conf. -DFMLC2007 -, ASME.
[33] Sakao, T. and Shimomura, Y., 2007, Service Engineering: A Novel Engineering Discipline for Producers to Increase Value Combining Service and Product. J. Cleaner Production, Special issue on sustainable production and consumption - making the connection, 15(6), 590-604.
[34] Hara, T., Arai, T. and Shimomura, Y., 2006, A Concept of Service Engineering: A Modeling Method and a Tool for Service Design. Proc. of IEEE International Conf. Service Systems and Service Management 2006 (SC SSSM 2006), 13-17, Troyes, France.
[35] Lovelock, H.C., 1983, Classifying Services to Gain Strategic Marketing Insights. Journal of Marketing, 47 (Summer), 9-20.
[36] Tukker, A., 2004, Eight Types of Product-Service System: Eight Ways to Sustainability? Experiences from Suspronet. Business Strategy and the Environment, 13, 246 - 260.
[37] Williams, A., 2007, Product service systems in the automobile industry: contribution to system innovation?. Journal of Cleaner Production, 15, 1093-1103.
[38] Akao, Y., 1990, Quality Function Deployment, Productivity Press, Cambridge, Massachusetts.
[39] Satty, T.L., 1980, The Analytic Hierarchy Process: McGraw-Hill.
[40] Manassero, G., Semeraro Q. and Tolio T., 2004, A New Method to Cope with Decision Makers' Uncertainty in the Equipment Selection Process, Annals of the CIRP, 53(1), 389-392.
[41] Shimomura, Y. and Sakao, T., 2007, A Service Evaluation Method for Service/Product Engineering, 16th Intl. Conf. on Engineering Design -ICED07-, France.
[42] Yoshimitsu, Y., Kimita, K., Hara, T., Shimomura, Y. and Arai, T., 2007, Proposal of a Measuring Method of Customer's Attention and Satisfaction on Services. the 14th CIRP International Conf. on Life Cycle Engineering -LCE2007-, Tokyo, Japan, CIRP, 417-422.
[43] Takeda, H., Sakai, H., Nomaguchi, Y., Yoshioka, M., Shimomura, Y. and Tomiyama, T., 2003, Universal Abduction Studio -Proposal of A Design Support Environment for Creative Thinking in Design-. Proc. 14th Intl. Conf. on Engineering Design -ICED03-, Stockholm.

Manufacturing System Evaluation

A General Economic Model for Manufacturing Cost Simulation

Mathias Jönsson, Carin Andersson, Jan-Eric Ståhl
Division of Production and Materials Engineering, Lund University, Sweden

Abstract
The described technical-economic model clarifies the influence of different production technological factors on the processing cost of a part. Influential factors can be weighted against each other, which leads to different production development scenarios and their effects on the processing cost can be studied. This implies a way to generate a basis of decision by which a company can base their production related development goals. The model describes influence of technical factors on the manufacturing cost and thereby represents the important link between technical development and economy.

Keywords:
Manufacturing Economy; Cost Model; Deterministic Production Development

1 INTRODUCTION

A majority of all manufacturing companies are working with production development and improvements to meet the global competition of today. There are a number of methods and philosophies for working with continuous improvements, where the success of lean production [1], is the most widely spread. An important question is if considerations and decisions made regarding investments and development actions are based on correct and adequate knowledge in order to achieve the highest efficiency benefits.

The outsourcing debate has been going on for some time now [2], [3]. Decisions made about moving production-units to low-wage countries are often based on limited information, giving wages too big influence over the decisions. Existing economic models are inadequate in utilizing estimation of the development potential of a production system and possible development actions.

There are many questions to be asked when considering major improvement changes in a production facility. The most common questions the company management would like to have answered are:

- How much better do we have to be to compete with for example low-wage countries and what and where in the production facility do we have to improve?
- What are the bases of decision required to formulate goals for production development, and what are reasonable goals for the actual production system?

The economic model presented in this paper can help to give answers to these questions, if the required data about the production performance is known.

2 PURPOSE AND LIMITATIONS

The purpose of this economic model is to describe the costs added to the cost of a part at every processing step. The model is not intended to be used only to describe the present cost situation, but also to function as a simulation tool to simulate different development scenarios and their effect on the part cost. Thereby it can be used as a support tool in manufacturing development activities.

The economic model presented is defined to comprise the direct production cost. The overhead costs are excluded at this level, because they have little to do with developing the production system. Factors tied to the income side of the production are not considered in the model. The model primarily describes batch production and is summarised to describe one processing step or a so-called planning point (a planning point is a set of machines and robots where the cycle time is determined by the slowest machine in the line). This simplification enhances the principle of comparing the influence of different cost items on the total production cost. These factors influence on the production cost can therefore constitute the foundation for choosing research and development actions.

3 LIST OF SYMBOLS

The parameters in the list of symbols are partly tied to the factor groups described above. The economic parameters are described in the Swedish currency krona (kr).

Table 1: List of symbols

t_0	Nominal cycle time per part.	min
t_m	Machine time	min
t_h	Handling time	min
t_{vb}	Tool switch time	min
t_p	Production time per part	min
t_S	Average down time per part	min
q_S	Down time rate	-
N_0	Nominal batch size	unit

N	Total amount of required parts to be able to produce N_0 parts	unit
N_Q	Amount of scrap parts in a batch of N_0 parts	unit
q_Q	Scrap rate	-
t_{0v}	Cycle time including production rate losses	min
q_P	Production rate	-
T_{su}	Set up time of a batch	min
T_{pb}	Production time of a batch	min
t_{pb}	Production time per part of a batch with N_0 parts	min
k	Part cost	kr/unit
k_B	Material cost per part including material waste	kr/unit
k_{CP}	Hourly cost of machines during production	kr/h
k_{CS}	Hourly cost of machines during down time and set up	kr/h
k_D	Wage cost	kr/h
K_B	Material cost of a batch including scrapped parts and material waste	kr/batch
k_{B0}	Material cost of the manufactured part without material waste	kr/unit
q_B	Material waste factor	-
m_{tot}	Total consumption of material per part	weight
m_{part}	Remaining material in the machined part	weight
U_{RP}	Degree of occupation	-
n	Rational number > 0	-
∂z	Partial change in arbitrary variable.	-
Δz	Change in arbitrary variable	-
x_p	Process development factor for the cycle time	-
x_{su}	Process development factor for set up time	-

4 LITTERATURE REVIEW

Several different models have been developed for the purpose of calculating the manufacturing cost. According to Tipnis, *et al*. [5], the models can be divided in microeconomic and macroeconomic models. In the microeconomic models specific process parameters influence on the part cost is described. Microeconomic models dealing with machining has been described by Colding [6], [7] and Alberti, *et al*. [8] among others and Knight, *et al*. [9] has developed a corresponding model for forging. Within the field of machining a microeconomic model can describe how for example the cutting rate, feed or working margin influence the part cost. In a macroeconomic model several parameters are aggregated. An example of a macroeconomic model is when the cost calculations are based on the cycle time and not the factors influencing the cycle time. The fundamental principles for developing macroeconomic models are described by Kaplan and Anderson [10]. The authors have not developed any models that are directly applicable to calculate the part cost but leaving these activities to the reader.

Macroeconomic models have previously been illustrated by Groover [11]. In this model only one production loss parameter is taken into consideration; the scrap rate. Ravignani and Semeraro [12] have developed a model that combines the micro- and macroeconomic views by noticing both cutting technological conditions and the batch size. Non production loss parameters are regarded.

It can be stated that the microeconomic models are specific for different processing methods. Numerous models have been developed to describe the cutting cost of machining. The models are describing the connection between the cutting rate, the wear rate of a cutting tool and the tool switch time. In these models the tool cost is highly prioritized. Costs of down time and the scrap rate are not often taken into consideration.

A cost model for assembly is introduced by Teng and Garimella [13]. This model is based on inventory costs, assembly costs and costs associated with diagnostic and rework activities. The model has a high resolution concerning cost of different types of equipment in the assembly line. The model is based on average cycle times where also the scrap rate is considered. Boothroyd [14] is describing a specific cost model for robot assembly which is noticing the down time costs in the assembly line.

Production cost regarding design has been discussed by Locascio [15], Liebers and Kals [16] and Shehab and Abdalla [17]. Locascio is assuming that all cycle times of the processing steps is known in advanced. Any specific connection to production loss parameters is not considered. Shehab and Abdalla is describing an interesting model that estimate the manufacturing cost of machining for different choices of material where both the material cost and the processing cost is taken into consideration.

The model described below is general and can be regarded as a macroeconomic model but with the possibility to consider the microeconomic parameters. The model is intended to describe the part cost of various specific or aggregated processing methods without any major modifications.

5 MODELLING OF THE PART COST

The nominal processing time (cycle time) t_0 for a part is comprised of machine time, handling time and tool change time:

$$t_0 = t_m + t_h + t_{vb} \tag{1}$$

The equation assumes that the events are performed in a sequential order and can be considered as a planning point. The real processing time t_p will be longer than the nominal time due to disturbances and downtime. The rate of the disturbance and downtime can be expressed as the quotient between the downtime t_S and the observed production time t_p described in equation 2. The sum of the downtime and nominal processing time gives the real processing time t_p according to equation 3. Combining equation 2 and 3 the processing time can be determined based on the nominal cycle time and the downtime rate q_S:

$$q_S = \frac{t_S}{t_p} = \frac{t_p - t_0}{t_p} \quad (2)$$

$$t_p = \frac{t_0}{1-q_S} = t_0(1+\frac{q_S}{1-q_S}) \quad (3)$$

To obtain N_0 number of correct parts, N number of parts has to be manufactured due to scrapped parts. The rate of scrapped parts is expressed by q_Q:

$$q_Q = \frac{N_Q}{N} = \frac{N - N_0}{N} \quad (4)$$

$$N = \frac{N_0}{1-q_Q} = N_0\left(1+\frac{q_Q}{1-q_Q}\right) \quad (5)$$

Losses in production rate are a fact when the cycle time has to be increased from t_0 to t_{0v} to maintain the quality level or avoid unplanned downtime. The relative loss in production rate is described as:

$$q_P = \frac{t_{0v} - t_0}{t_{0v}} \quad (6)$$

$$t_{0v} = \frac{t_0}{1-q_P} \quad (7)$$

To changeover the production from manufacturing part A to part B a certain amount of setup time T_{su} is required. The production time for a batch including the setup time is:

$$T_{pb} = T_{su} + N \cdot t_p = T_{su} + \frac{N_0 \cdot t_0}{(1-q_Q)(1-q_S)(1-q_P)} \quad (8)$$

The average production time for a batch of N_0 number of correct parts is calculated as:

$$t_{pb} = \frac{T_p}{N_0} \quad (9)$$

In the presented model there are primarily three cost items specified; equipment costs k_C, wage costs k_D and material costs k_B. Equipment costs for a machine or a production line can be split up into a cost during production k_{CP} and a cost k_{CS} when the machine or production is not running. For the case in question both these cost items include all of the costs that can be related to the equipment as investment cost, local cost, cost of maintenance, tool costs etc. The cost of wages per hour k_D are presumed to be independent of if the machine is running or not and also presumed to be unchanged during setup.

To study the material cost including scrapped parts and material waste, a material waste factor q_B is introduced:

$$K_B = \frac{N_0 \cdot k_{B0}}{(1-q_B)(1-q_Q)} \quad (10)$$

$$q_B = \frac{m_{tot} - m_{part}}{m_{tot}} \quad (11)$$

where k_{B0} is the material cost of the manufactured part and K_B is the material cost of the batch including scrapped parts and material waste. The material waste factor q_B consider the total consumption of material m_{tot} per part and comprises also material that are machined or cut off as for example chips during turning or milling and retainer surfaces during sheet metal forming. The remaining material in the machined part is denoted m_{part}.

Reduced occupation in a manufacturing system leads to consequences for all manufactured parts. This situation can be considered in different ways, hence the free production resource can be considered both as an economic asset and a disadvantage depending on the situation. In a long term view the manufactured parts must carry the costs for the over capacity. The over capacity time can be distributed over all the batches in relation to their production time T_{pb} by introducing a degree of occupation U_{RP}, calculated as the quotient between real production time T_{prod} and planned production time T_{plan}:

$$U_{RP} = \frac{T_{prod}}{T_{plan}};\ T_{plan} = T_{prod} + T_{free} \quad (12)$$

The extra free capacity $T_{free,b}$ to be added to a specific batch is calculated according to equation 13. The free time can be considered as a setup time at the same time as the equipment is available for manufacturing:

$$T_{free,b} = \frac{1-U_{RP}}{U_{RP}} T_{pb} \quad (13)$$

The manufacturing costs per part k, including the previously described parameters and assumptions can be expressed as:

$$\begin{aligned} k = {} & \frac{K_{sum}}{N_0} + \left(\frac{k_B N_0}{N_0(1-q_Q)(1-q_B)}\right) + \\ & \left(\frac{k_{CP}}{60N_0} \cdot \frac{t_0 N_0}{(1-q_Q)(1-q_P)}\right) + \\ & \frac{k_{CS}}{60N_0}\left(\frac{t_0 N_0}{(1-q_Q)(1-q_P)} \cdot \frac{q_S}{(1-q_S)} + T_{su} + \frac{1-U_{RP}}{U_{RP}} T_{pb}\right) + \\ & \frac{k_D}{60N_0}\left(\frac{t_0 N_0}{(1-q_Q)(1-q_S)(1-q_P)} + T_{su} + \frac{1-U_{RP}}{U_{RP}} T_{pb}\right) \end{aligned} \quad (14)$$

In some cases it can be necessary to introduce a disturbance factor q_{Ssu} to handle spreading in the nominal setup time.

The cost item K_{sum} in equation 14 comprises different types of costs that are not described separately in the model. A more complete economic model has a higher resolution and includes more of the separate terms that are now included in K_{sum}. A developed model can for example consider tool costs, cost of maintenance, remainder value of waste material, fixture costs, stock/buffer and transportation costs, surrounding equipment, costs arising due to environmental or recycling actions for example to eliminate cutting fluids or oils.

6 DETERMINISTIC PRODUCTION DEVELOPMENT

To be able to manage production development efficiently, clear goals has to be established for the development activities. Many companies today have implemented lean manufacturing to some degree, or they are by other methods developing and improving the manufacturing process. With this model those activities can be performed in a more deterministic, goal oriented way. The reasons for this is that an implementation of this model for every product in every processing step, enables the most critical factors from a cost perspective to be acquired. When you have this information it will be possible to establish concrete economic goals and to simulate the consequences these goals have on the parameters constituting the part cost. The consequences could for example be how much a given parameter must be changed to reach the established goal.

The development activities can be performed in relation to the present production conditions of the company or in relation to the competitors and other terms of the market. Example of production development goals are reduction of the manufacturing costs with 20% for a certain part type, a 50% reduction of setup time or an increase of production rate from 100 to 120 parts per week with unchanged cost parameters.

Considering that a lot of factors, isolated or in cooperation, influence the cost of a specific part, different changes in these factor can lead to same cost effects. To be able to separate the influence of these different factors on the part cost, different development factors are introduced to the parameters in equation 14.

$$k = \frac{K_{sum}}{N_0} + \left(\frac{k_B \cdot N_0}{N_0(1-q_Q)(1-q_B)} \right) + \left(\frac{\kappa_C \cdot k_{CP}}{N_0 60} \cdot \frac{x_p \cdot t_0 \cdot N_0}{(1-q_Q)(1-q_P)} \right) + \frac{\kappa_C \cdot k_{CS}}{60 N_0} \left(\frac{x_p \cdot t_0 \cdot N_0}{(1-q_Q)} \cdot \frac{q_S}{1-q_S} + x_{su} \cdot T_{su} + \frac{1-U_{RP}}{U_{RP}} T_{pb} \right) + \frac{k_D}{60 N_0} \left(\frac{x_p \cdot t_0 N_0}{(1-q_Q)(1-q_S)} + x_{su} \cdot T_{su} + \frac{1-U_{RP}}{U_{RP}} T_{pb} \right) \quad (15)$$

In equation 15 the development factor x_p operates on the cycle time and enables therefore analysis of changes in cycle time. The development factor x_{su} operates on the setup time and enables therefore studies of changes in setup time. The cycle time and setup time are the most important parameters describing the capacity and flexibility of a production system. A development factor given a value less than 1.0 result in a reduction in cycle time and setup time, if the factors are given for example the value 0.5, the production time and setup time has been reduced to half of the original size. The development factors can therefore be regarded as improvement variables in a goal function.

A cost development factor κ_C is introduced to describe an investment cost that can be connected to a change in cycle time. The cost factor operates on the equipment costs k_{CP} and k_{CS}. This factor is used to model changes in costs in primarily existing equipment, and can be used to determine the limit of investment justified by for example a decrease of the downtime rate to a certain value. For example does κ_C = 1.20 corresponds to an increase in equipment cost with 20%.

7 COST DERIVATIVES

Changes in part cost caused by a limited change in an arbitrary variable *z,* is calculated by partial derivative, and is described in linear form as:

$$\Delta k_i = \frac{\partial k_n}{\partial z} \cdot \Delta z \quad (16)$$

The changes in part costs can be calculated with respect to different parameters as for example changes in wage costs and share of downtime Δq_{Si}. Equation 17 is exemplifying changes in part costs due to changes in different governing parameters.

$$\Delta k_i = \frac{\partial k_i}{\partial k_{Di}} \cdot \Delta k_{Di} + \frac{\partial k_i}{\partial q_{Si}} \cdot \Delta q_{Si} \quad (17)$$

Cost neutral changes in each variable can be studied by putting the change in part costs $\Delta k_i = 0$. Equation 17 is written in a cost neutral form in equation 18, describing the size of the reduction in downtime share required to compensate for a change in wage costs.

$$\Delta q_{Si} = -\Delta k_{Di} \cdot \frac{\frac{\partial k_i}{\partial k_{Di}}}{\frac{\partial k_i}{\partial q_{Si}}} \quad (18)$$

The influence of a specific variable can be studied by calculating cost derivatives. A change in a variable giving a large influence on the part cost also gives large cost derivative values. It is hazardous to uncritically compare different cost derivatives with each other since the possibility of changing each variable is different. A weighting of the cost derivative can be made by multiplying the cost derivative with its functional value. A weighted cost derivative is a better indication of the impact each variable has on changes in the cost derivatives. All changes Δz in the variable *z* becomes relative with respect to the absolute value of the variable. By introducing a relative variable $\Delta z_0/z_0$, the changes expressed as a percentage for a specific variable can be compared with changes expressed as a percentage for another variable. This principal is expressed in equation 19.

$$\Delta k = \frac{\partial k(z_0)}{\partial z} \cdot z_0 \cdot \frac{\Delta z_0}{z_0} \quad (19)$$

8 MODEL EXAMPLE

In the present section the usefulness of the model will be shown by implementing the model using fictive input data. The example will illustrate what kind of analyses that could be performed and what decision-making bases you can get.

The costs for two different production cases can be studied by introducing an index *i* tied to the parameters and variables in equation 16 in order to separate them. In the following examples the part costs k_1 and k_2 are calculated for the presumption valid for each case. Below the developed model is exemplified by inserting technical and economic data according to Table 2.

Table 2: Applied data for the model.

t_0	10	min
T_{su}	100	min
k_{Cp}	1000	kr/h
k_{CS}	700	kr/h
k_{D1}	200	kr/h
k_{D2}	50	kr/h
k_B	20	kr/part
q_B	0	-
K_{sum}	0	kr/batch

Figure 1 illustrates the part cost k as a function of the nominal batch size N_0 with two separate values of the wage cost. In case 1 (dotted graph) is the wage cost unchanged, i.e. k_{D1} = 200 kr/h and in case 2 (continous graph) is the wage cost reduced; k_{D2} = 50 kr/h. The difference in wage cost can for example illustrate two manufacturing plants in different countries with different wage costs. All other parameters are unchanged. In the figure you can see that the value of k is clearly higher for case 1 because of the higher wage cost.

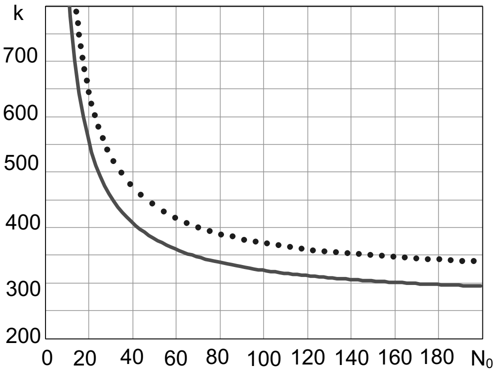


Figure 1: The part cost of the productio ıses 1 (dotted graph) and 2 (continuous graph) as a function of the nominal batch size N_0. In both case 1 and 2 is q_Q = 5 % and q_S = 40 %, x_p and x_{su} is 1.0.

For the plant in case 1 to be able to compete with the plant in case 2 it must take actions to alter one or more of the parameters building up the cost k. In Figure 2 has the plant in case 1 managed to decrease the down time losses from q_S = 40 % to q_S = 35 % and the process development factor x_{p1} has decreased from 1.0 to 0.95. By these changes the difference in part cost between the two cases has more than halved, even if it differ a factor 4 in wage cost.

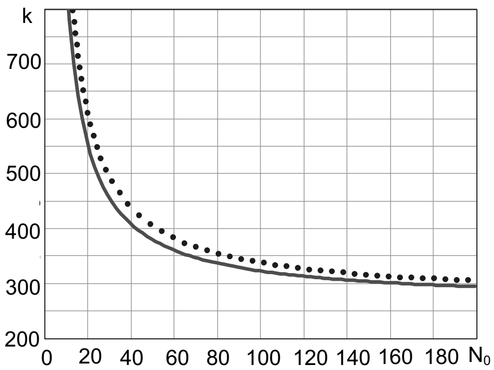


Figure 2: The part cost of the production cases 1 (dotted graph) and 2 (continuous graph) as a function of the nominal batch size N_0. In both case 1 and 2 is q_Q = 5 % and q_S = 40 % for case 2 and 35 % in case 1, x_p is 1.0 in case 2 and 0.95 in case 1.

In Figure 3 below, the down time factor q_{S1} has been further decreased with 5 % to 30 % and the process development factor x_{p1} is reduced to 0.80. In this situation the part cost for the plant in case 1 has been reduced and becoming 30 kr lower than the plant in case 2 for batches larger than 100 parts.

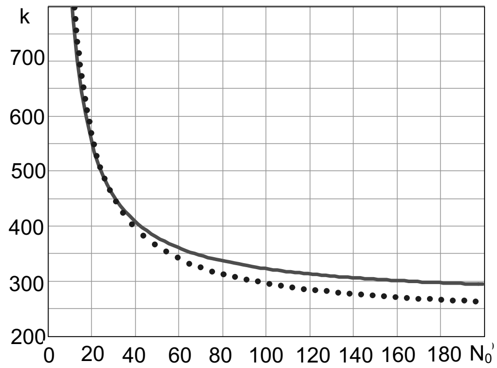


Figure 3: The part cost of production cas dotted graph) and 2 (continuous graph) as a function of the nominal batch size N_0. In both case 1 and 2 is q_Q = 5 %. q_S = 40 % for case 2 and 30 % in case 1, x_p is 1.0 in case 2 and 0.80 in case 1.

In Figure 4 the cost derivative is exemplified. The cost change Δk_1 is illustrated as a function of change in down time loss Δq_{S1} and change in relative wage cost $\Delta(k_{D1}/k_{D1})$. In the figure you can observe that a increase in part cost by 40 kr can either be received by increasing the down time loss 10 % or the wage cost by 70 %. In this linear model the corresponding decrease applies in the described variables.

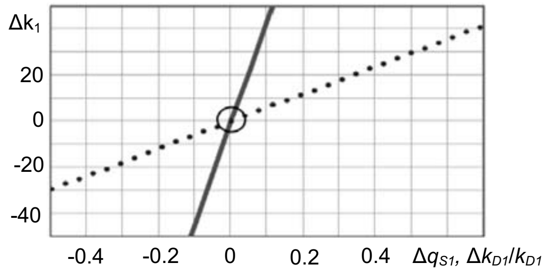


Figure 4: The relationships between the cost change Δk_1 and change in down time loss Δq_{S1} and change in relative wage cost $\Delta k_{D1}/k_{D1}$. In this case is q_Q = 5 % and q_S = 40 %, x_p and x_{su} is 1.0, the batch size 200 parts.

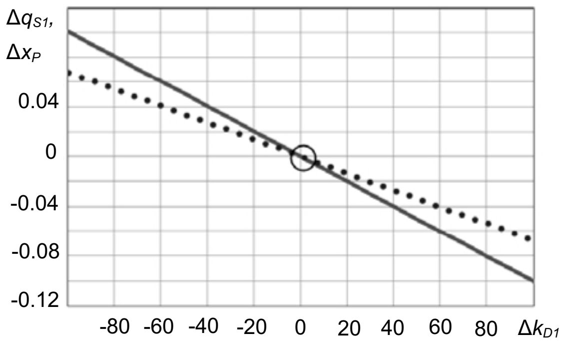


Figure 5: Cost neutral changes in wage cost and in down time losses (dotted graph) and also in wage cost and process development factor (continuous graph). In this case is q_Q = 5 % and q_S 40 %, x_p and x_{su} is 1.0 and batch size N_0 = 200 parts.

In Figure 5 the cost neutral changes are shown, which illustrate the balance for a change in wage cost and down time loss and also a change in wage cost and process development factor. In the figure it can be established that a wage increase by 40 kr per hour i.e. 20 %, corresponds a cost neutral improvement i the process development factor

Δx_p by about 4 % or a decrease in down time losses Δq_S by almost 3 %.

9 DISCUSSION AND CONCLUSIONS

The developed model enables analyses and economic estimations of various technical and organisational development alternatives. The model example shown in the section above illustrates for example how a higher wage cost can be compensated by technical and organizational improvements. Through studies of cost derivatives different alternatives related to production development can be judged. High cost derivatives shows the strength of a certain variable. The investment cost in research and development necessary to reduce x_{p1} from 1.0 to 0.80 and q_{S1} from 0.40 to 0.30 can for instance be weighted against alternative costs. The theoretical and practical possibilities to realize the necessary development for example in the case above must of course be estimated in each specific case. The conditions are highly governed by the present level of development and the belonged remaining development potential.

The difficulties of using the described model are that the model demands accurate input data. A systematic registration of the disturbances building up the parameters q_Q, q_S and q_P and parameters such as the set up time is of great importance. From experience the equipment costs represent though the greatest difficulties. These problems are dealt with by Ståhl (2007) among others.

10 REFERENCES

[1] Womack J.P., Jones, D.T., Roos D., 1990, The machine that changed the world, Rawson Associates, New York.

[2] Bengtsson, C., Berggren, L., Lind J., 2005, Alternativ till outsourcing, Liber AB, Malmö.

[3] Larsson, J., Malmqvist, C.-G., 2002, Outsourcing – erfarenheter av outsourcing i svenska företag*,* Olle Sjöstedt information AB. Västerås

[4] Ståhl, J.-E., 2007, Industriella Tillverkningssystem*,* Division of Production and Materials Engineering, Lund University, Lund.

[5] Tipnis, V.A., Mantel, S.J., Ravignani G.J., 1981, Sensitivity Analysis for Macroeconomic and Microeconomic Models of New Manufacturing Processes, Annals of the CIRP, 30:401-404.

[6] Colding, B., 1978, Relative Effects of Shop Variables on Manufacturing Cost and Performance, Annals of the CIRP, 27:453-458.

[7] Colding, B., 2003, Tids- och kostnadsanalys vid skärande bearbetning – Optimering för synkront detaljflöde, Colding International Corporation, Stockholm.

[8] Alberti, N., Noto La Diega, S., Passannanti, A., 1985, Interdependence Between Tool Fracture and Wear, Annals of the CIRP, 34:61-63.

[9] Knight, W.A., Poli, C.R., 1982, Design for Economical Use of Forging: Indication of General Relative Forging Costs, Annals of the CIRP, 31:159-163.

[10] Kaplan, R.S., Anderson, S.R., 2007, Time-Driven Activity-Based Costing; a simpler and more powerful path to higher profits, Harvard Business School Publishing Corporation, Boston.

[11] Groover, M.P., 1987, Automation, Production Systems, and Computer Integrated Manufacturing, Prentice-Hall Inc, Englewood Cliffs, NJ.

[12] Ravignani, G.L., Semeraro, Q., 1980, Economics of Combined Lot and Job Production with Consideration for Process Variables, Annals of the CIRP, 29:325-328.

[13] Teng, S.-H., Garimella, S.S., 1998, Manufacturing Cost Modelling in Printed Wiring Board Assembly, Journal of Manufacturing Systems, 17:87-96.

[14] Boothroyd, G., 1984, Economics of General-Purpose Assembly Robots, Annals of the CIRP, 33:287-290.

[15] Locasio, A., 2000, Manufacturing Cost Modelling for Product Design, The International Journal of Flexible Manufacturing Systems, 12:207-217.

[16] Liebers, A., Kals, H.J.J., 1997, Cost Decision Support in Product Design, Annals of the CIRP, 46:107-112.

[17] Shehab, E.M., Abdalla, H.S., 2001, Manufacturing cost modeling for concurrent product development, Robotics and Computer Integrated Manufacturing, 17:341-353.

Implementation of an Economic Model to Simulate Manufacturing Costs

Mathias Jönsson, Carin Andersson, Jan-Eric Ståhl
Division of Production and Materials Engineering, Lund University, Sweden

Abstract
An economic model describing manufacturing costs is implemented within the frame of a case study. The implemented economic model is developed to enable analyses of the cost items and parameters influencing the cost of a part or a batch and also to make simulations for the purpose of investigating the economic outcome of future development activities. The aim of the case study was to identify activities in a production unit that could lead lower manufacturing costs by using the method described in this paper.

Keywords:
Manufacturing Economy; Cost Model; Economic Simulation

1 INTRODUCTION

The globalisation has influenced the manufacturing sector considerably; it has become more important than ever to constantly increase the productivity in the manufacturing plants to be able to meet the increasing competition, especially from companies in low-wage countries. Many companies deal with these circumstances by offshoring or by offshore outsourcing. These types of relocations are solely based on cost reductions, in contrast to relocations based on market aspects for the purposes of getting closer to a certain market for competitive reasons. But to relocate the manufacturing plants to low-wage countries doesn't have to be the only way out to maintain a high competitiveness for companies in countries with high wages. Focusing on the optimal production development regarding organizational issues and production technology can in many cases compensate the higher level of wage costs.

In order to make the right decisions concerning the offshoring based on cost reductions, it is necessary to be able to make correct analyses of the performance in the manufacturing plant from an economic point of view. One approach is to analyse the manufacturing costs of the products produced in the plant. Common methods to calculate the manufacturing costs are the traditional full costing methods and Activity-Based Costing (ABC). Traditional full costing is mostly used for cost-price calculations and is not an equally suitable method to use when seeking detailed and accurate information solely about the manufacturing costs, essentially because of the methods volume based approach. One of the main purposes behind the development of ABC was to in a more accurate way than traditional full costing allocate the overhead costs [1]. A lot of research has been done over the years since ABC was first introduced with the intention of implementing and analyzing the method, for example Thyssen, *et al*. [2].

The model presented in this paper is developed with the purpose of calculating and analysing the part cost associated with the manufacturing. The model has some similarities to ABC, for example when it comes to recognising the cost of unused capacity and the different types of parts individual consumption of the manufacturing resources. The main differences are that the model presented here only describes costs related to manufacturing. The focus of this model is to describe the relations between economy and manufacturing performance. Another difference is that this model allocates batch level activities to the unit level and also allocates cost of unused capacity to products.

Models describing the manufacturing cost can roughly be divided into a micro- and macroeconomic approach according to Tipnis, *et al*. [3]. The macroeconomic models are essentially based on aggregated information while the microeconomic models include process data like cut rate, gas flow, current intensity etc. For example Colding [4], [5], Alberti, *et al*. [6] and Knight and Poli [7] have all described microeconomic models while Groover [8] has described a macroeconomic model.

The model applied below can be described as a macroeconomic model. The applied model differs from the models presented in the previous section by the inclusion of all the production loss parameters; scrap rate, down time rate and production rate.

2 PURPOSE

The performance of a defined production unit that manufactures gearwheels will be analysed, using a manufacturing economic model described in section 6. The purpose was to identify activities in the production unit that can lead to lower manufacturing costs by using the method described in this paper.

3 METHOD

To be able to test the model in real a context a case study at a company was considered as the most appropriate choice of method. The case study was chosen to be limited to one section at the factory in order to make the data collection manageable from a practical perspective and it was also considered sufficient in order to to comply with the purpose of the study. The case study began with a careful study of the chosen production unit to get an understanding about its characteristics regarding the process and function. The accuracy of the economic data obtained by the cost model is dependent on the accuracy of the collected data. Therefore a systematic data collection was made regarding the scrapped parts, down time, production rate and the set up time.

4 SYSTEMATIC PRODUCTION ANALYSIS

The systematic data collection mentioned in section 3 was performed by implementing a method called Systematic Production Analysis (SPA) [9]. The method has been developed to determine the existing production condition. In this method the result parameters downtime rate, scrap rate and production rate are measured for each processing unit involved in the manufacturing of a specific product. The possible downtime, scrapped parts and loss in production rate are related to a factor found in one of the following factor groups: A Tool and tooling system; B Work piece material; C Manufacturing process and process data; D Personnel, organization and outer logistics; E Maintenance and wear tied to A, C, D and G; F Special process behavior/factors; G Surrounding equipment and inner logistics; H Unknown or unspecified factors

Table 1 shows a method of presenting a SPA. Q_1 to Q_n describe different quality deviations leading to scraped parts, where every Q has a separate column. Analogous to the quality parameters, S_1 to S_n describe different types of down time losses and P_1 to P_n describe different production rate deviations. The factor groups describe causes leading to the different result parameters. Every factor group contain individual factors, for example factors A_1 to A_n, where every individual factor has a separate row. After an implantation of this method, where every disturbance has been registered in the right place in the table, you can sum up the result for every row and column and then find critical result parameters and factors for the specific processing unit. This method makes it possible to directly get an indication towards which of the result parameters and which of the individual factors that causes losses in the production efficiency. Coupling this to economic parameters it is possible to determine the part cost under the influence of the result parameters. In section 6 it will be shown how these result parameters together with time and batch size parameters and also economic data build up the part cost in a specific processing step.

Table 1: Systematic production analysis.

	Result parameters			
Factor groups	$Q_1,..,Q_n$ (unit)	$S_1,...,S_n$ (min)	$P_1,...,P_n$ (min)	∑
$A_1,...,A_n$				
$B_1,...,B_n$				
$C_1,...,C_n$				
$D_1,...,D_n$				
$E_1,...,E_n$				
$F_1,...,F_n$				
$G_1,...,G_n$				
H				
∑				

5 LIST OF SYMBOLS

Table 2: List of symbols used in this paper. The economic parameters is described in the Swedish currency krona (kr).

Parameter	Description	Unit
t_0	Nominal cycle time per part	min
t_m	Machine time	min
t_h	Handling time	min
t_{vb}	Tool switch time	min
N_Q	Amount of scrap parts in a batch of *N* parts	unit
N	Total batch size, including scrap parts	unit
N_0	Amount of correct produced parts in a batch	unit
q_Q	Scrap rate	-
t_p	Production time per part	min
t_S	Average down time per part	min
q_S	Down time rate	-
q_P	Production rate	-
t_{0v}	Cycle time including production rate losses	min
T_{su0}	Nominal set up time	min
T_{su}	Set up time including deviations from nominal set up time	min
q_{Ssu}	Ratio between the nominal set up time T_{su0} and the real set up time T_{su}	-
k	Part cost	kr/unit
k_A	Tool cost	kr/unit
k_B	Material cost	kr/unit
k_{CP}	Equipment cost during production	kr/h
k_{CS}	Equipment cost during downtime and set up	kr/h
k_D	Wage cost	kr/h
x_p	Process development factor for the cycle time	-

x_{su}	Process development factor for set up time	-
κ_c	Equipment cost development factor	-
Δz	Change in an arbitrary variable z	-
k_{VA}	Value added part of the part cost	kr
k_{NVA}	Cost of non value added activities	kr
k_{Tsu}	Costs related to set up	kr
k_Q	Costs connected to the scrap rate	kr
k_S	Costs related to the down time rate	kr
k_P	Costs connected to the production rate	kr

6 ECONOMIC MODEL

The economic model is priviously described in [9]. In this section a summerized description is presented.

The nominal cycle time t_0 in a machine or a line is defined in equation 1, t_m is the machine time, t_h handling time and t_{vb} tool change time.

$$t_0 = t_m + t_h + t_{vb} \tag{1}$$

The scrap rate q_Q is defined in equation 2, where N_Q is the number of scrap parts, N is the batch size and N_0 the number of correct, non scrapped parts of the batch.

$$q_Q = \frac{N_Q}{N} = \frac{N - N_0}{N} \tag{2}$$

The down time rate q_S is defined in equation 3, where t_s is the down time per cycle and t_p the actual cycle time.

$$q_s = \frac{t_s}{t_p} = \frac{t_p - t_0}{t_p} \tag{3}$$

The production rate q_P describes the ratio between the nominal cycle time t_0 and the real cycle time t_v and is defined in equation 4.

$$q_P = 1 - \frac{t_0}{t_{0v}} \tag{4}$$

The downtime rate during set up q_{Ssu} describes the ratio between the nominal set up time T_{su0} and the real set up time T_{su}, see equation 5.

$$q_{Ssu} = 1 - \frac{T_{su0}}{T_{su}} \tag{5}$$

The production time for a batch including the setup time can then be defined as in equation 6:

$$T_{pb} = \frac{T_{su0}}{(1-q_{Ssu})} + \frac{N \cdot t_0}{(1-q_S)(1-q_P)} \tag{6}$$

Reduced occupation in a manufacturing system leads to consequences for all manufactured parts. This situation can be considered in different ways, hence the free production resource can be considered both as an economic asset and a disadvantage depending on the situation. In a long term view the manufactured parts must carry the costs for the over capacity. The over capacity time can be distributed over all the batches in relation to their production time T_{pb} by introducing a degree of occupation U_{RP}, calculated as the quotient between real production time T_{prod} and planned production time T_{plan} according to according to equation 7 and 8. T_{free} is the time for the free, non occupied production time.

$$T_{plan} = T_{prod} + T_{free} \tag{7}$$

$$U_{RP} = \frac{T_{prod}}{T_{plan}} \tag{8}$$

The extra free capacity $T_{free,b}$ to be added to a specific batch is calculated according to equation 9. The free time can be considered as a setup time at the same time as the equipment is available for manufacturing:

$$T_{free,b} = \frac{1-U_{RP}}{U_{RP}} T_{pb} \tag{9}$$

With these parameters together with economic data, the cost of production per part can be calculated. The economic parameters included in the model is the following :

- Tool cost k_A
- Material cost k_B
- Equipment cost during production k_{CP}
- Equipment cost during down time and set up k_{CS}
- Wage cost k_D

The cost of production per par can then be calculated using equation 10.

$$\begin{aligned} k = {} & \frac{k_A \cdot N}{N \cdot (1-q_Q) \cdot (1-q_P)} + \frac{k_B \cdot N}{N \cdot (1-q_Q)} + \\ & \frac{k_{CP} \cdot t_0 \cdot N}{60 \cdot N \cdot (1-q_Q) \cdot (1-q_P)} + \\ & \frac{k_{CS}}{60 \cdot N \cdot (1-q_Q)} \cdot \frac{t_0 \cdot N}{(1-q_P)} \cdot \frac{q_S}{(1-q_S)} + \\ & \frac{k_{CS}}{60 \cdot N \cdot (1-q_Q)} \left(\frac{T_{su0}}{(1-q_{Ssu})} + \frac{1-U_{RP}}{U_{RP}} T_{pb} \right) + \\ & \frac{k_D}{60 \cdot N \cdot (1-q_Q)} \left(\frac{t_0 \cdot N}{(1-q_S) \cdot (1-q_P)} \right) + \\ & \frac{k_D}{60 \cdot N \cdot (1-q_Q)} \left(\frac{T_{su0}}{(1-q_{Ssu})} + \frac{1-U_{RP}}{U_{RP}} T_{pb} \right) \end{aligned} \tag{10}$$

To be able to simulate the effect of an improvement of the production process, a number of factors are introduced. The development factors are x_p, x_{su} and the cost factor κ_c, where x_p describe the improvement in cycle time that is achieved due to the development of the process. Likewise, x_{su} describes the improvement in set up time. κ_c is used to model changes in costs in primarily existing equipment, and can be used to determine the limit of investment justified to for example a decrease of the downtime rate to a certain value.

The manufacturing economic model including development and cost factors is described in equation 11.

$$
\begin{aligned}
k = &\frac{k_A \cdot N}{N \cdot (1-q_Q)\cdot(1-q_P)} + \frac{k_B \cdot N}{N\cdot(1-q_Q)} + \\
&\frac{\kappa_C \cdot k_{CP} \cdot x_P \cdot t_0 \cdot N}{60 \cdot N \cdot (1-q_Q)\cdot(1-q_P)} + \\
&\frac{\kappa_C \cdot k_{CS}}{60 \cdot N \cdot (1-q_Q)} \cdot \frac{x_P \cdot t_0 \cdot N}{(1-q_P)} \cdot \frac{q_S}{(1-q_S)} + \\
&\frac{\kappa_C \cdot k_{CS}}{60 \cdot N \cdot (1-q_Q)} \left(\frac{x_{su} \cdot T_{su0}}{(1-q_{Ssu})} + \frac{1-U_{RP}}{U_{RP}} T_{pb} \right) + \\
&\frac{k_D}{60 \cdot N \cdot (1-q_Q)} \left(\frac{x_P \cdot t_0 \cdot N}{(1-q_S)\cdot(1-q_P)} \right) + \\
&\frac{k_D}{60 \cdot N \cdot (1-q_Q)} \left(\frac{x_{su} \cdot T_{su0}}{(1-q_{Ssu})} + \frac{1-U_{RP}}{U_{RP}} T_{pb} \right)
\end{aligned}
\tag{11}
$$

Changes in part cost caused by a limited change in an arbitrary variable *z*, is calculated by partial derivative, and is described in linear form in equation 12.

$$\Delta k_i = \frac{\partial k_n}{\partial z} \cdot \Delta z \tag{12}$$

Cost neutral changes in each variable can be studied by putting the change in part costs $\Delta k_i = 0$. Equation 12 is written in a cost neutral form in equation 13, describing the size of the reduction in downtime share required to compensate for a change in wage costs.

$$\Delta q_{Si} = -\Delta k_{Di} \cdot \frac{\dfrac{\partial k_i}{\partial k_{Di}}}{\dfrac{\partial k_i}{\partial q_{Si}}} \tag{13}$$

7 THE CASE STUDY

The implementation of the method SPA resulted in data connected to a number of products produced at the factory at the choosen manufacturing unit. This paper will present an analysis of one of these products, called product x. Besides as an analysis of this product, this case study presentation also can be viewed as an example of how you can apply the cost model described in equation 11, when an SPA has been made and other necessary production data is available.

From the data collection phase of the case study, including a systematic registration of scrap, down time and production rate according to the method described in section 4, the values of the parameters in equation 10 was obtained and is presented in Table 3. In this case study the cost of reduced occupation has not been considered and there is no data for the deviation from the nominal set up time for this product. Inserting the values of the parameters in Table 3 into equation 10, the cost of production per part *k* equals 386,8 kr, which implies that this processing step adds 144,27 kr to the cost of the product.

Table 3 : Calculated values of the parameters in the economic model

q_Q	q_S	q_P	q_{Ssu}	T_{su0}	t_0
0.0230	0.4273	0	0	120	7.2
N_0	k_A	k_B	k_{CP}	k_{CS}	k_D
1000	13.7	242.53	420	420	150

During the analysis of the collected data, new ways to present the costs was developed. This was done by starting out from equation 10 and then divide the cost *k* into the different parts that add costs in the chosen processing step. The chart in Figure 1 below shows the part cost *k* together with k_B and these new cost items; k_{VA}, k_{NVA} k_{Tsu}, k_Q, k_S, and k_P, where *k* equals the sum of k_B, k_{VA}, k_{NVA} and k_{Tsu}, see equation 14.

$$k = k_B + k_{VA} + k_{NVA} + k_{Tsu} \tag{14}$$

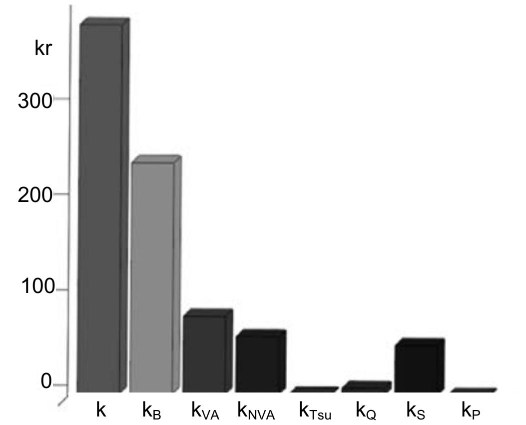


Figure 1: The part cost *k* and different parts of *k*.

k_{VA} is the value added part of the cost and is defined in equation 15. k_{VA} describes the cost added in this processing step, without considering q_Q, q_S, q_P and T_{su}.

$$k_{VA} = k_A + (k_{CP} + k_D)\cdot t_0 \tag{15}$$

k_{NVA} is the cost of non value added activities and constitutes of all costs related to the loss parameters q_Q, q_S and q_P, and is defined in equation 16.

$$k_{NVA} = k_Q + k_S + k_P \tag{16}$$

k_{Tsu} constitutes of the costs related to set up and is defined in equation 17.

$$k_{Tsu} = \frac{k_{CS} + k_D}{60 \cdot N \cdot (1-q_Q)} \cdot \frac{T_{su0}}{(1-q_{Ssu})} \tag{17}$$

By dividing the part cost *k* into k_B, k_{VA}, k_{NVA} and k_{Tsu} you get a quick overview of the cost condition of the product. The size of k_{NVA} indicates the potential cost reduction by decreasing the disturbances in the production. k_{NVA} make up 42.3 % of the total cost added in this processing step, or 61.0 kr per part. Correspondingly, k_{VA} make up 56.9 % of the total cost added, or 82.1 kr expressed in cost per part. k_{Tsu} is calculated to 1.2 kr per part. The combination of a large batch size and a long cycle time makes k_{Tsu} relatively insignificant regarding the part cost of product x at this chosen processing step.

A division of k_{NVA} can be made into k_Q, k_S and k_P, with the purpose to find out the specific costs contributed by each result parameter q_Q, q_S, q_P. k_Q describes the costs connected to the scrap rate q_Q and is defined in equation 18.

$$k_Q = \frac{N \cdot q_Q}{N \cdot (1-q_Q)} \cdot \left(k_A + k_B + \frac{k_{CP} \cdot t_0}{60} + \frac{k_D \cdot t_0}{60} \right) \quad (18)$$

The largest part of k_{NVA} consists in this case of costs related to the down time rate q_S. This cost is named k_S and is defined in equation 19. k_S was calculated to 52.7 kr per part and constitutes 86.3 % of k_{NVA}.

$$k_S = \frac{k_{CS} \cdot t_0 \cdot N \cdot q_S}{60 \cdot N \cdot (1-q_Q) \cdot (1-q_S) \cdot (1-q_P)} + \frac{k_D \cdot t_0 \cdot N \cdot q_S}{60 \cdot N \cdot (1-q_Q) \cdot (1-q_S) \cdot (1-q_P)} \quad (19)$$

k_P describes the costs caused by q_P and is defined in equation 20 and was calculated to 0.7 kr per part.

$$k_P = \frac{k_A \cdot N \cdot q_P}{N \cdot (1-q_Q) \cdot (1-q_P)} + \frac{k_{CP} \cdot t_0 \cdot N \cdot q_P}{60 \cdot N \cdot (1-q_Q) \cdot (1-q_P)} + \frac{k_D \cdot t_0 \cdot N \cdot q_P}{60 \cdot N \cdot (1-q_Q) \cdot (1-q_P)} \quad (20)$$

The size of k_{NVA} for this product at this production unit implies that there is a substantial amount of money to be saved if k_{NVA} can be reduced, if the total annual volume of the product is considered. The annual volume of product x is estimated to 9300 units by the company. If assuming that the value of k_{NVA} is intact over a year, the theoretical cost reduction becomes roughly: $k_{NVA} \cdot 9300 = 61 \cdot 9300 = 567300$ kr.

To reduce k_{NVA} you have to know the result parameters and the factors connected to these parameters that together constituting the value of k_{NVA}. When combining the result from the SPA and the parameters calculated from this data, it will be possible to obtain the influence on the manufacturing cost of every result parameter and factor registered in the SPA. In Figure 2 the factors having the largest effect on the part cost of product x as a result of the performed SPA are shown as their cost per part. The figure illustrates that apart from the tool factor A2, the major factors behind the size of k_{NVA} are related to the process C and organizational issues D.

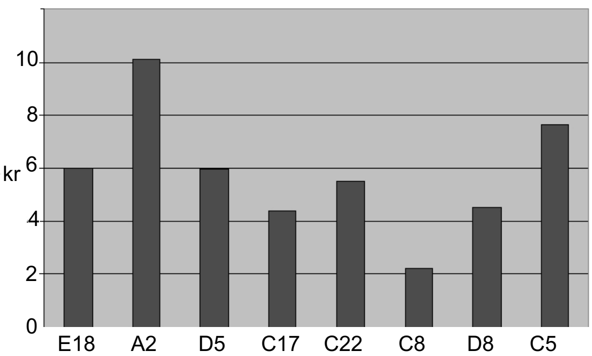


Figure 2: The factors having the largest effect on the part cost of product x

After quantifying the cost of the production disturbances and the factors in the different units or cells of the manufacturing plant, it will be possible to compare different disturbances influence of the manufacturing cost and thereby for example be able to make priorities between different development projects concerning the manufacturing process.

At this stage, when the critical parameters and factors connected these losses are obtained and economically quantified, then cost derivatives can be used to analyse different development scenarios. Figure 3 illustrate the cost neutral relationship between a change in the development factor $\Delta\kappa_C$ and a change in downtime rate Δq_S for product x. The figure shows for example that if the down time rate can decrease by 0.15, you get $\Delta\kappa_C$ to 0.35. This means that the equipment costs, k_{CS} and k_{CP}, can be increased by up to 35 % without increasing the part cost if the decrease in q_S can be accomplished.

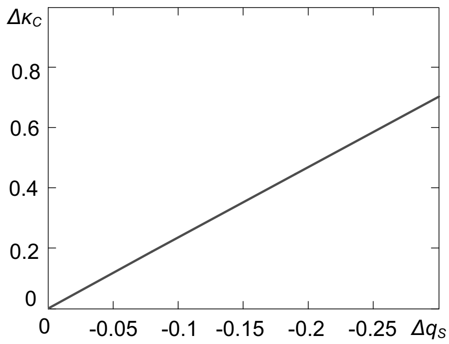


Figure 3: Cost neutral change between an increase in equipment cost factor $\Delta\kappa_C$ and reduction of down time rate Δq_S.

As Figure 2 illustrates, the largest cost factor for product x is A2, which is a factor connected to the tools in the machine. Figure 4 shows the cost neutral relationship for product x between an increase of the tool cost Δk_A and a decrease of the process development factor Δx_P. This relationship illustrates the maximum cost increase of improved tooling capable of reducing the cycle time to a certain level. The dotted graph shows the relationship when new but more expensive tools enables a decrease in x_P, but at the same time causes an estimated increase in the scrap rate q_Q by 0.01. The continuous graph shows the relationship without any increase in q_Q.

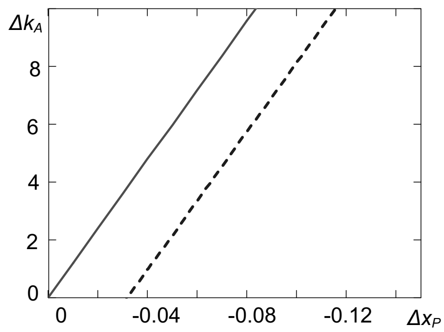


Figure 4: Cost neutral changes between an increase in tool cost Δk_A and reduction of the development factor Δx_P (continuous graph) and an increase in Δk_A plus a 0.01 increase of q_Q and a reduction of Δx_P (dotted graph).

This analysis is made for a single product in a processing step where several other products is produced. These other products and the costs connected to them must of course be taken into consideration to get a general picture of the costs in this processing step.

8 DISCUSSION AND CONCLUSIONS

To sum up the analysis described in the previous section, the high value of q_S is making a significant impact on the manufacturing cost of this product at this processing step. This case study shows that there is a potential in analysing the manufacturing performance of today and alternative simulated scenarios with the method described in this paper. An important prerequisite for these analyses to be reliable is the accuracy of the input data and how detailed this data is. If the model is implemented at every unit or cell at a factory it will then be possible to obtain which cost items that builds up the total manufacturing cost of a part or a group of parts and size of these items. Figure 5 shows the division of the part cost k made in section 7. This division makes it easy to get a clear view of the costs added in a processing step. Having this detailed information accessible for products in all processing steps would enabeling a greater insight and understanding about where in the manufacturing process there is the highest potential to lower the costs and thereby function as a basis for prioritys concerning production development activities.

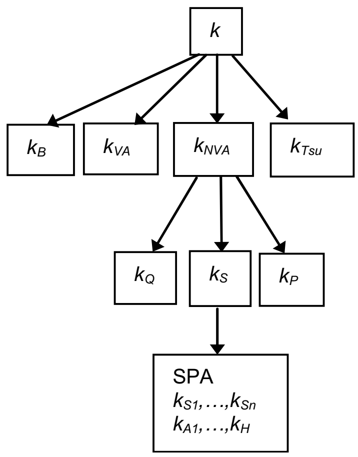


Figure 5: The division of k made in the presented case study analysis.

By combining production data and economic data into this economic model it is also possible to establish manufacturing economic development goals. A development goal could for example be to reduce the cost of manufacturing per part with 10 % for a product or a group of products. By the implementation of this model a plan to reach that goal could be established.

9 REFERENCES

[1] Cooper, R., Kaplan, R.S., 1992, Activity-Based Systems: Measuring the Costs of Resource Usage, Accounting Horizons, 6/3:1-13.

[2] Thyssen, J.T., Israelsen, P., Jorgensen B., 2006, Activity-based costing as a method for assessing the economics of modularization – A case study and beyond, International journal of production economics, 113:252-270.

[3] Tipnis, V.A., Mantel, S.J., Ravignani, G.J., 1981, Sensitivity Analysis for Macroeconomic and Microeconomic Models of New Manufacturing Processes, Annals of the CIRP, 30:401-404.

[4] Colding, B., 1978, Relative Effects of Shop Variables on Manufacturing Cost and Performance, Annals of the CIRP, 27:453-458.

[5] Colding, B., 2003, Tids- och kostnadsanalys vid skärande bearbetning – Optimering för synkront detaljflöde, Colding International Corporation, Stockholm.

[6] Alberti, N., Noto La Diega, S., Passannanti, A., 1985, Interdependence between Tool Fracture and Wear, Annals of the CIRP, 34:61-63.

[7] Knight, W.A., Poli, C.R., 1982, Design for Economical Use of Forging: Indication of General Relative Forging Costs, Annals of the CIRP, 31:159-163.

[8] Groover, M.P., 1987, Automation, Production Systems, and Computer Integrated Manufacturing, Prentice-Hall Inc, Englewood Cliffs, NJ.

[9] Ståhl, J.-E., 2007, Industriella Tillverkningssystem, Division of Production and Materials Engineering, Lund University, Lund.

A Dynamic Cost Model for the Effect of Improved Process Flexibility in Steel Plants

Joakim Storck[1] and Bengt Lindberg[2]

[1] National Post Graduate School in Metal Forming, Högskolan Dalarna, SE-78188 Borlänge, Sweden

[2] Department of Production Engineering, Royal Institute of Technology, SE-10044 Stockholm, Sweden

Abstract

Reduced setup times in the rolling mill generate flexibility which allows shorter leadtimes through continuous casting and hot rolling. Traditionally known as schedule-free rolling, this flexibility allows the rolling mill to handle variations without the need for buffering. Cost models based on system dynamics methodology are used to assess the economic potential. Effects on inventory, energy and work roll consumptions are analysed. The simulation results show that investments in flexible processes can be evaluated with dynamic cost models. There is an opportunity for significant cost reduction, but also lowered environmental impact due to reduced energy consumption.

Keywords:

Hot rolling; Setup time reduction; Cost estimation

1 INTRODUCTION

This paper presents a model for the influence of setup time reductions in the hot rolling mill (HRM) on manufacturing costs in steel production. The model concerns an integrated steel mill with continuous casting (CC) and hot rolling of slabs for production of stainless steel strip. Setup time reduction is a generally recognised method to gain the flexibility needed for one-piece just-in-time (JIT) flow [1,2,3]. However, as noted by Nye et al. [4], most authors follow the advice of Shingo [5] and recommend that changeover times are reduced to less than 10 minutes. The actual relation between setup times and production performance has not been much discussed and it is not clear how large investments in setup time reduction that can be economically justified.

Steel production is a very capital and energy intensive business, and producers must take every opportunity to reduce manufacturing cost. If the meltshop/CC processes are isolated from the rolling mill through buffering, this will cause the average leadtime for workpieces (**slabs**) to be longer than in a process with less buffering [6]. Slabs cool more the longer they are stored, and the heat from the melting process is lost. Decoupled operation of meltshop/CC and HRM is thus known as **cold charging**, since slabs are cool when they enter the rolling mill. The opposite, i.e. integrated production with short lead times, is known as **hot charging**.

The economic potential in setup time reduction is mainly due to energy savings when transfer times are short between meltshop, CC and HRM. It was previously reported [7] that setup time reduction may facilitate savings in the order of 4 EUR/ton even when no measures are taken to reduce work roll consumption due to the increased amount of roll conditioning. This represents a total a cost reduction of 4 MEUR on a yearly production of one million tons, translating directly into increased profit. If work rolls are utilised more efficiently the potential is even bigger, an issue that will be discussed in the present paper.

1.1 The role of flexibility

When many different products are made in the same production line, the need for quick changeovers arises to allow production to rapidly shift to the product currently in demand. Browne et al. (cited in [8]), termed this **process flexibility**, i.e. 'the ability to produce a given set of part types'. Process flexibility in the context of steel production can be interpreted as the ability, at any given time,

- of the meltshop to produce a particular steel grade,
- of the continuous caster to cast a particular steel grade and slab geometry,
- of the hot strip mill to roll a slab of a particular grade, width and thickness into the desired target thickness.

Setup times in a modern hot strip mill as seen in Figure 1 are typically in the order of 15 minutes, but it is not unusual that setups in older mills are in the order of one hour or more. If a

Figure 1: Hot strip mill with replacement roll pair ready.

modern finishing mill is paired with an older roughing mill, overall production planning will be controlled by the flexibility of the older (less flexible) mill. It becomes the job of production planning to ensure efficient resource utilisation on expense of increased buffering and leadtimes.

Extensive research has been done on production scheduling in steelmaking. Lee et al. [9] reviewed scheduling research until 1995. Dorn and Shams [10] implemented an expert system at Böhler Uddeholm in Austria. More recently Tang et al. presented optimisation models for meltshop/CC [11] and hot rolling [12], while Cowling and Rezig [13] presented an optimisation model for integrated scheduling of meltshop/CC and HRM. Singh et al. [14] presented an optimisation model for minimising material handling in the slab yard.

Production scheduling aims to produce schedules that e.g. maximise equipment utilisation and charging temperature while minimising buffering and work roll consumption. The above cited examples of scheduling research show that there is a potential for improvement which can be realised through advanced production scheduling algorithms. However, the quality of an optimal schedule deteriorates as the number of grades and geometries increases, i.e. in plants that produce a large number of low volume high-grade products.

Instead, the constraints on job order can be relieved if each process step can be made more flexible, with the ability to adapt to the current situation. Buffer levels and leadtimes may then be reduced [6]. We argue that inflexible processes is a major source of waste in steel production, and that setup time reduction and similar investments in improved flexibility should have the highest priority for steel producers.

1.2 Problem structure

The conceptual problem, shown schematically in Figure 2, is to relate operating conditions in the plant to manufacturing costs. Figure 2 does answer how this should be done; it only shows some of the considered process parameters, and the division of costs into three components: gas for reheat furnaces (energy), tied capital (inventory/WIP) and work rolls.

In the following sections we present two models based on the **system dynamics** methodology [15,16]. System dynamics (SD) is a generic method for describing complex causal

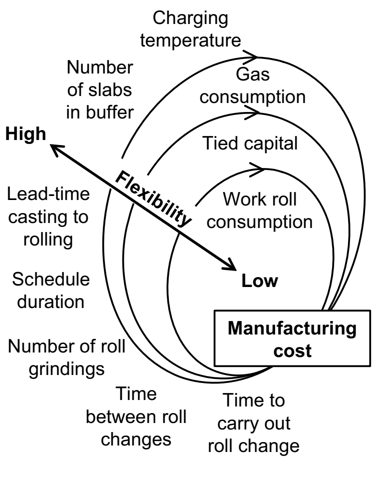


Figure 2: Structure of the cost model. The level of process flexibility balance costs for work roll consumption against buffering and reheat energy.

relations with stocks and flows [16], which are visualised as **causal loop diagrams** (CLD). The method is based on continuous simulation [17], with an emphasis on feedback structures. The models presented in this paper are implemented and run in the Vensim® simulation environment.

The first model, which is described in [7], is called the **basic model**, and estimates the cost of WIP, reheat energy and work rolls without respect to the dynamics of manufacturing over time. As shown schematically in Figure 2, this model accounts for the costs of;

(a) **Energy**: If leadtime and buffering is reduced, some heat from the melting and casting is preserved, and the mean temperature of slabs entering the rolling mill increase. Since slabs must hold about 1250°C during hot rolling, and reduced leadtime allow fuel consumption in the reheat furnaces to be lowered.

(b) **Material**: The amount of WIP and hence tied capital is reduced, which is of particular interest to producers of stainless steel due to high raw material prices.

(c) **Tools**: Increasing the number of roll changes also require more frequent conditioning of the work roll surface. This may result in raised overall roll consumption, hence causing the tool costs to increase.

The second, more elaborate model, is called the **dynamic model**, and is discussed in detail in [18]. It is in effect a manufacturing simulation model designed to predict resource consumption during production, and to produce hourly cost rates during a simulation run. As with the basic model, the dynamic model is used to predict costs that arise from WIP, reheating and work roll consumption. The principles behind this are presented in the following section. Details specific to the basic and dynamic models will be discussed in Section 3 and 4 respectively, while the simulation results are presented and analysed in Section 5.

2 WIP, ENERGY AND WORK ROLLS

According to Little's law [2], the amount of WIP is the product of throughput (tons/h) and the average cycle time. If all workpieces that are scheduled for rolling within the next HRM program are temporarily stored in the slab yard while the preceding program is processed, the average transfer time becomes equal to the duration of the HRM program, i.e. the time between roll changes (setups).

The cost of tied capital is defined as the internal interest rate of return on WIP. If the amount of WIP and the material price is known, this can be readily calculated. The material price depend on raw material prices and is at present around 3000 EUR/ton for an 18% Cr, 8% Ni stainless steel. For a throughput of 100 tons/h, and 10 h between roll changes, average WIP becomes 1000 tons. This yields 3 MEUR in yearly interest on tied capital for a 10% internal rate of return.

Energy costs cover only fuel consumption in the reheating furnace of the HRM. The energy required to reheat a workpiece to rolling temperature can be estimated if the initial and final temperatures are known along with the geometry and boundary conditions. A relation between lead time and workpiece temperature was found [18] through simulated cooling in STEELTEMP® [19]. If transfer time from casting to rolling is 6 h, workpieces hold about 500°C on arrival to the rolling mill. The actual fuel consumption will then depend on e.g. furnace efficiency and the reheat cost follows from gas

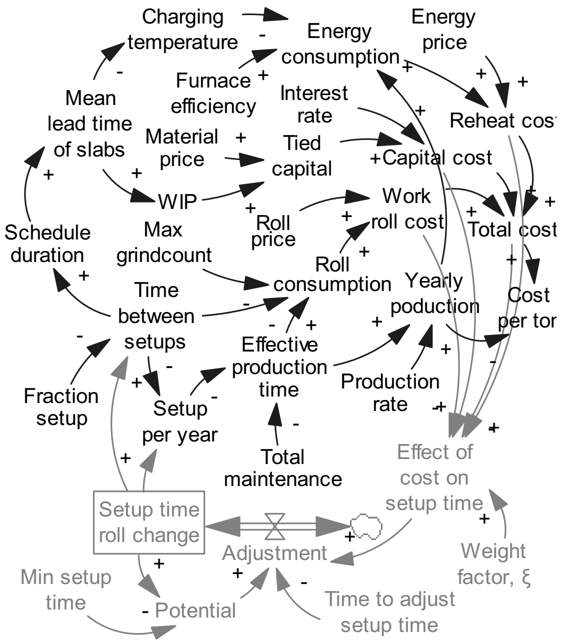


Figure 3: Causal loop diagram for the basic cost model.

price and consumption based on heat content, charging temperature and furnace efficiency.

The work roll cost is a consequence of roll wear due to frictional and thermal forces during rolling [20,21]. The rolls are therefore conditioned in a roll grinding machine after each roll change. Work rolls normally have a cast iron core covered by a steel surface. When the surface layer has been worn down the roll pair is scrapped and replaced. Provided that throughput is 100 tons/h and that a 0.5 h changeover is initiated every 10 h, production during one program is 950 tons. If a roll pair lasts 40 grindings and costs 0.1 MEUR, the approximate work roll cost becomes 2.6 EUR/ton.

The actual relation between roll wear and operating conditions is complex, and Munther and Lenard [20] concluded that the wear depends on process parameters such as temperature, velocity, load and geometry, but also that the material properties of the rolls and the rolled metal are of equal importance. Still, experiments by Pellizzari et al. [21] indicated that after a transitory running in period, wear was in effect proportional to the rolled distance. It therefore seems reasonable to make the assumption that wear is proportional to the production volume (tons).

A fundamental assumption in both models is that the total amount of setup is constant even though the duration of a single setup can change. Hence, if the setup time can be decreased through investment in setup time reduction, more frequent roll changes can be carried out while the total fraction of setup is unchanged. This results in less buffering and shorter lead times since the average time that slabs wait for the previous program to be completed is shortened.

Rolling mill schedules are designed by selecting workpieces that can be processed in sequence during a single setup of the mill [9]. A schedule lasts from one roll change to the next, and it is assumed that all workpieces which are to be rolled within the next schedule must be produced in advance and buffered while the current program is completed. Hence, the average amount of WIP depends on the setup time, and the dynamics of the entire model can be controlled by altering the setup time in the rolling mill.

3 THE BASIC MODEL

The basic model implements the cost equations in an SD model, the CLD of which is seen in Figure 3. This model can be compared to the conceptual model in Figure 2. When executed, it runs through a transient period as it converges against equilibrium due to a negative feedback loop where the setup time is adjusted based on the cost distribution.

The relative size of cost components is controlled by the parameter ‘effect of cost on setup time’, seen in the lower part of Figure 3. The balance between capital costs and reheat energy on one hand and work roll consumption on the other is adjusted through the weight factor, ξ. Changing ξ corresponds to altering the flexibility; an increase reduces buffering and lead times, i.e. capital and reheat costs, while a decrease reduces work roll consumption (cf. Figure 2).

The basic model assumes that a roll pair lasts for a given number of changes and that the grinding depth during conditioning is unchanged when the frequency of roll changes increase. The rolls may therefore have to be scrapped unnecessarily often, causing the work roll cost to increase sharply for quick setups with frequent roll changes.

4 THE DYNAMIC MODEL

The dynamic model can be used to simulate production over extended periods, e.g. several months. The output includes hourly estimates of the power and energy consumption in the reheat furnace, throughput in the rolling mill and build-up and depletion of WIP in the slab yard. The model is described in detail in [18], but some of its main features are discussed in the following paragraphs.

The dynamic model targets some of the limitations in the basic model mentioned in [7], namely:

- Leadtimes and WIP depend on the current state of the model and vary over time in response to events such as roll changes and maintenance stops.
- It accounts for actual roll wear and adjusts the amount of conditioning depending on produced volume on the current roll pair.

The excess conditioning of the work rolls that occur in the basic model with increased setup frequency (cf. Section 3) is handled by lowering the grinding depth to compensate for the reduced wear. The extent to which this is done is controlled by the **grinding efficiency**, which is introduced in the model as seen in Figure 4. The parameter ‘rolling rate’ in this figure

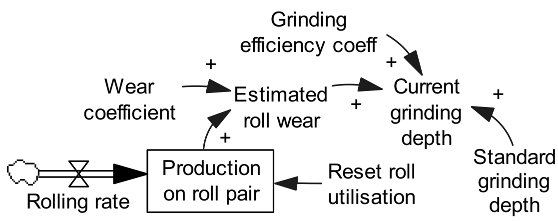


Figure 4: Estimation of roll wear and grinding depth in the dynamic model.

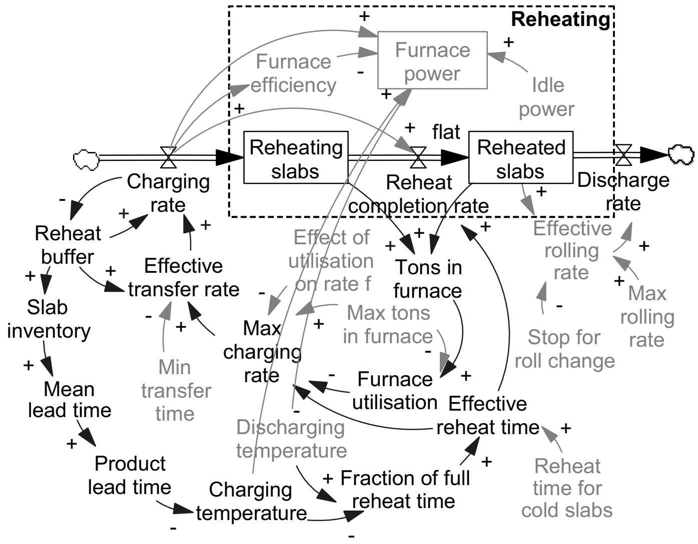


Figure 5: Causal loop diagram of the reheating process in the dynamic cost model.

is equal to the 'discharge rate' of Figure 5. The 'current grinding depth' is a fraction of a standard grinding depth which follows from the maximum number of grindings that a roll pair is estimated to last, i.e. 'max grindcount' of Figure 3.

The part of the CLD representing the rolling mill reheat furnace is shown in Figure 5. Reheating is modelled as a two stage process, where cold slabs are turned into hot slabs with a rate that depends on the time needed to reach rolling temperature. This 'effective reheat time' (Figure 5) is a function of the charging and discharging temperatures. The discharging temperature is set to the rolling temperature (1250°C), while the charging temperature depends on the cooling time, i.e. the leadtime from casting to rolling.

Energy consumption is calculated from the furnace power which, as seen in Figure 5, depends on the charging rate (tons/h), as well as on the initial and final temperatures. The charging rate in turn depends on current furnace utilisation and effective reheat time.

5 RESULTS AND DISCUSSION

The dynamic model was used to simulate 20 weeks of production with a time step of 0.0625 h (3.75 min) for 126 different combinations of setup time and grinding efficiency. The experiment covered 21 different setup times in the range from 0.05 to 1.5 h, while six different values of the grinding efficiency coefficient, ranging from 0 to 1, were used. Material price was set to 3000 EUR/ton (cf. Section 2), and the energy price was 50 EUR/MWh. Each run resulted in one value per cost component. The results from all these runs were aggregated to produce the plots of Figures 6, 7 and 8.

Two extremes were considered; that the grinding depth is the same regardless of rolled volume, or that it is proportional to the rolled volume. Figure 6 shows the costs of work rolls, reheating and tied capital for the case of constant grinding depth. As seen in the figure, the contribution of work roll costs increase dramatically for low changeover times (i.e. frequent changeovers). This corresponds to the results of the basic model reported in [7].

If proportional wear and grinding is assumed, this results in constant work roll cost irrespective of the changeover frequency. In the ideal case, conditioning removes only the wear that actually occurred since last setup, essentially only a touch up of the surface finish when the rolls are changed very often. Figure 7 shows the cost components as function of setup time under these conditions.

In reality, it is unlikely that the ideal of Figure 6 can be achieved. The actual amount of conditioning can be expected to fall between the cases in Figures 6 and 7. Figure 8 shows

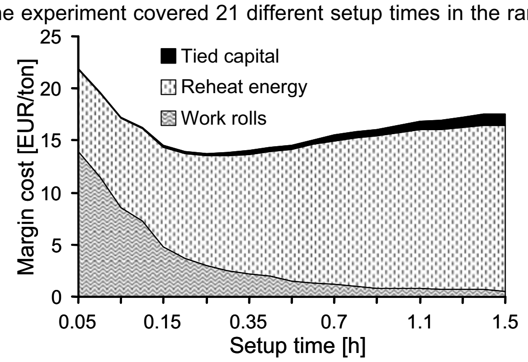


Figure 6: Cost as function of setup time in the dynamic model with constant roll grinding (q=0).

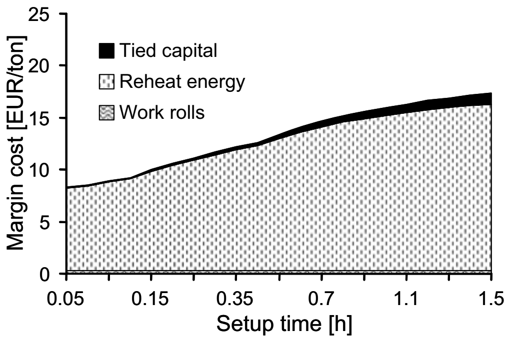


Figure 7: Cost as function of setup time in the dynamic model with ideal proportional roll grinding (q=1).

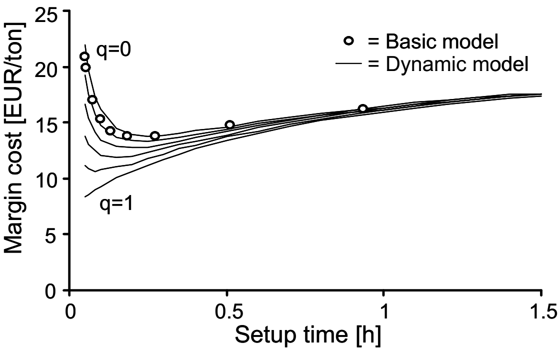


Figure 8: Total cost for the two models. *q* indicate grinding efficiency (constant=0, ideal=1) in the dynamic model.

the sum of the cost components based on the results of the dynamic model for constant grinding as well as for the ideal and intermediate proportional grinding cases. Figure 8 also includes the results from the basic model, which can be seen to correspond well with the case of constant grinding depth in the dynamic model. The real outcome of an investment in improved process flexibility can be expected to be represented by an intermediate curve, depending on to what extent unnecessary conditioning can be avoided.

As shown in Figure 6, and in accordance with [7], setup time reduction for the case of fixed grinding depth (Figure 8, *q*=0) give a maximum 21% cost reduction of 3.7 EUR/ton at 0.25 h setup compared to a margin cost of 17.5 EUR/ton at 1.5 h setup; in itself a considerable improvement. If proportional grinding is employed, i.e. any of the curves representing *q*>0 in Figure 8, further cost reductions are possible. The case of ideal proportional grinding (Figure 7 and Figure 8, *q*=1), allows a total 47% cost reduction, or 8.2 EUR/ton at 0.1 h setup. The savings potential at 0.25 h setup is 6.4 EUR/ton, i.e. 37% of the cost at 1.5 h setup time. Hence, a mill with one million tons yearly production may increase its revenue by 8 MEUR per year.

Throughout this paper it was assumed that improved flexibility is achieved through setup time reduction. However, the same problem is traditionally targeted in **schedule free rolling**, which can be accomplished in several ways. One is to use inline roll grinding [6], constantly refreshing the work roll surface. Another option is to use very durable work rolls, and change these before any pronounced wear contour develop. The investment in quick setups is thereby replaced by investment in more expensive work rolls. In practice, a combination of both may be of interest.

Results from a run with 1.5 h setup time and 30 h mean schedule duration are shown in Figure 9. The upper plot

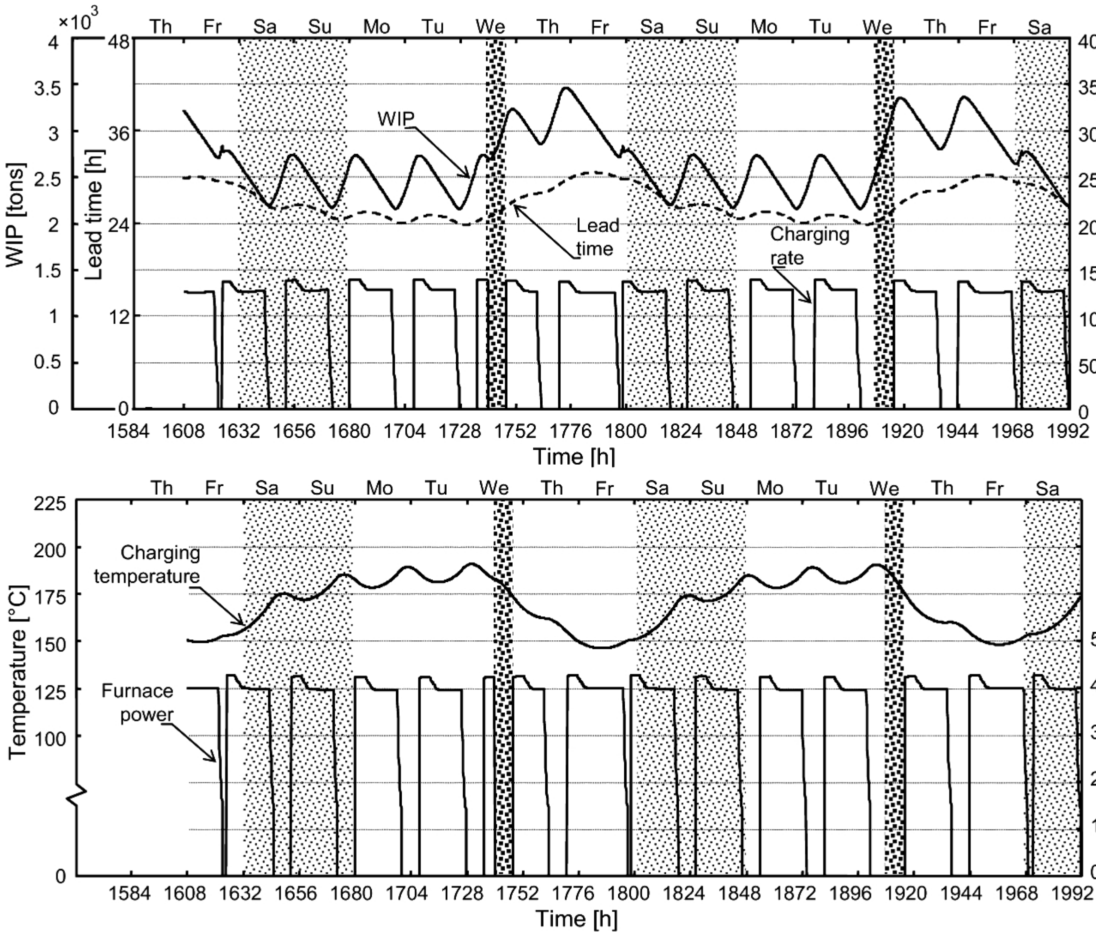


Figure 9: Time series from simulation with the dynamic cost model. Top: WIP, lead time and production rate. Bottom: charging temperature and furnace power. Shading on Wednesdays indicates maintenance stops.

shows WIP, cooling time and throughput over two weeks starting about 10 weeks into the simulation, while the lower plot shows charging temperature and furnace power for the same period. Roll changes are carried out in the periods when charging rate is zero. These periods also correspond to completion of the previous and initiation of a new rolling mill schedule. The amount of WIP varies from 2200 tons to 2500 tons, a variation that is mainly caused by the scheduling mechanism. The magnitude of this variation grows when the time between scheduling events increase, i.e. with longer setup times. Since variation is buffered by some combination of inventory, capacity or time [2], reducing setup times has an overall beneficial effect. This is shown by the results presented in this paper.

6 CONCLUSIONS

Scheduling requirements for the hot strip mill were conceived to ensure maximum utilisation of the work rolls and minimise the number of roll changes in order to maximise utilisation of the mill. They are the result of a strong focus on roll economy, and reflect the assumption that

- setup times are fixed and unalterable, and
- the cost of energy and tied capital is small in comparison to the cost of work rolls.

Two models for estimation of the effect of setup time on manufacturing cost in continuous casting and hot rolling of steel were presented. The main findings were:

- Setup time reduction and more frequent roll changes can reduce manufacturing costs significantly.
- Quick setups allow more frequent changes and shorter rolling mill programs, which stabilises WIP on a lower level with less variation.

The savings potential is mainly due to reduced energy consumption. As previously stated [7], increasing energy prices are a strong incentive for improved manufacturing flexibility and shorter setup times.

Based on the results of the models presented in this paper, the claim that inflexible processes is a major source of waste in steel production seems to be justified. Setup time reduction and similar investments in improved flexibility should therefore gain higher priority for steel producers in their future attempts to improve competitiveness.

7 ACKNOWLEDGMENTS

This research was sponsored by the Swedish Steel Producers' Association (Jernkontoret), Outokumpu Stainless AB and the Knowledge Foundation (KK-stiftelsen).

8 REFERENCES

[1] Ohno, T., 1988, Toyota Production System: Beyond large-scale production, Productivity Press, New York.

[2] Hopp, W.J., Spearman, L., 2000, Factory Physics: Foundations of manufacturing management, 2nd ed., Irwin/McGraw-Hill, New York.

[3] Liker, J.K., , 2004, The Toyota way: 14 Management principles from the world's greatest Manufacturer, McGraw-Hill, New York.

[4] Nye, T.J., Jewkes, E.M., Dilts, D.M., 2001, Optimal investment in setup reduction in manufacturing systems with WIP inventories, Eur. J. Oper. Res., 135:128-141.

[5] Shingo, S., 1985, A revolution in manufacturing: The SMED system, Productivity Press, Cambridge, MA.

[6] Storck, J., Lindberg, B., 2007, A lean production strategy for hot charge operation of a steel mill, Proc. of the 2007 IET Int. Conf. on Agile Manufacturing (ICAM'07), 158-167, 9-11 July 2007, Collingwood College, Durham University, Durham, U.K.

[7] Storck, J., Lindberg, B., 2007, A cost model for the effect of setup time reduction in stainless steel strip production, Proc. of the 1:st Swedish Production Symposium, Stahre, J. (Ed.), Chalmers University of Technology, Göteborg.

[8] De Toni, A., Tonchia, S., 2005, Definitions and linkages between operational and strategic flexibilities, Omega, 33:525-540.

[9] Lee, H.S., Murthy, S.S., Haider, S.W., Morse, D.V., 1996, Primary production scheduling at steelmaking industries, IBM J. Res. Dev., 40/2:231-252.

[10] Dorn J., Shams, R., 1996, Scheduling high-grade steelmaking, IEEE Expert, 11/1:28-35.

[11] Tang, L., Liu, J., Rong, A., Yang, Z., 2000, A mathematical programming model for scheduling steelmaking-continuous casting production, Eur. J. Oper. Res., 120:423-435.

[12] Tang, L., Liu, J., Rong, A., Yang, Z., 2000, A multiple traveling salesman problem model for hot rolling scheduling in Shanghai Baoshan iron & steel complex, Eur. J. Oper. Res., 124:267-282.

[13] Cowling, P., Rezig, W., 2000, Integration of continuous caster and hot strip mill planning for steel production, Journal of Scheduling, 3:185-208.

[14] Singh, A.K., Srinivas and Tiwari, M.K., 2004, Modelling the slab stack shuffling problem in developing steel rolling schedule and its solution using improved parallel genetic algorithm, Int. J. Prod. Econ., 91:135-147.

[15] Forrester, J.W., 1961, Industrial Dynamics, MIT-press, Cambridge, MA.

[16] Sterman, J.D., 2000, Business Dynamics: Systems thinking and modeling for a complex world, Irwin/McGraw-Hill, New York.

[17] Law, A.M., Kelton, W.D., 2000, Simulation modelling and analysis, 3rd ed. MacGraw-Hill Int.

[18] Storck, J., Lindberg, B., 2007, Dynamic cost modelling in steel production, submitted to Int. J. Prod. Econ., December 2007.

[19] Leden, B., 1986, STEELTEMP – A program for temperature analysis in steel plants, Scand. J. Metall., 15:215-223.

[20] Munther, P., Lenard, J.G., 1995, Tribology during hot, flat rolling of steels, CIRP Annals, 44/1:213-216.

[21] Pellizzari, M., Molinari, A., Straffelini, G., 2005, Tribological behaviour of hot rolling rolls, Wear 259:1281-1289.

Manufacturing Characteristics of Subcontractor SMME:s – an Empirical Study

Malin Löfving[1,2], Christer Johansson[1,2] and Mats Winroth[1]

[1] Department of Industrial Engineering and Management, School of Engineering, Jönköping University, Sweden

[2] Departments of Product and Production Development, Division of Production Systems, Chalmers University of Technology, Sweden

Abstract

This paper presents empirical findings on manufacturing characteristics at subcontractor small and medium sized manufacturing enterprises (SMME). The SMME:s play a significant role in today's economy, but they do not act to the same extent on the global market as the international larger companies. To remain competitive on the global market SMME:s should improve manufacturing.

The study has been accomplished in different industries and different sizes to identify their characteristics. The results indicate that SMME:s are focused on process technologies and not on the entire manufacturing system. The companies also have difficulties in locating and hiring skilled people.

Keywords:

Small and medium sized enterprises, decision criteria, manufacturing characteristics

1 INTRODUCTION

Small and medium sized manufacturing companies (SMME) play a significant role in today's economy [1]. 99 % of the companies are SMME:s [2]. Many SMME:s are subcontractors to larger companies [3] and are therefore highly dependent on them. Historically, a subcontractor works for a limited number of larger customers [4]. Competitive advantages have been low price and whether the subcontractor is located near the customer [4]. This is however changing and more and more SMME:s are aware of this challenge and try to increase their customer base.

The challenges for the SMME:s have increased [4] during the last years. The challenges come from the increasing globalization and the increasing customer demands. The globalization increases the access to new technologies, to new knowledge, and to new markets, but the competition also becomes more severe [5]. For SMME:s, the globalization can be seen as a larger challenge than for larger international companies. The SMME:s are not acting to the same extent on the global market as the international larger companies. International larger companies can drive national SMME out of the market, since SMME:s do not have the same financial strengths or resources as the larger companies [6]. To remain competitive on the global market the SMME should improve productivity [1], management, and manufacturing.

Methods, tools, and philosophies used by large manufacturing companies, like Lean Production, have not yet been adopted by all SMME:s [7, 8]. To be able to adopt new manufacturing and managerial methods, financial resources, time, and skilled employees, both operators and management are needed [8, 9]. SMME:s suffer in general from financial constraints and lack of human resources [8, 9]. This makes it more difficult for smaller companies to implement methods that were developed by larger companies. Due to this, smaller manufacturing companies must have the ability to become more efficient. The first step to develop manufacturing in SMME:s is to the identify their manufacturing characteristics. The next step is to develop methods and tools that can support manufacturing in SMME:s.

Research into small and medium sized enterprises has increased the last years [10], especially in the business perspective. Studies about SMME have focused on growth and development of the company [10]. When manufacturing is mentioned in the SMME literature, it is often described as a part of the company that affects the market and the growth of the company [11] [12]. For the SMME:s the manufacturing is often a central part of the company. SMME:s' concentrate largely on manufacturing and technical issues, rather than on management, organisation, or customer service [4]. Even if the SMME in general focuses on manufacturing, there is no general research about characteristics of manufacturing in SMME:s.

2 RESEARCH METHODOLOGY

The research presented here was performed as an empirical study. The purpose of the study was to identify manufacturing characteristics in subcontractor SMME:s. Data was collected through interviews and observations at 20 different subcontractor SMME:s in Southern Sweden. The companies participated in three different projects, depending on type of industry. Before visiting the companies, semi structured interview questions were designed. Semi structured interviews [13] were chosen as interview method to enable follow up questions. The first questions were about the company and ownership. The next questions were about the future, strategy, specific questions about the project, and manufacturing system. The interviews lasted not more than 2 hours. The observations were carried out in the manufacturing plant after the interviews and followed a structure that was developed prior to the study.

After the investigations of the companies were done, the interviews were transcribed.

The analysis of the data has followed Miles and Huberman [14] cross-case analysis. Cross-case analysis was chosen to enable a deeper understanding and to enhance the possibility to generalize.

Miltenburg's [15] six decision criteria for a manufacturing system were chosen for the cross-case analysis. The six decision criteria were chosen since they support the manufacturing system and provide the necessary conditions for successful manufacturing. After forming the characteristics of each decision criteria (see table 1), the entire manufacturing system can be described. The six decision criteria are:

- Human resource
- Organisation structure and controls
- Sourcing
- Production planning and control
- Process technology
- Facilities

Each decision criteria had certain characteristics that were listed in a table (see table 1). Each company was analysed in terms of the characteristics of the six decision criteria. The industries were then analysed and compared in terms of size, ownership, and manufacturing process.

2.1 Studied SMME:s

According to Cagliano [4], there are two main business models for SMME:s: original manufacturers and subcontractors. The manufacturers produce their own products, developed and engineered in house. The subcontractors provide manufacturing capacity and process technology skills, but are not endowed with design skills [9]. In this paper the subcontractors are studied.

The studied SMME:s are subcontractors from three different industries: foundry industry, polymer industry, and

Table 1: Table of the manufacturing characteristics

	Small foundry	**Medium sized foundry**	**Small polymer company**	**Medium sized polymer company**	**Small automotive part company**	**Medium sized automotive part company**
Human Resources						
Promotion opportunities	Easy	Medium	Easy	Easy	Easy	Medium
Level of education	Low	Low	Low	Low	Low	Low
Multi-skilled	Yes	No	Yes	No	Yes	No
Participation of employees in problem solving and improvement activities	Low	Low	Medium	Medium	Medium	Medium
Organisation structure and controls						
Organisational structure	Flat	Hierarchical	Flat	Hierarchical	Flat	Hierarchical
Centralised or decentralised organisation	Centralised	Centralised	Centralised	Centralised	Centralised	Centralised
Informal or formal	Informal	Formal	Informal	Formal	Informal	Formal
importance of line or staff	Staff	Line	Staff	Line	Staff	Line
Sourcing						
Relationship with supplier	Long-term	Dependent of the owner	Long-term	Long-term	Long-term	Long-term
Procedure of deciding whether a product will be produced internally or obtained from a supplier	CEO	Sales department	CEO	Sales department	CEO	Sales department
How supplier are chosen	Contacts	Price	Customer decide	Price	Contacts	Price
Production planning and control						
Whether the systems are centralised or decentralised	Centralised	Centralised	Centralised	Centralised	Centralised	Centralised
Push or pull	Push	Push	Push	Push	Push	Push

automotive parts industry. The definition of SMME used in this paper is made by the European Union [16]. According to this, companies under 49 employees are defined as small and companies, whilst companies between 50 to 250 employees are defined as medium sized [16]. The companies in this study have between 27 and 150 employees; 11 of the companies are small and 9 are medium sized. The visited companies are in the middle or southern part of Sweden, 88 % are located in the Gnosjö Region. The Gnosjö Region is famous for its strong entrepreneurial spirit, known as the Gnosjö spirit [17]. In Gislaved municipality in the Gnosjö region we can find 0,3 % of the Swedish population and 6,5 % of the Swedish polymer companies [18].

The studied SMME:s have different kinds of ownership: family business, venture capital, or a company group. Groups that own the studied SMME:s are often owned by families or private investors. When ownership was considered to affect the company characteristic, it is mentioned in the paper.

3 RESULT

The research analysis shows that four of Miltenburg's [15] six decision criteria of manufacturing system are dependent on size (see chapter 2 for the six criteria). The criteria discussed in this paper are:

- Human resource
- Organisation structure and controls
- Sourcing
- Production planning and control

The analysis indicates that these four criteria above focus on the organisation rather than the technology and facilities. The analysis also indicates that these four criteria are dependent on size and ownership, rather than what is manufactured. The other two criteria, process technology and facilities are dependent on what is manufactured rather than on size. The two criteria are not analysed further in this paper.

3.1 Human resource

The study indicates that a qualified person can relatively easily be promoted within the small company, but for a person with higher education there are not many positions with challenging tasks within the small company.

The study shows that personnel are well-known to the CEO and management in the studied small company. The CEO or management knows who is skilled and experienced and who to promote. There is however a difficulty to be promoted in the studied smaller company since the organisational structure is flat and the promotion opportunities are scarce. Promotion opportunities are scarce because there are not many positions in the small company; another reason is that the turnover of staff is low. Manufacturing Management in the studied small companies often stay in a position for many years, before changing position.

In the studied medium size companies there are more positions in the organisation and there are more hierarchical levels. The CEO or Management does not know all personnel and relies on the manufacturing Management knowledge on skilled and experienced personnel that can be promoted.

A company representative stated that "when appointing a position, there are two alternatives: internal or external recruitment". Both of the studied small and medium sized companies prefer to promote internally. The result of the analysis indicates that the personality and knowledge about the manufacturing processes are more important than education and management skills. Production managers or foremen should have deep knowledge in the manufacturing processes and technology. A skilled and experienced operator with knowledge in the technologies is often chosen to be manager. This is general for all of the companies, independent of size and industry. According to the studied companies, especially the foundries, it takes years of practice to learn the manufacturing processes. Management can be studied at the university or be taught by a consultant, but manufacturing processes must be learnt by experience. The analysis shows that the different companies have a lot in common, independent of the type of industry, but the companies often believe that they are unique, i.e. the manufacturing process are difficult to learn. The manufacturing processes are different in the studied companies, especially in the foundries were they often have specialised processes. We found in the study that it is often not necessary to have the deepest knowledge in manufacturing to become manager and sometimes it is better for the development of the company and the manufacturing to promote or appoint a person with new ideas and management skills, rather than the person with the deepest manufacturing knowledge.

In the studied companies, there are limited numbers of people with higher education, independent of the type of industry or size. The results of the analysis show that:

- The studied companies recruit thorough personal acquaintance.
- Difficulties in finding and employing people with higher education.
- The CEO or owner is not familiar in discussing with people with higher education.
- The company does not have the financial resources to employ people with higher education.

The study indicates that personnel, both operators and managers, are often recruited through personal acquaintances.

People with higher education often want to live in urban areas, where their friends and family live, or in larger cities with larger and well-known companies, many companies stated. If moving to a town without knowing anyone, it can be difficult to feel at home even if the work is interesting. The well-educated people want to have possibilities to be promoted, however there are limited opportunities in the studied companies.

The studies indicate that if the owner or CEO does not have any higher education, he or she does not know what the people with higher education can provide the company. The financial resources also affect when employing.

The study also indicates that it is difficult for the companies to find and employ skilled operators, who want to move to the town where the company is located. The studied companies are often located in smaller towns like Gnosjö. Gnosjö Region is known for its large number of manufacturing companies and the employment rate is high. The studied companies do not want to employ the unemployed people in the town, because they are not considered suitable for work in manufacturing. To be able to grow, many of the studied companies have recruited skilled immigrants during the last years. The analysis shows that the trend is in employing immigrants, and to employ persons from Eastern Europe via unique agencies. The agencies act as an intermediary between the company and skilled persons in Eastern Europe. The immigrants and skilled people from Eastern Europe are also willing to move to the town where the company is located. In a studied foundry company, 50 % of the employees had an international background with different nationalities.

The study shows that the employees in the studied small companies are more multi-skilled than the medium sized companies. For example, in a foundry company the operators should survey the machine, remove excess material, and do maintenance. Some of them could also operate the CNC-machines. In a medium sized foundry company an operator only monitored the foundry machine, another did maintenance, and a third monitored the CNC-machine.

Participation of employees in problem solving is dependent of the ownership and size of the company. In general, in the studied companies the participations of employees in problem solving are quite low. In the studied small company the owner can have control of everything and decide everything by him- or herself, in the studied medium size company the decisions are often made by the foreman or the production manager.

3.2 Organisation structure and controls

The organizational structure is comparatively flat in most of the studied small companies. The study shows that between the operator and the CEO there is one level in the small companies, i.e. the production manager or the foreman. The studied small companies often had a foreman, or a number of foremen, instead of a production manager. The production manager's or foreman's role in the small company is mostly operational with a strong focus on daily and monthly activities. The study shows that the studied small companies have short term and operational focus in the manufacturing. Therefore the manufacturing processes are prioritized, not the long term development of the manufacturing system. Only a small number of the studied companies worked with systematic production development tools and methods. Many had however tried some methods, but due to the short term horizon and lack of resources that had time for implementing long term changes, the development tools and methods had not worked.

The foreman carries out many different work tasks in the studied companies, manager, planner, industrial engineer, and sometimes even operator. He or she has little time for each task and often does not have knowledge, experience, or interest in all tasks.

In the studied medium sized companies, the organization has more hierarchical levels than the small companies. The study shows that between the operator and CEO there are often two levels, both foreman and production manager. In the studied small company the owner is often the CEO. The CEO in the studied small companies often takes a major part of the decisions. Neither the studied small nor medium sized companies are decentralized. In the studied medium sized companies the CEO, production manager, and quality manager make the decisions that affect manufacturing. The study indicates that the decisions that are made are tactical or operational decisions, not strategic and long-term. The operational focus means that the companies have process focus. None of the studied companies has a system focus, probably because of the short time perspective and the lack of competence.

The studied companies do not have a production development department, but they have multi skilled and experienced employees that can carry out a large number of different work tasks in manufacturing when necessary. The operators can come up with a request to the CEO and the CEO listens to them. Decisions can be made immediately. The small companies, independent of industry, are more informal than the medium sized companies because of the size; everybody knows the owners and the owners do not hesitate to visit the manufacturing premises, discussing with personnel study what is going on. In the medium sized company, the owner or CEO does not have the same contact with the operators and rely more on secondary information from the planning department, quality, and production manager.

The line consists of a team of operators that works with a machine or machine group. The line is more important in the medium sized companies than in the small company, where staffs are important. In the small company the relations are closer and each employee is important in order to be able to fulfil the daily orders. The result of the analysis seems to be that medium sized company is too large to have a close relationship with every employee, i.e. the line is more important.

3.3 Sourcing

The studied companies, independent of size and type, purchase raw material from a limited number of suppliers. The studied companies manufacture components from raw material. The components are sent to an assembly company or to the customer's manufacturing plant to be assembled to a final product.

The study indicates that the relationship between the studied companies and their raw material supplier is long-term and based on existing acquaintance. A company representative, from a medium sized company, said “we are more professional today because we rely on performance when choosing suppliers rather than on existing acquaintance”. This company has grown from small to medium size during the last years and the market share has increased. The studied company can affect the price because of the larger size of quantities of material that are being purchased. This is general for the studied medium sized automotive part companies and foundry companies that are owned by a company group.

In the polymer industry, the studied companies' customers choose and own the injection moulding tool as well as the raw material. The customers also choose the raw material supplier as well as the tool supplier. This result is unique for the polymer industry in this study, see table 1. The customer often chooses the cheapest supplier, independent of the quality of the material. This affects the studied companies; they can not influence what material is chosen and must manufacture the products in the time the customer wants, independent of material quality and how the material works in the machines. The customers also own the tools, so the studied companies only own the injection moulding machines and the facilities. The result of the analysis indicates that this is a somewhat problematic situation for the companies; they can not influence the development of the products they manufacture or the choice of material. The study shows that the single competitive advantage often is the price and the customer often chooses the injection moulding company with the lowest price. The studied companies get decreased profitability when offering the customer a low price, and cannot afford to purchase new machines or develop organisation and manufacturing. One medium sized polymer company lost their largest customer this way; they could not offer the lowest price. To survive they changed their business plans and goals, from relying on the old customer to offering complete solutions in injection moulding. They are also developing a new department which specialty is plastic material component development.

The studied polymer companies do not have own competence about material, the material supplier or machine supplier is the expert of plastic material. The studied polymer companies take advices from the suppliers when they need information about material or what machine they should purchase. In the studied foundry companies the operators and purchasers have knowledge about the raw materials that are used in the products, in this case iron, steel, or non-ferrous metal. The studied foundries are proud of having their own experts in material, both within the companies and at their own material research centre which educates personnel in both foundry technology and casting materials. In the studied automotive part companies, the customer decides material but the companies own and purchase the material.

The study shows that the CEO and/or production manager decides whether a component should be manufactured internally or purchased externally. The studied companies can purchase the components if the component is cheaper to purchase from 3rd party in a low wage countries, or if there is a lack of capacity. The study also shows that a new order from a customer can trigger the purchase of a new machine. In the studied foundries, new machines are often purchased to replace old ones. Polymer companies can decide to produce a product at a competitor in the Gnosjö region, instead of buying there own machines, this because of the Gnosjö spirit and near relationship between the competitors in the region.

3.4 Production planning and control

The study shows that all the studied companies, independent of size or type of industry, use the manufacturing forecasts. The studied companies use forecast planning because they have long delivery times, 1 week to 16 weeks. Normally the customer puts the order 3 to 4 weeks before delivery. To be able to handle the delivery time the studied companies trust the forecast. The result of the analysis shows that the components are sometimes manufactured before the customer lays the orders, i.e. the component is pushed through the manufacturing system.

In the studied polymer and automotive part companies, one batch can take one month to produce. The customer often wants to receive the entire order at a certain time, which leads to increased inventory for the company. The studied companies accept this manufacturing situation to get the order.

The study shows that none of the studied companies knows the lead time in manufacturing, but all of them are aware of the machine time, i.e. the time for a component to be manufactured in a machine. The time before and after the manufacturing of a component has not been important in the studied companies due to the importance of the pay-back time. The empirical study indicates that the studied companies therefore focus on manufacturing processes, not the whole manufacturing system.

In the studied foundries and automotive part companies, the delivery time of raw material sometimes take several months, depending on the demand of the material. The studied polymer companies have shorter delivery times of plastic material.

In this time of economic boom, the studied companies have problems in delivering the components to the customer on time, because the companies cannot process the volume of incoming orders. The study indicates that prioritized customers get their products on time, whilst others must wait. Prioritized customers are often the largest and most important customers that the studied companies cannot afford to lose. The studied companies, however, try to decrease the largest customers' share and increase the number of customers, to be able to survive if the largest company cancels the contract. In the studied foundry companies the delivery precision is 90 %, the other two industrial sectors have higher. The analysis of the study shows that many companies have capacity problems, and have started or are going to start a new shift. This is done instead of working with manufacturing improvement methods.

The studied medium sized companies have their own production planner, but in the smaller companies it is the foreman who plans the manufacturing. The foreman is often promoted from manufacturing and does not have any knowledge in production planning and production planning systems. The planning of the production can be seen as a time consuming task rather than as a help for the foreman.

4 CONCLUSION

The result of the study shows that there are more differences between small and medium sized companies, independent of the type of industry (see table 1).

Today, in the economic boom, companies get customers' orders due to the lack of manufacturing capacity in Swedish companies. Today, the polymer industry is affected by the competition from low wages countries and the customer chooses in most of the cases subcontractors on price. The Swedish polymer SMME:s can compete up to a certain level with prices, when they will start losing money. To break this trend, in the polymer industry and the other studied industries, the companies must take action and develop the company in the long-term to survive. There are a limited number of people with higher education in the studied companies. To be able to develop the manufacturing in the companies, persons with higher education should be employed in manufacturing. The companies have difficulties to employ people with higher education because of the location of the company, and because the lack of financial resources. The companies also have difficulties employing skilled operators from the town the company is located in; often they employ immigrants or people from Eastern Europe through agencies. If they had not employed skilled immigrants, the companies would have problems to grow and to survive.

Another reason for the insecure situation of the studied companies is the fact that they do not develop and sell their own products. In the studied polymer companies, as well as other companies, the customer chooses and owns the plastic material. The studied companies provide manufacturing processes and technology without being able to affect the situation. Some of the studied companies are aware of this and have taken different actions to be more than a provider of manufacturing processes, however, with various successes. Today the studied companies have operational and/or tactical perspective in the manufacturing. To survive in the long term perspective the studied companies should have a long term focus in the manufacturing to be able to develop the manufacturing system further.

The studied companies, independent of size and type of industry, mostly focus on the different manufacturing processes; they do not have a holistic view on manufacturing system. The studied companies know the machine times, but not the lead time. To be able to rise the delivery time the studied companies also should measure the lead time. Today the prioritized customer gets the components on time; other customers must often wait for the delivery. To survive and be able of keeping their customers, the companies must focus on the whole manufacturing flow, not only the different machines. The companies are not aware of the lead time and do not have the competence to measure and analyse the lead time. Often, when more capacity is needed, a new shift is started to gain more capacity. The time for a component in the machine is often reduced, but nothing is done to the manufacturing system, due to lack of competence.

Only a limited number of the studied companies are working with manufacturing system development. Today, when the low cost countries are getting better and better at manufacturing quality products, there are many challenges for the studied companies. To survive in this time of globalization the companies must focus on manufacturing system development. Today, the lack of resources of employees with higher education and lack of financial resources makes it more difficult to compete on the global arena.

5 ACKNOWLEDGEMENT

This result in this paper is based on findings from three different projects, “Strategic competence needs in the foundry industry”, a collaboration with Swerea SweCast AB, sponsored by Vinnova, “Future research areas in the polymer industry”, sponsored by Sparbanken Alfa Stiftelsen and Efficient Manufacturing – implementing Lean production in small and medium sized manufacturing companies in the Gnosjö Region, a collaboration with Industriforum and IUC Gnosjö.

We would also like to thank all the participating companies in this study.

6 REFERENCES

[1] Gunasekaran, A., Forker, L., Kobu, B., 2000, Improving operations performance in a small company: a case study, International Journal of Operations & Production Management, 20/3:316-226.

[2] European Commission, 2007, Putting SMEs First, DG Enterprise and Industry.

[3] Kungl. Ingenjörsvetenskapsakademin (IVA) 2005, Made in Sweden - Produktion för konkurernskraft - Syntesrapport, Stockholm.

[4] Cagliano, R., Spina, G., 2002, A comparison of practice-performance models between small manufacturers and subcontractors, International Journal of Operations & Production Management, 22/12:1367-1388.

[5] Nutek, 2007, Årsbok 2007, Stockholm.

[6] Teknisk Framsyn 2003, Produktionssystemet - Sveriges välfärdsmotor, Stockholm.

[7] Greatbanks, R., 1998, Can SMME's be World Class?, Proceedings of Conference on Stimulating Manufacturing Excellence in Small and Medium sized Enterprises, Sheffield.

[8] Achanga, P., Shehab, E., Roy, R., Nelder, G., 2006, Critical success factors for lean implementation within SMEs, Journal of Manufacturing Technology Management, 17/4:460-471

[9] Grando, A., Belvedere, V., 2006, District's manufacturing performances: A comparison among large, small-to-medium-sized and district enterprises, International Journal of Production Economics, 104/1:85-99.

[10] Pasanen, M., 2003, In Search of Factors Affecting SME Performance, University of Kuopio, Kuopio.

[11] Steiner, M.P., Solem, O., 1988, Factors for success in small manufacturing firms, Journal of Small Business Management, 26/1:51-56.

[12] Julien, P.A., 1998, The State of the Art in Small Business and Entrepreneurship, Ashgate, Aldershot.

[13] Bryman, A., Bell, E., 2003, Business Research Methods, Oxford University Press, Oxford.

[14] Miles, M.B., Huberman, M.A., 1994, Qualitative Data Analysis, Sage Publications, Thousand Oaks.

[15] Miltenburg, J., 2005, Manufacturing Strategy, Productivity Press, New York.

[16] European Union, 2003, SME Definition.

[17] Karlsson, C., Larsson, J., 1993, A macroview of the Gnosjö entrepreneurial spirit, in Small business dynamics, Edited by Karlsson, C., Storey, D.,Johannisson, B.,Routledge, London.

[18] Berlin, L., 2006, GNOHOW - Gnosjöregionens klustersatsning, Gislaved.

Integrated Information as an Enabler for Change Impact Evaluation in Manufacturing Life-cycle Management

Victoria Rogstrand[1], Johan Nielsen[2] and Torsten Kjellberg[1]
[1] Department of Production Engineering, The Royal Institute of Technology, Sweden
[2] Eurostep AB, Sweden

Abstract
To be able to completely evaluate and make a reliable cost analysis, all aspects of a planned change must be known. A study of change impact evaluation in Manufacturing Life-cycle Management (MLM) has been performed. This article discusses the need for an integrated information representation of the product, process, and resource domains to enable this. It also proposes the use of manufacturing requirements in order to completely evaluate the impact. Process planning was studied to identify manufacturing requirements, which have been represented in the information standard STEP 10303-239 to demonstrate the possibilities.

Keywords:
Information Representation, Decision Making, Change Impact Evaluation, Manufacturing

1 INTRODUCTION

Today, the life span of manufacturing systems is in most cases much longer than the product market life. This means that existing manufacturing systems must be altered and re-used for manufacturing of new products to secure cost-effective product realization. When introducing a change in manufacturing it is essential to be able to identify and understand the effect that the change will have. To ensure effective production it is important to know the capability of the manufacturing system and how it will limit the product. The knowledge of the capability is today most often in form of knowledge of experienced engineers. The current trend is that engineers are changing positions more often and it is more and more uncommon that people work at the same place for 25 years. Some companies even have policies that engineers should change position within the company every other year to be good candidates for a management position. This has the effect that long-term experience within a certain application area e.g. process planning is getting rarer. In order to safeguard the experience built up within the company, without hindering the new mobility for engineers, it is important to manage the manufacturing information in such a way that it is possible to trace why a certain decision was taken. An information architecture gives an integrated information view of how the different domains of manufacturing are related and influence each other. [1]

However, merely having the information architecture does not enable the tracing of why a decision was taken. To be able to completely evaluate different alternatives, all aspects of a change must be known. In order to make well based decisions, information is needed that indicated why there are constraints and how the company prioritizes between different alternatives. The use of manufacturing requirements as defined in [2] and section 5, would support this process.

Explicit manufacturing requirements capture the information stating why e.g. a process and resource are related, or why a process was chosen. Manufacturing requirements enable better decisions to be taken since it is possible to trace the reason for a connection or choice. When you have access to information regarding why previous decisions were taken, you will be able to successfully evaluate the impact of a proposed change.

This indicates the need for an information platform that manages data throughout the manufacturing life-cycle. Therefore, two hypotheses of the need for integrated information were stated. A study was performed of the process planning of a new manufacturing system for a cylinder head. The objective of the study was to find manufacturing requirements and test the hypotheses through implementation of the manufacturing requirements using ISO 10303-239 (PLCS) [3]

2 METHOD

This article assumes the following hypotheses.

Hypothesis 1:

‘Complete and integrated information structures of the product, process, and resource domain is a prerequisite for performing a change impact evaluation’

Hypothesis 2:

‘Requirement information regarding the different domains and the relationships between them is necessary in order to make informed decisions based on the impact evaluation’

These hypotheses were tested [4] through a study identifying manufacturing requirements and how they affect process planning during a major change of a manufacturing system for a cylinder head.

A case study was performed at a large truck manufacturer in Sweden to identify manufacturing requirements during process planning. To provide multiple sources of evidence, [5], both interviews and documents were used to collect data. The interviews were open-ended and focused [5]. There were a mix of questions concerning facts and opinions of the interviewees. The interviewees were engineers with more than ten years experience. The language used was Swedish and the interviews were 1.5 – 2 hours long. The data was collected during the winter of 2007/2008. The analysis of the interviews was made through explanation-building [5]. Between every set of interviews careful analysis was done to evaluate the stated hypothesis. The collected documents (tenders from suppliers of manufacturing equipment) were developed during the purchase and process planning of a new manufacturing line in the plant, and contain information of manufacturing resources and processes. They also contain requirements of the manufacturing line and its processes.

The identified manufacturing requirements and data about the product, process and resources was represented through the use of ISO 10303, STEP, STandard for the Exchange of Product model data, in particular ISO 10303-239 Product Life-cycle Support (PLCS). PLCS was chosen because it has been shown to be capable of representing information about product, process, resource, and manufacturing requirements [2]. The data from the study was implemented in the platform Share-A-space™. Share-A-space™ was chosen because it is a commercial system which supports PLCS [6].

3 PROCESS PLANNING

Today, decisions during process planning are mostly based on the knowledge of experienced engineers [7]. This knowledge, which is built up over many years, is essential for making good decisions. The information available is usually managed by a variety of applications and systems, making it vital to know where to look and how to interpret the information. Very little of this knowledge is formalized in any form of knowledge management. It is the experience of the engineer that shows him/her where to find the information needed when making good decisions. The information needed during process planning is created and maintained in a variety of systems. The issue of finding the right information and consolidating it, is a task that depends on the individual persons performing the task. This makes it difficult, and sometimes impossible, to find and consolidate the information. The consolidation is a prerequisite in order to be able to evaluate and trace the impact of a change request. If the information about how the different domains are connected is not accessible, then it is impossible to evaluate how the different domains affect each other. Because of these circumstances, it is often not possible to fully predict and evaluate the impact of a change.

The knowledge of the engineer is also the base on which information consolidation depends during analysis. Today, the consolidation is done manually by the process planner. There is a desire to capture this knowledge, that has been developed in the company, in an information platform.

4 INFORMATION PLATFORM FOR MANUFACTURING LIFE-CYCLE MANAGEMENT

Information management is essential for any company and for any business process. All decisions are based on available information making it essential to be able to have an overall picture, an information architecture, of the information owned by the company. The information architecture is not the same as the IT-architecture. The IT-architecture is the structure of available applications and support systems. The information contained in different applications and support systems are most often not integrated and consolidated with each other.

The information architecture on the other hand, should give a clear picture of where to find wanted information, regardless if the information is stored in an application, a document or the mind of a person. The information architecture makes it possible to find important information, how information is related; between different domains, but also in between different levels of detail. This information architecture is hereafter referred to as the Manufacturing Life-cycle Management (MLM) platform.

The information supporting process planning is in large quantities and complex. There is a great variety of information supporting process planning. Process planning may start while the product is still in its conceptual stage. The resources and its processes may be planned for the new product but it may also be already existing resources in the current manufacturing system. The existing manufacturing system contains what is here referred to as physical resources, actual manufacturing equipment that you can touch and that are located on the shop floor. Therefore it is essential to be able to manage data of real existing equipment and its processes as well as planned resources and processes.

During early development phases the MLM platform should provide support for evaluating e.g. feasibility or cost of, a certain design decision. The product is in the conceptual phase stated by functional requirements [8,9]. During the continued product realization process, the functional requirements are realized by design parameters and embodied in technical solutions. The process also has a functional view of what is needed to be performed e.g. a hole. Based on what resources are available in the existing manufacturing system there is a variety of types of possible processes, e.g. drilling or milling, that can be chosen to realize the design parameters. During process planning the MLM platform should provide a process library from which most of the information needed can be retrieved. Only information that is new as a result of the current process planning should be added. The process library and its knowledge base will ensure that best practise is followed and philosophies such as kaizen are enabled.

[10, 11] present a methodology for developing product families that conform to the limits of the current manufacturing system. An MLM platform should support all aspects and life-cycle stages of the manufacturing of a product, cf. Figure 1. The purpose of the MLM platform is to support methodologies such as those described in [10, 11].

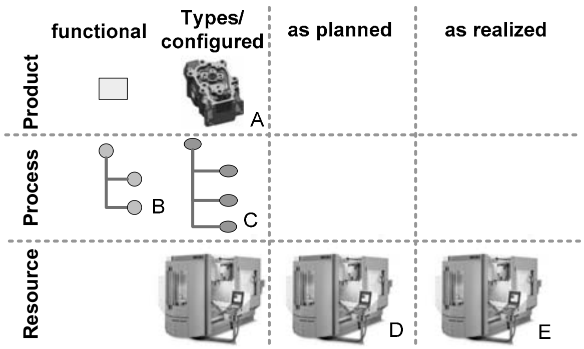


Figure 1: Domains and life cycle stages in manufacturing

The product and process also have planned and realized representations, e.g. to evaluate an outcome of a process in comparison to the planned. However, this is not in the scope of this article.

The integrated information platform that the MLM platform represents is essential when trying to evaluate how the proposed change will impact the existing manufacturing system.

5 MANUFACTURING REQUIREMENTS

As stated previously, it is essential to have a structured representation of product, process, and resource information in order to make a change impact evaluation. However, it is not possible to make a decision based on that information alone. In order to make qualified decisions of what processes, or combination of process and resource, are of vital importance to the success of the company, there is also a need for manufacturing requirements.

Manufacturing requirements are requirements that originate in manufacturing. Their purpose is to ensure efficient production. Efficient production can be achieved through the re-use of existing systems with as little alteration as possible. Another aspect of efficient production is to limit the flora of tools and other resources used. By reducing the variety of resources, maintenance and support can be streamlined. This reduction in variety can be achieved through the use of manufacturing requirements in the MLM platform to evaluate if similar processes may use identical tools. Manufacturing requirements will also facilitate the use of methods such as Design For Manufacturing (DFM), where the manufacturability of the product is considered during design. In Figure 2 DFM is represented by the manufacturing requirements that act on the product.

Therefore manufacturing requirements are mostly considered as constraints, as they constrain the product. For example, the current manufacturing system will limit (put constraints) on what is possible to manufacture.

Constraints are as [12] describes a type of requirement, and needs to be represented as any other requirement [13]. It needs to be versionable, possible to classify, indicate validity, relate to other information etc.

The manufacturing requirement information will add the crucial information pieces that allow the decision maker to see whether or not a specific change proposal will have an effect on products, processes, or resources that have been identified to be of vital importance to the success of the company. Manufacturing requirements gives extra information about why decisions are taken the way they are, stating what governs the decisions.

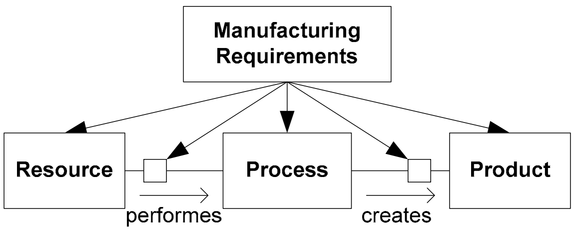


Figure 2: Manufacturing requirements

In the case of the new manufacturing system for the cylinder head, that was studied, the operation order was changed. Previously, the finishing of the top surface was the last operation, now it is part of the first operation. This change in the order of the operations led to a change of gripping points on the product. It has also introduced a requirement on all operations that come after the finishing of the top surface; they must not damage that surface. This scenario is represented by the manufacturing requirements that act on the process in Figure 2.

5.1 Manufacturing requirements' origin

The origin of a manufacturing requirement differs a lot. The resources constrain the type of processes that can be performed. For example, a geometry limitation of the product may originate in the manufacturing equipments capability in positioning a cutting tool in a certain set up. The requirements on a manufacturing process may come from a variety of sources; ergonomics (manual assembly, sound levels), environmental (restrictions of choice of fluids), goals of quality and cost.

Manufacturing requirements can also originate directly from products, processes, or resources. They can also originate from the relationship between a process and a resource, or a product and a process. This is an important aspect, since it might not be the process or resource themselves that are the source for a requirement, but rather the combination of the two, e.g. when a machine performs a certain process.

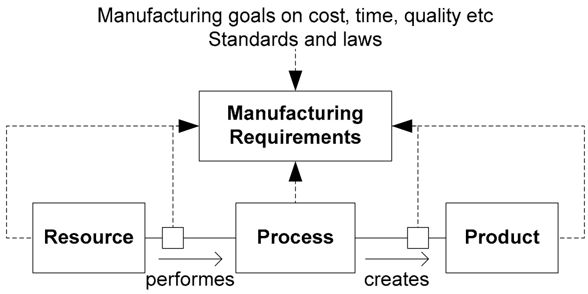


Figure 3: Origin of manufacturing requirements

In the case identified during the case study, where the order of the operations have changed, a requirement on flexibility in the manufacturing is the reason for choosing machining centres in favour of a transfer line. This requirement has in turn led to the requirement on all processes in the machining centres not to damage the finished contact surface of the cylinder head.

5.2 Benefits of the use of Requirements

The use of manufacturing requirements in the MLM platform allows for better management of the information that is contained in the different systems. Without the requirement information, the reason for the connections is lost.

In the case that a specific manufacturing resource is the source for a manufacturing requirement, the requirement will change or become invalid if the resource is replaced. Even if such a change causes the requirement to stop acting on the current manufacturing, the requirement will still be available through the history available in the MLM platform.

During the change impact evaluation, there will be multiple requirements that have to be evaluated and balanced against each other. The balancing of requirements is facilitated by classifying the requirements according to importance.

5.3 PLCS as a base for Manufacturing Life-cycle Management

The information about product, process, and resource is created in different systems during different life-cycle stages of the product and the manufacturing system. The variety of systems suggests that an information standard is needed to be able to integrate the information across the different domains.

The international standard ISO 10303-239 Product life-cycle support (PLCS), has been developed to support the information about a product over its complete life-cycle, from requirements to maintenance.

The difference between a product and a resource is simply the point of view. A product for one actor is a resource for another. PLCS, which support the entire life-cycle of a product, is therefore a good match to use as a starting point for an MLM platform. The use of PLCS as an MLM platform was tested through the use of Share-A-space™.

6 DISCUSSION

Automotive manufacturers have a need to estimate the cost of a change very early in the change process. The use of manufacturing requirements in change impact evaluation which has been presented in this paper is intended to facilitate this early cost estimate.

The decisions made during the change impact evaluation depend on a functioning information architecture and the traceability that manufacturing requirements provide.

Consolidating the information through the use of an information system such as Share-A-space™ eliminates the problem of having to access the information from every system and application in the IT-architecture, and manually keeping track of how they are connected.

Even if you are not using a standards based information solution to consolidate your information, you still have the problem of harmonizing terminology in different domains – especially when the information is created in different systems, by different persons who work in different contexts. PLCS uses reference data that solve this issue. The reference data ensures that different actors use the same terminology. The reference data also provide a mechanism through which the semantics of the standard can be extended. This is highly beneficial in the case of manufacturing requirements, since the standard itself does not have to be updated to support what has been presented in this paper. It is only the reference data that has to be updated.

7 CONCLUSIONS

An integrated information architecture of all product, process and resource information enables Manufacturing Life-cycle Management (MLM).

Furthermore, by having explicit requirements as relations between the product, process, and resource domains, it is possible to evaluate the full impact of a proposed change.

When introducing a change the use of manufacturing requirements will indicate where the changes will have an impact. The origin of the manufacturing requirement will indicate why it exists, and the priority of the requirement will indicate the impact of the change.

8 ACKNOWLEDGMENTS

This paper is a result of the research conducted in the ModArt (Model driven parts manufacturing) project. ModArt is a collaborative research project with several research organizations and industry partners, which is financed by VINNOVA. The participants include KTH, IVF, Scania, Sandvik Coromant, ITT-Flygt, Haldex Brake Systems, NMW, UGS, Dassault Systèmes, PTC, Eurostep, SIS, Epsilon, and Microtech.

9 REFERENCES

[1] Rogstrand, V., Kjellberg, T., 2005, Representation of Manufacturing Requirements in Collaborative Environments, International Conference on Manufacturing Research

[2] Rogstrand, V., Kjellberg, T., 2007, The use of Production Requirements in Model-Driven Parts Manufacturing, Proc. 4th International Conference of Digital Enterprise Technology

[3] ISO/TC184/SC4, 2004, Industrial Automation Systems and Integration – Product Data Representation and Exchange – Part 239: Application Protocol: Product Life-cycle Support, International Organization for Standardization, ISO 10303-239

[4] Popper, K., 2003, The Logic of Scientific Discovery, Routledge Classics, ISBN 0-415-27844-9

[5] Yin, R. K., 1990, Case Study Research - Design and Method, Sage Publications, ISBN: 0-8039-5665-2

[6] CIMdata, Inc., 2007, Eurostep's Share-A-space, A CIMdata Program Review, available on http://eurostep.com/ (2008-01-27)

[7] Rogstrand, V., Kjellberg, T., 2004, Interaction of Requirements in Manufacturing System Development, Proc. 14th International CIRP Design Seminar.

[8] International Council Systems Engineering INCOSE "Systems Engineering Handbook", 2004, http://www.incose.org/ProductsPubs/products/sehandbook.aspx (2005-05-24)

[9] Suh, N. P., 1990 Principles of Design, Oxford University Press, ISBN: 0-19-504345-6,

[10] Kimura, F., Nielsen, J., 2005 A Design Method for Product Family under Manufacturing Resource Constraints. Annals of CIRP, vol.54, no.1, pp.139-142

[11] Nielsen, J., Kimura, F., 2005 A Resource Capability Model to Support Product Family Analysis. LEM21, Nagoya, Japan.

[12] Harwell, R., Aslaksen, E., Hooks, I., Mengot, R., Ptack, K., 1993, What is a Requirement?, Proc. 3rd Third international Symposium of the INCOSE

[13] Kar, P., Bailey, M., 1996, Characteristics of Good Requirements, Proc. 6th Sixth international symposium of the INCOSE

Total Performance Analysis of a Downsized Manufacturing System

Nozomu Mishima[1], Shinsuke Kondou[1] and Keijiro Masui[1]

[1] Advanced Manufacturing Research Institute, National Institute of Industrial Science and Technology, Japan

Abstract

Micro mechanical fabrication is one of the key manufacturing technologies. However, in terms of cost and environmental impact, existing systems are not very efficient. It should be leaner. As an answer, AIST developed the first prototype of microfactory and insisted it could reduce environmental impact of micro mechanical fabrication. However, the effect has not been quantified. In this report the authors propose a system efficiency index (TPI) based on system throughput, cost and environmental impact. TPI is useful in evaluating the configuration of a microfactory-like system. In addition, taking flexibility into account, modified index is considered to simulate the effect on diverse-types-and-small-quantity production.

Keywords:

Microfactory, System efficiency, Total performance index, System configuration

1 INTRODUCTION

In recent world, there are many small mechanical parts and product are used for mobile phones, medical devices, home appliances, and so on. However, manufacturing systems for those devices are large and complex. Manufacturing systems are not goals. They are just methods to create the target parts or products. So, manufacturing systems should be small as possible within satisfying requirements in the production. "Microfactory" was a concept of a future manufacturing system as an answer to the situation. It was proposed in the Japanese national R&D project named "Micro Machine Project [1]." The concept of the microfactory was very simple. The development team including one of the authors thought if it is possible to build "a super-miniature factory" for micro mechanical fabrication, environmental impact of manufacturing can be decreased greatly. In 1999, AIST developed the first prototype of a microfactory that consists of miniature machine tools and miniature manipulators. (Fig.1) The microfactory was able to perform a series of fabrication and assembly within a small desktop [2,3]. The result of the test production led us to conclude that the microfactory had considerable capability of micro mechanical fabrication.

As the results, "microfactory" has become a rather common concept. The development team insisted that the microfactory would reduce environmental impact and costs of "diverse-types-and-small-quantity production", "one-off production" or "variety-and-variant production". Since the smallness of the machines enables flexible layout changes, it can control the increase of the costs when the product designs have been modified. However, since there have been no effort to evaluate efficiency of the microfactory comparing with conventional factory quantitatively, the advantage to reduce environmental impact is still uncertain. The purpose of this report is to explain briefly about the microfactory and propose a simple and useful efficiency index to support system configuration design of microfactory-like system.

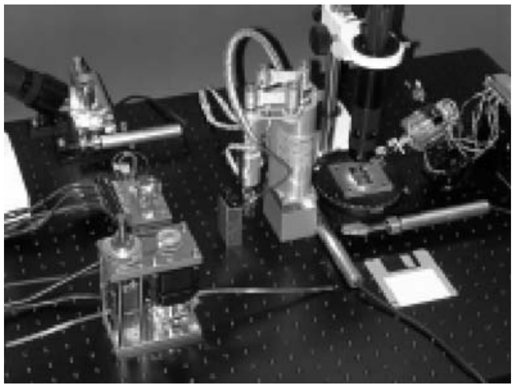


Figure 1: Developed Microfactory (1999, AIST)

2 PROPOSAL OF AN EFFICIENCY INDEX; TPI

2.1 TPI for products

Hereby, the paper tries to define an index to evaluate the efficiency of manufacturing process. In an existing research [4], the authors have proposed an index to evaluate real "eco-efficiency[5]" of products, by considering products' utility values, costs and environmental impacts, through product life-cycle. Efficiency index is defined by (1).

$$TPI = \frac{UV}{\sqrt{LCC}\sqrt{LCE}} \qquad (1)$$

TPI: total performance indicator
UV: utility value of the product
LCC: life-cycle cost of the product
LCE: life-cycle environmental impact of the product

Value per cost is often used to evaluate product performance in quality engineering. And additional value per environmental impact, so-called eco-efficiency is also a common index in recent design for environment [6], for evaluating another aspect of product performance. However, these existing evaluation indexes cannot evaluate the environmental and economical aspects simultaneously. In addition, since the "value" in the eco-efficiency is a fixed value, it cannot consider change of the value throughout the product life-cycle. The proposed index is the simplest combination of the both. In our proposal, because the utility value of the product can be expressed by integral of occasional value through life-cycle, it can simulate value deviation. So, our proposing TPI can be an answer to the problems in existing eco-performance indicators.

2.2 TPI for manufacturing processes

To evaluate the efficiencies of manufacturing processes or systems, the same idea can be applied. Because the design engineers and manufacturers have long histories of serious effort to reduce cost of manufacturing, they might not accept an index which does not evaluate cost and functionality. We define the total performance of the manufacturing process by (2). The equation expresses the balance of the product value created by the manufacturing process and cost and environmental impact necessary to fabricate "one" product.

$$Process\ \ TPI = \frac{UV}{\sqrt{\sum_{i=1}^{i=n} PE_i \cdot \sum_{i=1}^{i=n} PC_i}} \tag{2}$$

PCi: cost of the individual process
PEi: environmental impact of the individual process
n: number of processes

2.3 TPI for manufacturing systems

The final extension shown in (3) is to evaluate system efficiency by the same idea. The equation means that the efficiency of the system is defined by sum total of product values fabricated by the manufacturing system within a certain period of time. Cost and environmental impact during the corresponding period are also considered. Cost and environmental impact to build the manufacturing system itself (so-called initial cost and initial environmental impact) should be divided by lifetime duration of the manufacturing system, and assigned to PC and PE in the equation. However, it is not easy to quantitatively calculate (3), because when the system is for "variety-and-variant production" or "one-off production", and produces some different products, it is necessary to quantify all the value of the different products. But, fortunately, when the product of the manufacturing system is always the same, the system TPI equation can be simplified to (4).

$$System\ \ TPI = \frac{\int_{t0}^{t0+T} \sum_{i=1}^{i=m} \Delta UV_i dt}{\sqrt{\int_{t0}^{t0+T} \sum_{i=k}^{i=l} PE_i dt \cdot \int_{t0}^{t0+T} \sum_{i=k}^{i=l} PC_i dt}} \tag{3}$$

T: period of estimation
m: kinds of product,
k: number of the first process of in the system
l: number of the first process of in the system

$$System\ \ TPI = \frac{Tp}{\sqrt{\frac{(\frac{C_m}{L} + C_L)}{1600}} \cdot \sqrt{E_e + \frac{E_m}{L \times 1600}}} \tag{4}$$

Tp: throughput of the system
(number of the product produced in an hour),
Cm: initial total cost of the machines (yen)
C_L: labor cost (yen/year)
L: life of the system (year)
Ee: environmental impact caused by electricity
Em: environmental impact caused by the machines

Instead of the utility value of a product defined in the original index, throughput of the manufacturing system; "*Tp*" is introduced. By defining the throughput by number of products fabricated within an hour, the total performance indicator of the manufacturing system can be calculated. "*C*" can be calculated by using a sum total of machine costs, labor costs and other costs during the corresponding time.

3 TOTAL PERFORMANCE OF THE MICROFACTORY

3.1 Manufacturing process of the microfactory

The next step is to apply proposed efficiency index to actual manufacturing systems. As it was written in the beginning we are insisting miniature manufacturing systems are hopeful hints for environmentally benign manufacturing. So, the paper takes the microfactory shown in Figure 1, as an analysis example. It consisted of five miniature machines. The five components were a lathe, a milling machine and a press machine for parts fabrication, and a transfer arm and a micro-hand for assembly. Every component was designed rather specifically and extremely smaller than corresponding conventional item. To show the capability of the microfactory as a manufacturing system, trial production was executed. The trial product was a miniature ball bearing, which was 900 microns in diameter, and consisted of 4 kinds of parts. All the parts except steel balls were fabricated using the miniature machine tools, and assembled using the micro manipulator.

As the result of the fabrication, it can be said that the microfactory has possibility as a future manufacturing system which produces many varieties of micro mechanical products. However, the productivity of the system was very low. To improve the productivity, an appropriate system configuration should be considered. To fabricate the miniature ball bearing, manufacturing process shown in Figure 2 was applied.

Every part starts from the material shown in the upper side, passes through some sub-processes shown in the block and reaches the assembly processes written in the lower side. Among the processes, assembly processes were very time-consuming, because the process should be done sequentially under a microscopic vision.

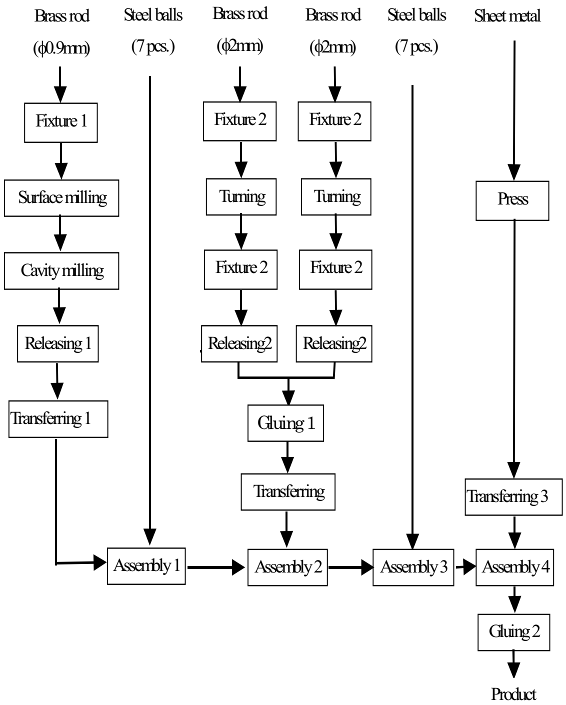


Figure 2: Manufacturing process used in the microfactory

3.2 Analysis of cost and process time

Table 1 indicates the average process time of the corresponding processes in Fig.4, after operators had been skilled enough. Number of operators required for each process is also shown in the table.

Table 1 : Required time for each process per unit

Process name	Process time	Num. of operators
Fixture 1	10 sec.	1
Fixture 2	5 sec.	1
Surface milling	1 min.	1
Cavity milling	2 min.	1
Turning	2 min.	1
Press	0.2 sec.	0
Releasing 1	10 sec.	1
Releasing 2	5 sec.	1
Transferring 1	1 sec.	0
Transferring 2	1 sec.	0
Transferring 3	1 sec.	0
Assembly 1	3 min. (per ball)	1
Assembly 2	3 min.	1
Assembly 3	3 min. (per ball)	1
Assembly 4	3 min.	1
Gluing 1	1 min.	1
Gluing 2	2 min.	1

According to Fig.5 and Table 1, it is evident that the assembly processes by the micro hand were the bottlenecks for the throughput. However, it is necessary to be aware that the flexibility of the process was assured by the ability of the micro hand. Machine costs are also necessary to calculate the TPI shown in (4). Table 2 shows the rough estimation for the cost of the machines used in the microfactory. And also the energy consumption of each machine is an important factor to consider system efficiencies. Table 3 shows the average power consumption of the machines during the operation. Both tables show that the micro hand was the most critical components for cost and energy, as well.

Table 2: Machine costs

Machine	Milling	Turning	Press	Arm	Hand
Cost (millionYen)	0.7	1.2	2.0	3.0	5.0

Table 3 : Energy consumption

Machine	Milling	Turning	Press	Arm	Hand
Average power consumption (kw)	0.25	0.3	0.05	0.2	0.4

4 ANALYSIS OF SYSTEM CONFIGURATION

Analysis of the manufacturing process mentioned in the former section showed that the assembly processes performed by “micro-hand” is critical both for throughput and environmental impact. When the number of components or operators is not limited to be 1, a simple strategy to enhance system efficiency will be to increase the number of the “hands.” By assuming the annual operation time of the system is 1600 hours, the system TPI shown in (4) can be calculated. By changing the system configuration, bottleneck will not always be the micro hand. Case studies are necessary. Figure 3 shows the behavior of the system efficiency calculated throughout the case studies.

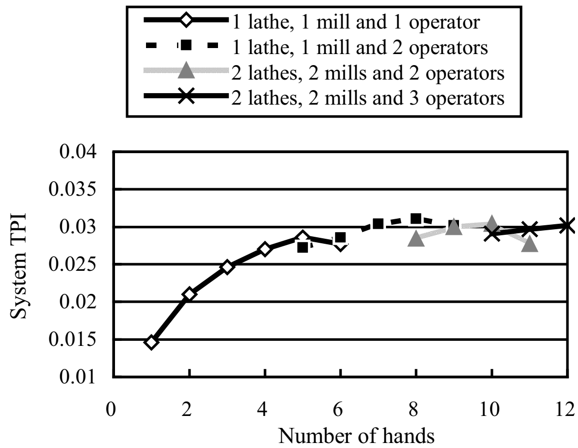


Figure 3 : Behavior of system TPI of the microfactory

In the figure, since press and transferring processes are not significant for the overall throughput, the figure shows the behavior of the system efficiency according to the change of the number of hands, lathes and mills. According to Figure 3, it can be said that there are some local maximums. The result suggested some simple strategies. Generally, the calculation shows that the system TPI has some peaks. It means that it is possible to optimize system configuration according to the evaluation result.

5 COMAPARISON WITH A MASS PRODUCTION SYSTEM

In the former section, effect of the system cinfiguration change has been considered. Focusing on diverse-types-and-small-quantity production of micro mechanical products, the final goal of this research is to prove that a microfactory-like system is more suitable than a large mass production manufacturing system. When frequent layout change of the system is necessary, high flexibility may cover low throughput. To prove this point, the paper compared the system TPI of the microfactory with that of a conventional mass production line. Figure 4 shows the block diagram of a typical manufacturing line for bearing.

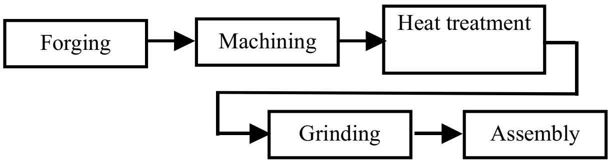


Figure 4: Commercial manufacturing process of ball bearing

A very rough estimation says the throughput of a commercial manufacturing line of bearing is about 200 thousand units per month. By assuming the line runs 20 days a month and 8 hours a day, it is equivalent to about 1200 units per hour. Another data says the initial cost of the manufacturing line is about 200 million yen, and the power consumption is estimated to be about 200kw. In the sumulation, we assumed that the manufacturing line change due to design change of the product occur rather frequently. Considering the environmental impact of machine material, die-set and electricity, the efficiency can be calculated. Figure 5 shows the rough estimation of the system TPI of the mass production system, corresponding to lifetime length of the system and average demand. The horizontal line in the figure shows TPI of the microfactory. TPI of the mass production system decreases with the decrease of demand and life of the system. On the other hand, TPI of the microfactory does not change remarkably.

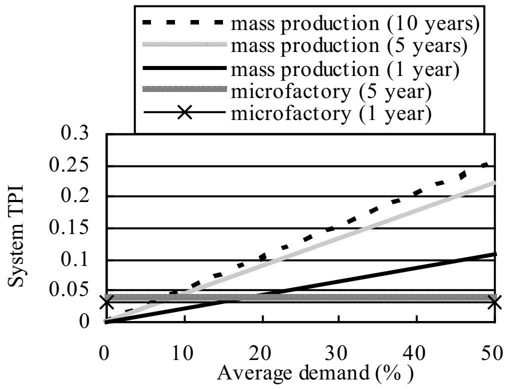


Figure 5: Comparison with a commercail system

As for micro mechanical parts used in mobile phones, electrical devices and so on, design change is rather frequent. After the design change it often happens that the demand is unusually high, because of «vertical strat-up. » According to an example of mobile phone, it is said that avewrage product life is about 6 months. If the demand for the product decreases linearly from the maximum towards zero during 6 months, and the life of the product line is 1 year, it means that the annual average demand is about 25%. In this case system TPIof the microfactory is almost same of ththat of the mass production system.

6 CONCLUSION

The microfactory developed by AIST was capable of micro mechanical fabrication. The capability was proved through the test production. However, although the advantage of the microfactory was said that it could reduce environmental impact of manufacturing, the actual effect of reducing impact has not been estimated quantitatively.

In this report, a simple index to evaluate the system efficiency of a microfactory-like system was proposed. The index (TPI) considered system throughput, machine costs, labour cost, and environmental impact by equivalent CO2 emission. As the results of the analysis of the test production process of the microfactory, it was able to show that the system had some suitable configurations, and it was possible to re-design system configuration of the microfactory.

The evaluation result was compared with a rough estimation of the efficiency of a conventional manufacturing system of ball bearings. The comparison indicated that the efficiency index of the microfactory was lower than that of a mass production system. However, it is necessary to consider recent short product lifecycle. By assuming that the lifetime of the system is relatively short and there are frequent design changes of the product, efficiency of the microfactory can be comparable or higher than that of a mass production line. And this can happen to products which use micro mechanical parts. The fact shows although "microfactory" is not a suitable system for mass production, it is a hopeful system for "diverse-types-and-small-quantity production. "

7 REFERENCES

[1] N. Kawahara, T. Suto, T. Hirano, Y. Ishikawa, N. Ooyama and T. Ataka, "Microfactories; New Applications of micro machine technology to the manufacture of small products," Microsystem Technologies, pp.37-41.

[2] N. Mishima, K. Ashida, T. Tanikawa and H. Maekawa, "Design of a Microfactory," Proc. of ASME/DETC2002, Montreal, Canada, 2002, DETC2002/DFM-34164.

[3] Y.Okazaki, Mishima N and Ashida K, "Microfactory - Concept, History and Developments-," Journal of Manufacturing Science and Engineering, Trans. ASME, Vol.126, 2004, pp.837-844.

[4] S. Kondoh, K. Masui, M. Hattori, n. Mishima and M. Matsumoto, "Total Performance Analysis of Product Life-cycle Considering the Deterioration and Obsolescence of Product Value," Proc. of CARE INNOVATION 2006, Vienna, Austria, 2006, 2.10.1.

[5] Y. Kudoh, K. Tahara and A. Inaba, "Environmental Efficiency of Passenger Vehicles: How Can the Value of a Vehicle be Determined?," Proc. of 7th International Conference on Ecobalance, Tsukuba, Japan, 2006, B3-7.

[6] M. Ernzer, M. Lindahl, K. Masui and T. Sakao, "An International Study on Utilizing of Design Environmental Methods (DfE) –a pre-study-," Proc. of Third International Symposium on Environmentally Conscious Design and Inverse Manufacturing, IEEE Computer Society, 2003, pp.124-131.

Environmental Burden Analysis for Machining Operation Using LCA Method

Hirohisa Narita[1], Nelfa Desmira[1] and Hideo Fujimoto[1]
[1] Graduate School of Engineering, Nagoya Institute of Technology
Gokiso-cho, Showa-ku, Nagoya, Aichi 466-8555, Japan

Abstract
We have developed an environmental burden analyzer for machine tool operations, which can evaluate an NC program from the view point of an environment burden by simulating a cutting process and using emission intensities. In this paper, global warming is selected as an impact category and the environmental burden analysis of dry, wet and MQL machining operations are carried out by using this analyzer with consideration of various situations. The influence of the peripheral devices of a machine tool, the spindle and the servo motors, the coolant, the lubricant oil, the cutting tool and the metal chips to global warming is introduced in detail.
Keywords:
Environmental burden, Machine tool, LCA, End mill operation

1 INTRODUCTION

Due to the recent requirement of the environmental protection, some approaches are tried for manufacturing processes [1, 2]. The concrete evaluation methods for machining operations, however, haven't been provided, yet. That is to say the detailed cutting conditions can't be evaluated by the conventional system.

In this research, an environmental burden analyzer for machine tool operations, which can calculates the accurate environmental burden, is developed based on LCA (Life Cycle Assessment) concept. Then, The influence of the peripheral devices of a machine tool, the spindle and the servo motors, the coolant, the lubricant oil, the cutting tool and the metal chips to global warming is analyzed in detail. The cutting fluid type and the disposal process of the cutting fluid are especially considered in this paper.

.

2 ENVIRONMENTAL BURDEN ANALYZER FOR MACHINE TOOL OPERATIONS

2.1 System overview

Conventional commercial analysis systems can evaluate just compare the differences of dry, wet, semi-dry machining, but this can't evaluate the differences of other cutting conditions such as depth of cuts, feed rate, tool path patter, etc. Furthermore, if the removal volume and the material type are same, environmental burden is calculated to same value. In other word, the conventional evaluation system can't provide enough information to decide the machining strategies. The analyzer developed in this research can solve the aforementioned problems.

Figure 1 shows an overview of the developed environmental burden analyzer for machining operations. When workpiece & cutting tool models and an NC program are input to the analyzer, all activities related to machining operation and machining process are estimated, and then the electric consumption of machine tool, the cutting tool status, the coolant quantity, the lubricant oil quantity and the metal chip quantity are calculated. The analyzer output environmental burden with using the background data and resource data, when a product is manufactured. Here, the background data means emission intensities to estimate the environmental burdens and the resource data means the data of machine tool spec, cutting tool spec. and so on.

This study focuses global warming, hence CO_2, CH_4 and N_2O are evaluated based on Japanese data. Here, global warming potential of 100 years [3] shown in table 1 is considered and equivalent CO_2 emission is evaluated as the environmental burden.

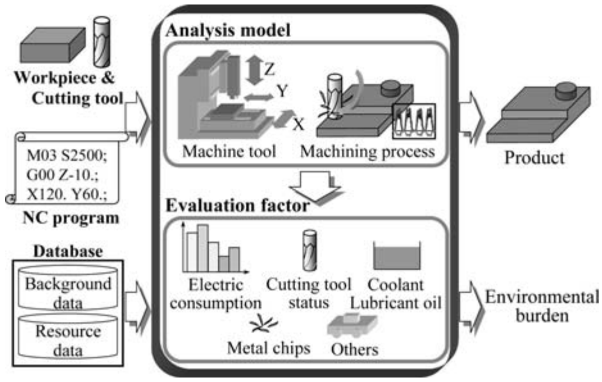


Figure 1: Processing flow of the developed environmental burden analyzer for machine tool operations.

Table 1: Characterization factors of global warming [3]

	CO_2	CH_4	N_2O
Global warming potential	1	21	310

2.2 Calculation algorithm for environmental burden

Total environmental burden is calculated from the following equation in this research.

$$Pe = Ee + Ce + LOe + \sum_{i=1}^{N} Te_i + CHe \quad (1)$$

Pe: EB of machining operation [kg-CO_2]
Ee: EB of machine tool component [kg-CO_2]
Ce: EB of coolant [kg-CO_2]
LOe: EB of lubricant oil [kg-CO_2]
Te: EB of cutting tool [kg-CO_2]
CHe: EB of metal chip [kg-CO_2]
N: Number of tool used in an NC program
EB: Environmental burden

Electric consumption of machine tool (*Ee*)

The environmental burden due to the electric consumption of machine tool is calculated as follows.

$$Ee = k \times (SME + SPE + SCE + CME + CPE + TCE1 + TCE2 + ATCE + MGE + VAE) \quad (2)$$

k: EI of electricity [kg-CO_2/kWh]
SME: EC of servo motors [kWh]
SPE: EC of spindle motor [kWh]
SCE: EC of cooling system of spindle [kWh]
CME: EC of compressor [kWh]
CPE: EC of coolant pump [kWh]
TCE1: EC of lift up chip conveyor [kWh]
TCE2: EC of chip conveyor in machine tool [kWh]
*ATC*E: EC of ATC [kWh]
EI: Emission intensity
EC: Electric consumption

In equation (2), the electric consumption of peripheral devices can be calculated from a running time. But one of servo and spindle motors is varied dynamically according to the machining process. Hence, a cutting force and a cutting torque models [4] are applied to estimate the electric consumption for these motors.

Coolant (*Ce*)

There are two types cutting fluid, so two equations are proposed for Ce evaluation.

First, water-miscible cutting fluid is explained. The coolant is generally circulated by coolant pump until coolant is updated. The cutting fluid is adhered to the metal chips, hence they are reduced bit by bit until the coolant is updated. Hence, the cutting fluid is supplied for compensation during this period. The dilution fluid (water) is also supplied at regular intervals due to loss through vaporization. Here, the environmental burden due to the coolant is calculated as follows by considering the aforementioned process.

$$Ce = \frac{CUT}{CL} \times \left\{ \begin{array}{l} (CPe + CDe) \times (CC + AC) \\ \qquad + WAe \times (WAQ + AWAQ) \end{array} \right\} \quad (3)$$

CUT: Coolant usage time in an NC program [s]
CL: Mean interval of coolant update [s]
CPe: EI of cutting fluid production [kg-CO_2/L]
CDe: EI of cutting fluid disposal [kg-CO_2/L]
CC: Initial coolant quantity [L]
AC: Additional supplement quantity of coolant [L]
WAe: EB of water distribution [kg-CO_2/L]
WAQ: Initial quantity of water [L]
AWAQ: Additional supplement quantity of water [L]

Second, water-insoluble cutting fluid is explained. In this case, discharge rate is an important factor. Hence, following equation is applied.

$$Ce = \frac{CUT \times CS}{3600 \times 1000} \times (CPe + CDe) \quad (4)$$

CS: Discharge rate of cutting fluid [cc/h]

Lubricant oil (*LOe*)

The lubricant oil is mainly used for a spindle and a slide way, so two equations are introduced. Minute amounts of oil is infused to the spindle part and the slide way in decided intervals. Hence, the following equations are adapted to calculate the environmental burden due to lubricant oil. Here, grease lubricant is not mentioned, but same equations can be adapted.

$$LOe = Se + Le \quad (5)$$

Se: EB of spindle lubricant oil [kg-CO_2]
Le: EB of slide way lubricant oil kg-CO_2]

$$Se = \frac{SRT}{SI} \times SV \times (SPe + SDe) \quad (6)$$

SRT: Spindle runtime in an NC program [s]
SV: Discharge rate of spindle lubricant oil [L]
SI: Mean interval between discharges [s]
SPe: EI of spindle lubricant oil production [kg-CO_2/L]
SDe: EI of spindle lubricant oil disposal [kg-CO_2/L]

$$Le = \frac{LUT}{LI} \times LV \times (LPe + LDe) \quad (7)$$

LUT: Slide way runtime in an NC program [s]
LI: Mean interval between supplies [s]
LV: Lubricant oil quantity supplied to slide way [L]
LPe: EI of slide way lubricant oil production [kg-CO_2/L]
LDe: EI of slide way lubricant oil disposal [kg-CO_2/L]

Cutting tool (*Te*)

Cutting tools are managed from the viewpoint of tool life. The cutting tools, particularly those for a solid end mill, are recovered by regrinding after reaching their life limit. In this study, environmental burden is calculated by comparing machining time with tool life and considering the aforementioned process.

$$Te = \frac{MT}{TL \times (TNR + 1)} \times ((TPe + TDe) \times TW + TNR \times RGe) \quad (8)$$

MT: Machining time [s]
TL: Tool life [s]
TPe: EI of cutting tool production [kg-CO_2/kg]
TDe: EI of cutting tool disposal [kg-CO_2/kg]

TW: Tool weight [kg]
T*NR*: Total number of recovery (re-grinding or dressing)
RGe: EI of re-grinding or dressing [kg-CO_2]

Metal chip (*CHe*)

Metal chips are accumulated, and coolant is separated from them. Then, they are recycled to material by electric heating furnace. This materialization process has to be considered. This equation is supposed to consider material kind, but electrical intensity of this kind of electric heating furnace is represent by kWh/t [5], so equation constructed in this research is calculated from total metal chip weight.

$$CHe = (WPV - PV) \times MD \times WDe \quad (9)$$

WPV: Workpiece volume [cm^3]
PV: Product volume [cm^3]
MD: Material density of workpiece [kg/cm^3]
WDe: EI of metal chip processing [kg-CO_2/kg]

3 NUMERICAL EXAMPLE

3.1 Parameter setting

Emission intensities related to evaluation factors are summarized in table 2. These values are listed from some reports, which are environmental report, technical report, home page and industrial table [5, 6, 7, 8, 9, 10, 11, 12]. About these data, production and disposal processes are considered and transportation and other processes are ignored, because these values are different by users.

We would like to focus the coolant influences to the environment in this paper, hence the emission intensities related to cutting fluid is summarized in detail. Regarding to the water-miscible cutting fluid, the disposal process related to A1, A2 and A3 types, and the distilling and condensing process are considered. The distilling and condensing process, which is often used recently, can be applied for all type of the water-miscible cutting fluid.

In order to decide the emission intensities of them, the electric consumption of the processing devices and the emissions due to the thermal disposal are investigated. Regarding to the water- insoluble cutting fluid, the thermal and the material recycles are considered. For the thermal recycle, the CO_2 emission of the disposal process is reduced by the electric consumption due to the calorific value of the waste oil. For the material recycle, the emission due to the thermal disposal is eliminated.

3.2 Comparison of cutting method

For the numerical example, the machine tool is MB-46VA (OKUMA Corp.), the cutting tool is carbide-square end mill with 10 mm diameter, two-flute and a 30° helical angle, and the workpiece is mideum carbon steel (S50C). The parameters of MB-46VA peripheral devices are shown in table 3 and other parameters input by machine tool users shown in table 4.

Figure 2 shows the product shape and the tool path pattern used for the evaluation. The spindle speed is 2500 rpm and the feed rate is 200 mm/min. Here, the tool life is assumed to be increased to twofold the original one due to the coolant effect.

Table 2: CO_2 Emission intensities

Electricity[kg-CO_2/kWh]	0.381
Cutting fluid production [kg-CO_2/L]	0.469
Cutting fluid disposal (water-miscible cutting fluid;A1 type) [kg-CO_2/L]	3.782
Cutting fluid disposal (water-miscible cutting fluid;A2 type) [kg-CO_2/L]	5.143
Cutting fluid disposal (water-miscible cutting fluid; A3 type) [kg-CO_2/L]	8.103
Cutting fluid disposal (distilling and condensing process) [kg-CO_2/L]	3.425
Cutting fluid disposal (water-insoluble cutting fluid: normal) [kg-CO_2/L]	2.555
Cutting fluid disposal (water-insoluble cutting fluid: thermal recycle) [kg-CO_2/L]	1.778
Cutting fluid disposal (water-insoluble cutting fluid: material recycle) [kg-CO_2/L]	0.261
Dilution fluid (water) [kg-CO_2/L]	0.189
Spindle and slide way lubricant oil production [kg-CO_2/L]	0.469
Spindle and slide way lubricant oil disposal [kg-CO_2/L]	0.0029
Cutting tool production [kg-CO_2/kg]	33.7478
Cutting tool disposal [kg-CO_2/kg]	0.01346
Re-grinding [kg-CO_2/number]	0.0184
Metal chip processing [kg-CO_2/kg]	0.0634

Table 3: Electric consumption of peripheral devices

NC controller [kW]	0.16
Cooling system of spindle [kW]	0.45
Compressor[kW]	1
MQL compressor[kW]	0.75
Coolant pump[kW]	0.25
Lift up chip conveyor[kW]	0.1
Chip conveyor in machine tool [kW]	0.6
ATC [Wh]	0.08
Tool magazine[Wh] (1 round)	0.087
Vampire energy[kW]	0.64

Table 4: Other parameters related to evaluation factor

Initial coolant quantity [L]	8.75
Additional supplement of coolant [L]	4.3
Total quantity of dilution fluid [L]	257.25
Mean interval between replacements of coolant in pump [Month]	5
Discharge rate of spindle lubricant oil [mL]	0.03
Mean interval between discharges for spindle lubrication [s]	480
Discharge rate for coolant of MQL[mL/h]	7
Lubricant oil supplied to slide way[mL]	228
Mean interval between supplies [hour]	2000
Tool life [s]	5400
Total number of re-grinding	2
Material density of cutting tool [g/cm^3]	11.9
Material density of workpiece [g/cm^3]	7.1
Air blow for MQL [NL/min]	150
Air pressure of MQL [MPa]	0.7

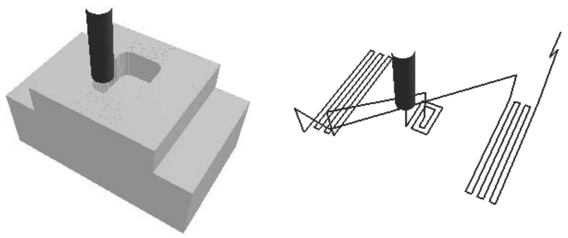


Figure 2: Machined shape and tool path pattern

Analyzed results are shown in figure 3. The MQL machining is preferable for this machining operation. When we use this evaluation system, we can decide the cutting conditions realizing the minimum environmental burden.

The electric consumption of the peripheral devices of the machine tool, except for the servo motors and the spindle motor and the cutting tool are the main factors in the machining operation from the viewpoint of global warming.

Anyway, the emission intensities of the coolant are very different and not small, however the environmental burden of the coolant is small. It is found that the values of the peripheral devices and the cutting tool must be reduced in order to reduce the equivalent CO_2 emission effectively.

Futhermore, the comparison of the impacts of CO_2, CH_4 and N_2O is carried out. CH_4 and N_2O emissions are occured by the thermal disposal. The impacts of CH_4 and N_2O calculated from table 1 correspond to below 0.001 g-CO_2 by comparing the amount of the emission of each condition. In other words, CO_2 is the dominant environmental burden in the machining operation concerning global warming.

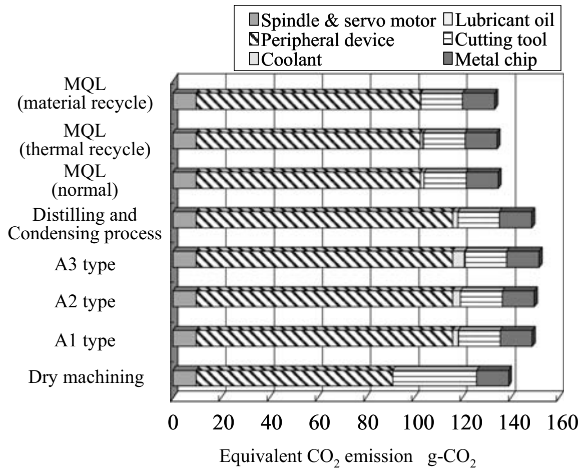


Figure 3: Analyzed environmental burden results of numerical example

4 CONCLUSIONS

1) The environmental burden analyzer for machine tool operations was introduced.
2) The influence of the peripheral devices of a machine tool, the spindle and the servo motors, the coolant, the lubricant oil, the cutting tool and the metal chips to global warming was analyzed. The electric consumption of the peripheral devices of the machine tool and the cutting tool is the main factors in the machining operation when an impact category is global warming.
3) CO_2 is the dominant environmental burden in machining operation concerning global warming by comparing CO_2 emission with the equivalent CO_2 emission of CH_4 and N_2O.

5 REFERENCES

[1] Sheng, P., Bennet, D., Thurwachter, S., / B.F. von Turkovich, 1998, Environmental-Based Systems Planning for Machining, Ann. of the CIRP, 47/1, 409-414.

[2] Shimoda, M., LCA Case of Machine Tool, Symposium of 2002 Japan Society for Precision Engineering Spring Annual Meeting "Leading-Edge Trend of Environmental Impact Evaluation for Inverse Type Design and Manufacturing", (2000), 37-41, (in Japanese).

[3] Japan Meteorological Agency, 1995, Climate Change 1995, Science of Climate Change.

[4] Narita, H., Shirase, K., Wakamatsu, H., Tsumaya, A. and Arai, E., 2002, Real-Time Cutting Simulation System of a Milling Operation for Autonomous and Intelligent Machine Tools, Int. J. of Production Research, 40, 15, 3791-3805.

[5] Mizukami, H., Yamaguchi, R., Nakayama, T., Maki, T., 2002, Off-gas Treatment Technology of ECOARC, NKK Technical Report, No.176, 1-5, (in Japanese).

[6] Tokyo Electric Power Company, 2005, The Earth, People & Energy TEPCO Sustainability Report 2004.

[7] Tokyo Waterworks, 2003, Environmental Report of Tokyo Waterworks 2002. (in Japanese)

[8] Nansai, K., Moriguchi, Y. and Tohno, S., 2002, Embodied Energy and Emission Intensity Data for Japan Using Input-Output Tables (3EID)–Inventory Data for LCA-, Center for Global Environment Research, National Institute of Environmental Studies, Japan.

[9] New Energy and Industrial Technology Development Organization (NEDO), 2003, Development of Environment-Conscious Technology Laboratory, First Section Meeting of "Development of Energy Saving Wastewater Treatment Technique", Program Original Register of Development of Energy Saving Wastewater Treatment Technique, (in Japanese)

[10] Ministry of Economy, Trade and Industry, 2002, About Declaration of Evaluation Report related to Budget Request in Fiscal Year 2002 (Policy Evaluation), Section 6 Construction of Industrial Facilities (Regional Economic Industry U), Development of Wastewater Treatment Technique, (in Japanese).

[11] Osaka prefecture, 2004, Osaka prefectural statistical yearbook, (in Japanese)

[12] Touma, S., Ohmori, S., Kokubo, K., Tateno, M., 2003, Evaluation of Environmental Burden in Eco-friendly Machining Method using Life Cycle Assessment Method – Estimation of Carbon Dioxide Emission in Eco-friendly Turing Method– J. of JSPE, 69, 6, 825-830, (in Japanese).

Efficient Maintenance of Machine Tools – Adapted Maintenance Activities and Assembly-Specific Maintenance Intervals

Jürgen Fleischer[1], Jan Wieser[1], Matthias Schopp[1], Alexander Broos[1]
[1] Institute of Production Science (wbk), Universität Karlsruhe (TH), Germany

Abstract
Besides purchasing costs of modern production equipment, guaranteed availabilities and maximum costs over the expected time of usage become more important. In order to meet these required availability limits, maintenance represents a key function as low MTTR and high MTBF values depend on systematic maintenance measures. Looking at maintenance policy in today's industry, the responsibility for maintenance is not only the duty of the service department but also a task of the machines' operators, who can be trained for minor maintenance tasks. The division of the responsibilities between maintenance experts and operators depends on the complexity of the respective tasks.
Keywords:
Machine Tool, Maintenance, Service Engineering

1 INTRODUCTION AND MOTIVATION

Increasing costs for modern production equipment oblige operators to use their machines most efficiently [1]. From a technical point of view, availability is key in offering great potential [2, 3]. Availability could potentially be increased by carrying out regular and preventive maintenance activities in order to avoid machine downtimes [4].

2 OBJECTIVES

The project "Maintenance-friendly machine tools – simple and robust maintenance activities to increase availability" aims at increasing machine tool availability by adapting maintenance activities and specific maintenance intervals to operating conditions. Availability can thus be improved by increasing MTBF (Mean Time Between Failures) for maintenance activities and by reducing MTTR (Mean Time To Repair) for servicing and repair activities (figure 1). This leads to a more efficient outcome of maintenance activities and allows for the use of TPM (Total Productive Maintenance) [5, 6].

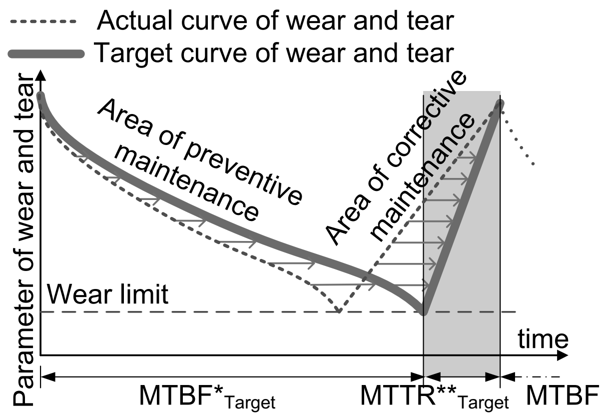


Figure 1: Objectives.

3 APPROACH

First of all, a modularisation approach for machine tools will be developed which takes into account how much effort is needed to carry out preventive and corrective maintenance activities for relevant components. This allows us to assess individual maintenance activities using the index for maintenance friendliness (IMF). The algorithm for the identification of collateral components (AIC) identifies components whose early replacement within upcoming maintenance activities is reasonable from an economic point of view. IMF and AIC were implemented into a software environment in order to simplify the application of these two methods. A detailed analysis of machine downtime causes during operation allows for the adaptation of the manufacturer's maintenance instructions to the respective operating conditions. The resulting measures for optimising preventive and corrective maintenance activities and intervals will be implemented into the machine control software. Furthermore, constant inspections monitor machine state and provide the basis for the calculation of a corrective maintenance index continuously assessing their efficiency. This model of intelligent maintenance can systematically increase availability (figure 2).

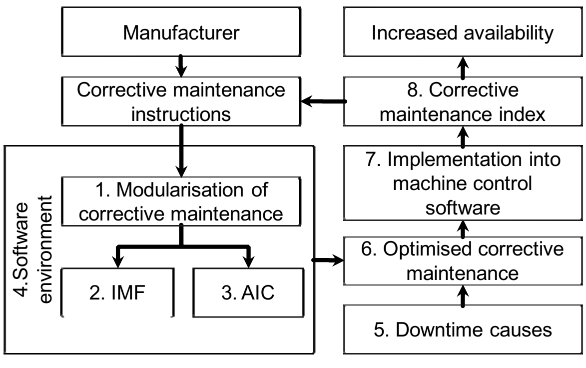


Figure 2: Approach.

3.1 Modularisation of maintenance activities

A modularisation approach for machine tools was elaborated which takes into account how much effort is needed to carry out preventive and corrective maintenance activities for relevant components [7]. Maintenance instructions that were provided by the manufacturer with delivery can thus be assigned to three machine tool levels: components, assemblies and the machine as a whole. The interdependence of components and assemblies in the event of replacement is taken into account. The time needed for identifying collateral components and replacing them at a subsequent point in time can thus be determined. The required time will either be determined by experts or, in terms of activities carried out regularly, by established time measurement methods such as MTM [8] or REFA (German Association for work design, industrial organisation and company development) [9].

3.2 Index for maintenance friendliness (IMF)

In order to identify potentials to increase maintenance friendliness, an index for corrective maintenance friendliness was elaborated (IMF) [10, 11]. The IMF allows for assessing the maintenance friendliness of an individual activity, component, assembly or the machine tool as a whole. The outcome of this assessment is a numerical value ranging from 0 % (maintenance-unfriendly) to 100 % (maintenance-friendly).

3.3 Algorithm for the identification of collateral components (AIC)

The AIC is used to identify so called collateral components. According to maintenance instructions, these components do not have to be replaced yet, but from an economic point of view it might be reasonable to replace them ahead of schedule when replacing a neighbouring component. The time needed for another dismantling of the assemblies sitting above these two components is calculated. Not only needs the primary time needed for carrying out the actual maintenance activity to be taken into account, but the secondary time as well. This is the time needed to remove components sitting above the elements to be replaced, such as covers for example. As opposed to [12, 13, 14], this approach takes the existing interdependence of components and assemblies into account when replacing them as part of regular maintenance activities. Another aspect to be taken into account when identifying collateral components is, besides additional time and staff needed, the respective remaining value of the component, which is left unused when replacing it ahead of schedule [7].

3.4 Software environment

Existing data from various sources are used for the application of this method (e.g. corrective maintenance instructions, spare parts lists, life-cycle specifications). A software environment was thus developed in MS Visio, in order to ensure a simple and rapid implementation of the steps presented above [15].

3.5 Downtime causes

Based on a field data analysis of machines' operating behaviour assemblies and components are identified according to [16], whose failure behaviour is key to machine availability. The aim is to improve the failure behaviour of these components by adapting corrective maintenance instructions to the operating conditions in place and thus to increase machine availability. Analyses of downtime causes based on these results and the identification of potential improvements provide, together with individual experiences of servicing and maintenance staff for the analysed machine type, the basis for optimised maintenance activities.

ABC-analysis II – Frequency as per component		ABC-analysis I – Total downtime as per component: A serious / significant	B medium	C low
A	frequent	1. Significant total downtime & Frequent occurence	2. Medium total downtime & Frequent occurence	3. Low total downtime & Frequent occurence
B	Not too frequent /several times	4. Significant total downtime & Several donwtimes	5. Medium total downtime & Several donwtimes	6. Low total downtime & Several donwtimes
C	rare	7. Significant total downtime & Rare occurence	8. Medium total downtime & Rare occurence	9. Low total downtime & Rare occurence

Figure 3: Pareto approach.

The Institute of Production Science (wbk) developed an pareto chart based on 2D ABC analysis that can be used to introduce the results into a matrix and compare them. The ABC analysis categorises downtime along the horizontal axis and downtime frequency along the vertical axis (figure 3).

A component is considered to be significant in its impact on availability if it figures in at least one of the two ABC analyses in class A and in none of the two ABC analyses in class C. In the ABC 2D analysis this corresponds to the fields 1, 2 or 4.

ABC-analysis II – Frequency [times]		ABC-analysis I – Total downtime [h]: A serious / significant	B medium
A	frequent	▪ Tooling system [3193,6h; 717] ▪ Machine control [2101,0h; 580] ▪ Turret [2053,1h; 326] ▪ Limit switch, sensor, proximity switch [1518,9h; 624] ▪ Feeding device [1919,4h; 316]	▪ Chip conveyor [1121,9h, 318] ▪ Coolant supply [907,3h; 277]
B	Not too frequent, several times	▪ Lines, fittings [1426,2h; 180]	

Figure 4: Results of the pareto.

The results shown in figure 4 stem from 33 lathes of the same type that are in operation in a single company. In the following these are seen as a single machine with a single period of operation. The resulting overall period of operation is based on the sum of individual periods of operation and therefore amounts to 193.64 years. Due to this company's shift schedule the theoretic overall period of operation of all machines amounts to 145.40 years. The overall downtime of all assemblies (27259h, app. 3 years) equals 100%. This is the result of 5860 error messages.

3.6 Optimised maintenance

The identified components are assessed according to their potential for optimisation (figure 5). The aim is to increase the availability of the whole machine by optimising the maintenance of these components. Every component is thus examined for its potential for optimisation along the lines of the four basic measures of machine maintenance which are preventive maintenance, inspection, corrective maintenance and improvement [17]. Based on their potential for optimisation components are analysed on whether it is reasonable for example to optimise them rather within the framework of preventive maintenance or as part of other basic corrective maintenance activities. Components can be classified into electronic and mechanical components. Electronic components (machine control, limit switch, sensor, proximity switch) usually offer less potential for optimisation than mechanical components. One reason is the failure behaviour of electronic components which is often referred to as digital and which represents an unpromising starting point for improvement. In practice, this can often be ascribed to the insufficient predictability of downtimes.

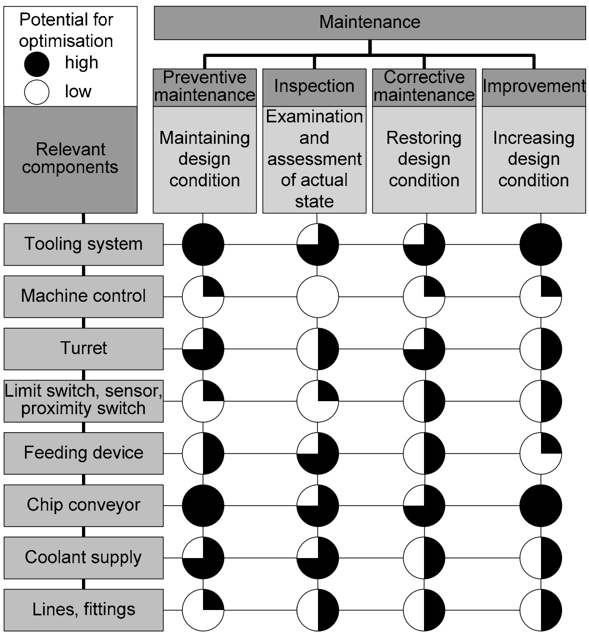


Figure 5: Potential for optimisation of relevant components.

The Institute of Production Science (wbk) developed a classification system in order to analyse the relevant components according to their downtimes. Potential causes for downtimes are restricted to the four most relevant downtime causes in terms of machine maintenance.

- **Soiling** is seen as part of an unintentional machine malfunction caused by a foreign object, e.g. chips.
- **Operating errors** include all downtime causes resulting form incorrect machine operation, e.g. incorrect CNC programming.
- **Consequential damage** occurs when downtime is caused by another component, e.g. by a defective pneumatic hose line (component: lines, fittings). This can lead to palletisation downtime (component: tooling system)
- **Tear and wear** includes surface abrasion of a component through frictional, grinding, rolling, smiting, scratching, chemical and thermal strain.
- Should error messages occur that cannot definitely assigned to a specific failure, these are assigned to the position "**no classification**". This position, however, should only be used in emergency cases.

A cluster analysis visualises the results (figure 6). The abscissa shows downtime causes, whereas the ordinate displays components. The frequency of downtime causes for each component (in percent) is indicated by bubble diameter. The operating behaviour documented above is based on a sufficient amount of field data and thus provides the basis for optimising existing instructions.

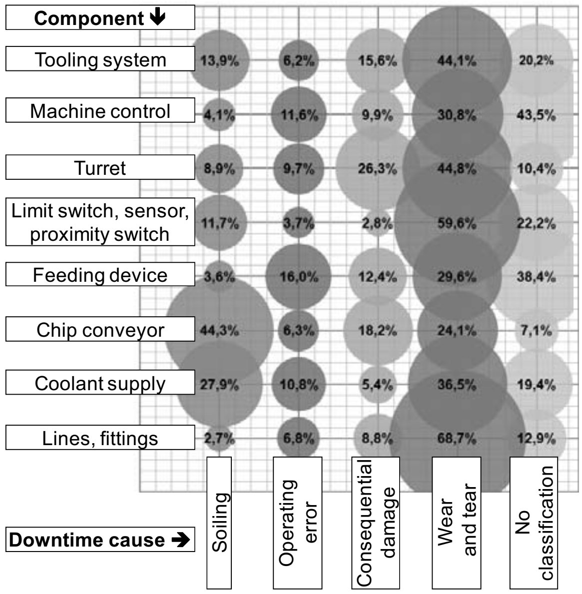


Figure 6: Downtime causes of the relevant components

The endeavour is to efficiently increase machine tool availability by adapting maintenance activities to operating conditions. The downtime causes "soiling" and "tear and wear" in particular can lead to changes in maintenance activities and intervals. Based on these findings and after consultation with the operator and the manufacturer, the respective instructions were elaborated. Optimised preventive and corrective maintenance instructions include the following information:

1. Activity: short description of the measure to be taken
2. Number: consecutive numbering of instructions

Table 1: Optimised preventive and corrective maintenance instructions sheet

1. Activity			2. Number
			3. Classification
			4. Machine operating mode
5. Interval	6. Duration	7. Number of persons	8. Service group
9. Caution!	specific aspects of this instruction		
10. Place	place of activity to be carried out		
11. Instruction	1.		

3. Classification: preventive maintenance, inspection or corrective maintenance [17]
4. Machine operating mode: Distinction of On – Set-up mode – Off
5. Interval: time specification for interval duration in min – h – a
6. Duration: time specification for duration in min – h – a
7. Number of persons: the number of persons needed for carrying out maintenance activities
8. Service group: in charge of maintenance activity depending on the level of difficulty: untrained employee, machine operator, maintenance staff (mechanics), maintenance staff (electrics), and hybrids [10].
9. Caution: consideration of specific framework conditions
10. Place: indication of the place where the activity is to be carried out
11. Instruction: description of the activity

The written instructions focus on a consistent documentation that is easy to understand for the employee in charge (table 1). Instructions may be completed by graphics and technical drawings.

3.7 Implementation into machine control software

In order to ensure that maintenance activities are carried out at the scheduled time, the optimised instructions were implemented into a software, which allows for a visualisation of maintenance instructions on the machine control screen. As part of the TPM maintenance concept, machine operators shall be increasingly included in machine maintenance by carrying out maintenance-friendly activities. As opposed to [18, 19], the Siemens user interface was designed in analogy to the 840D control interface (figure 7), in order to facilitate machine introduction and training for the user.

The user interface contains the same information as shown in table 1. In order to do justice to the TPM approach, a signal using the traffic light colours red – yellow – green was created which is shown on the control interface at all times within the machine operator's sight. This signal indicates, depending on the time required for issuing a maintenance warning for specific activities, which activities are to be carried out on the machine next.

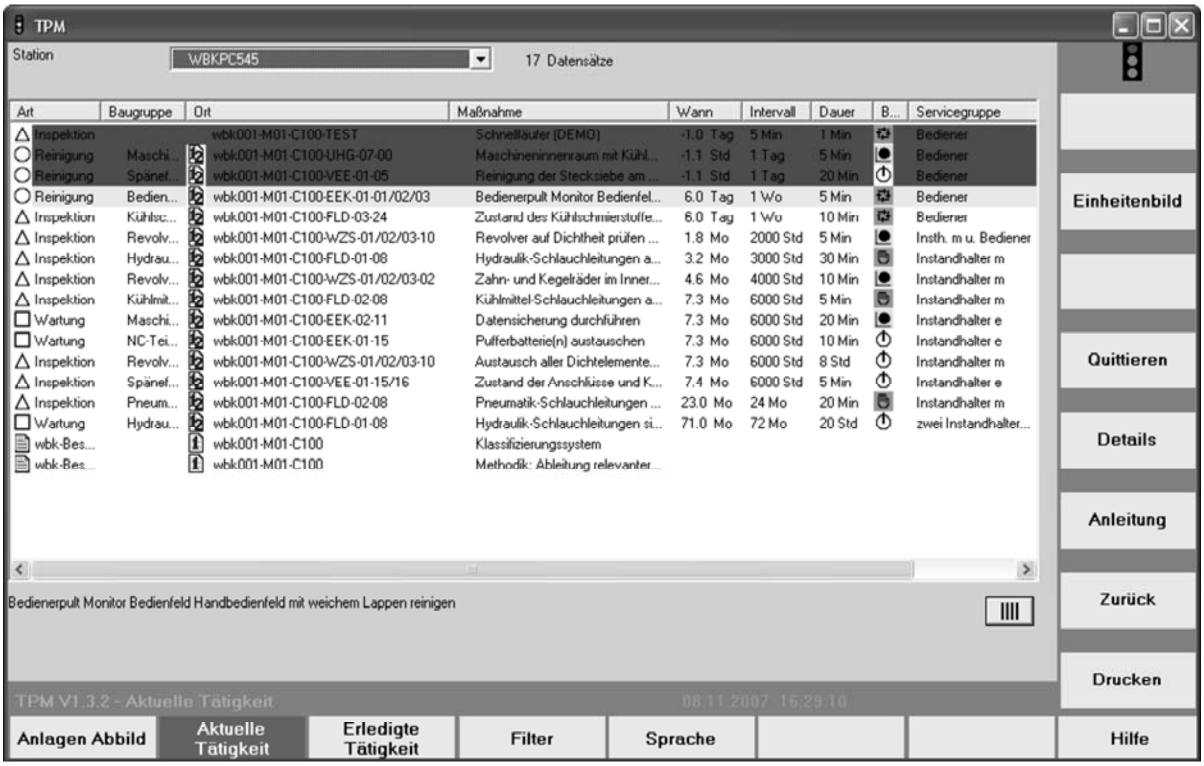


Figure 7: User interface Siemens MCIS TPM.

As long as the period of warranty applies, the manufacturer is particularly interested in the list of activities carried out on the machine in order to control the activities carried out by the operator in the event of damage. Furthermore, in analogy to individual service groups, user accounts are created in order to take corporate framework conditions into account. Among these figure for example adapting the interval and duration of an activity to the corporate framework conditions.

3.8 Corrective maintenance index based on operational behaviour

The efficiency of optimised preventive and corrective maintenance instructions in view of machine availability is assessed by an index. This represents a decisive step towards a reliable maintenance concept (RCM – Reliability Centered Maintenance) [20, 21]. The corrective maintenance index is the result of three essential criteria in order to assess the overall machine state:

- A. Availability: Machine availability is assessed for a specific period of time using downtime figures (i.e. by using production data acquisition systems such as SAP).
- B. Profitability: SAP determines machine tool costs for a specific period of time.
- C. Machine state assessment: the technical machine state is assessed as part of an inspection.

A scoring method for single criteria is used to determine the index. The targeted and continuous monitoring and assessment of the overall machine state enables us to carry out corrective maintenance activities no earlier than actually required, and this with a limited amount of risk (due to short machine monitoring intervals) and effort (i.e. through the targeted use of automated condition-monitoring-systems [22] in respective places). First observations, made by the wbk on a machine fitted with this software, provide very promising results.

4 SUMMARY

This article presented a method for intelligent machine tool maintenance. Based on a modularisation approach for corrective machine tool maintenance, the algorithm elaborated by the wbk enables us to assess machine tools' corrective maintenance friendliness with the IMF and to identify collateral components with the AIC. Both IMF and AIC were implemented into a software environment to allow for an easier application of both methods. Downtime causes of machine tools during operation were analysed in detail including the components that are relevant for machine tool availability. These figures provided the basis for adapting the corrective maintenance instructions issued by the machine manufacturer to the respective operational conditions and for optimising these instructions. These were implemented into the machine tool control software. The machine state is permanently monitored by inspection staff and its efficiency constantly assessed by the corrective maintenance index. The approach presented in this paper paves the way for systematically increasing overall machine tool availability by focusing on the targeted improvement of components relevant for availability.

5 ACKNOWLEDGMENTS

This article is publicizing a research project of the Institute of Production Science (wbk), Universität Karlsruhe (TH), which is financially supported by the KSB-Stiftung, Frankenthal, Germany. With our best thanks to the KSB-Stiftung for the support.

6 REFERENCES

[1] Takata, S., Kimura, F, Van Houten, F.J.A.M., Westkämper, E., Shpitalni, M., Ceglarek, D., Lee, J., 2004, Maintenance: Changing Role in Life Cycle Management, Annals of the CIRP, 53/2:1-13.

[2] Fleischer, J., Nesges, D., 2005, Identifying Availability Contribution of Life-Cycle-adapted Services, CIRP Life Cycle Engineering.

[3] Fleischer, J., Niggeschmidt, S., Wawerla, M., 2007, Optimizing the Life-Cycle-Performance of Machine Tools by Reliability and Availability Prognosis, CIRP Life Cycle Engineering, 329-334.

[4] Schuh, G., Grawatsch, M., 2005, Trends im Instandhaltungsmanagement. wt Werkstattstechnik online 95, 7/8:590-591.

[5] Al-Radhi, M., 2002, Total Productive Management, ISBN 3-446-21873-4, München, Wien: Hanser Verlag.

[6] Nakajima, S., 1995, Management der Produktionseinrichtungen, Total Productive Maintenance, Frankfurt: Campus Verlag.

[7] Fleischer, J., Hennrich, H., Wieser, J., 2008, Instandhaltungsfreundliche Werkzeugmaschinen, wt Werkstattstechnik online 98, 1/2:002-008.

[8] N.N., 1993, Verband für Arbeitsstudien und Betriebsorganisation e.V., Fachbuchreihe Betriebsorganisation. Ausgewählte Methoden des Arbeitsstudiums, München.

[9] D/JB, Deutsche MTM- Vereinigung e.V., 2002, MTM-Grundverfahren, Lehrgangsunterlagen.

[10] Fleischer, J., Wieser, J., Hennrich, H., Schopp, M., Broos, A., 2007, Entwicklung eines Bewertungssystems für Instandhaltungstätigkeiten an Werkzeugmaschinen, ZWF Zeitschrift für wirtschaftlichen Fabrikbetrieb 102, 11/12:756-764.

[11] Fleischer, J., Wieser, J., Schopp, M., Broos, A., 2008, Availability Increase of Machine Tools through Adapted Maintenance Activities and Assembly-Specific Maintenance Intervals, CIRP Life Cycle Engineering.

[12] Heisel, U., Schleich, B., 2006, Bewertungsverfahren zur Beurteilung der Instandhaltbarkeit in der Entwurfsphase, wt Werkstattstechnik online 96, 4:239-242.

[13] Deuse, J., Mende, K-M., Birkmann, S., 2006, Optimierung von Instandhaltungsumfängen, wt Werkstattstechnik online 96, 7/8:540-547.

[14] Deuse, J., Mende, K-M., Birkmann, S., 2007, Optimierung von Instandhaltungsumfängen – Teil 2, wt Werkstattstechnik online 97, 3:183-187.

[15] Fleischer, J., Wieser, J., Hennrich, H., Schopp, M., Broos, A., 2008, Optimization of Maintenance Activities

through Maintenance-Relevant Design, CIRP Design Conference.

[16] Fleischer, J., Broos, A., Schopp, M., Wieser, J., 2007, Datenbasis für lastabhängige Prozesseingriffe. wt Werkstattstechnik online 97, 7/8:491-497.

[17] N.N., DIN 31051, 2003, Grundlagen der Instandhaltung. Beuth Verlag GmbH, Berlin.

[18] Van Houten, F.J.A.M., Kimura, F., 2000, The Virtual Maintenance System: A Computer-Based Support Tool for Robust Design, Product Monitoring, Fault Diagnosis and Maintenance Planning, Annals of CIRP, 49/2:91-94.

[19] Takata, S., Inoue, Y., Kohda, T., Hiraoka, H., Asama, H., 1999, Maintenance Data Management System, Annals of CIRP, 48/1:389-392.

[20] Wald, G., 2003, Prozessorientiertes Instandhaltungsmanagement, ISBN 3-8322-1821-1, Aachen: Shaker Verlag.

[21] Niemeier, F., 2001, Prozesszustandsorientierte Verfügbarkeitslenkung von Produktionsanlagen. ISBN-3-8265-9252-2, Aachen: Shaker Verlag.

[22] Fleischer, J., Schopp, M., 2007, Sustainable Machine Tool Reliability based on Condition Diagnosis and Prognosis, CIRP Life Cycle Engineering.

CONTACT

Prof. Dr.-Ing. Jürgen Fleischer (fleischer@wbk.uka.de),
Dipl.-Ing. Jan Wieser (wieser@wbk.uka.de),
Dipl.-Ing. Matthias Schopp (schopp@wbk.uka.de),
Dipl.-Ing. Alexander Broos (broos@wbk.uka.de)

Universität Karlsruhe (TH),
Institute of Production Science (wbk),
Kaiserstr. 12,
76131 Karlsruhe, Germany

Manufacturing System Organization

Strategy-oriented Qualification Framework as a Supporting Function of Lean Production System Implementation in Small and Medium-sized Enterprises

Uwe Dombrowski[1] and Isabel Crespo[1]

[1] Institute of Production Management and Operations Research, Technical University Carolo-Wilhelmina of Braunschweig, Germany

Abstract

Small and Medium-sized Enterprises (SMEs) are facing the hard competition of global markets and the more specific and higher requirements of the customers everyday. In order to cope with these challenges many enterprises implement a lean production system (LPS). For the implementation of a LPS a continuous support of a well-structured qualification background is necessary. This paper presents a strategy-oriented qualification framework which has been developed and successfully implemented in a publicly financed research project in close cooperation with six SMEs. This framework includes a strategy-oriented scheduling of the necessary qualification activities before and during the LPS implementation, a concept to adapt the qualification contents to the different needs from the entrepreneur down to the shop-floor worker, and various innovative approaches (e.g. cascade-training) to realise the implementation.

Keywords:

Lean Production System Implementation, Qualification, Small and Medium-sized Enterprises

1 INTRODUCTION

Changes in the business environment as shorter life cycles, higher product variety, fluctuations of the production volume, rapidly changing technologies, as well as the customers' demand for low prices, and short lead times, force Small and Medium-sized Enterprises (SMEs) to improve their processes and organization [1]. In addition to the general business conditions, large enterprises expect from SMEs to be able to cope with its requests on flexibility and high product quality permanently [2]. Therefore SMEs have to use not only recent developments in production, information and communication technology but also have to apply current organizational concepts [3].

One of the successful strategies to deal with these changes and requirements is the implementation of a lean production system (LPS). This term, coined in *The Machine That Changed the World*, emblematises the efforts of many American and European production enterprises to copy and adapt the well-known and successful Toyota Production System which had been developed from the founder of Toyota, Sakichi Toyoda and the engineer Taiichi Ohno [4]. "Lean Production" is widely considered the next big step in the evolution of manufacturing beyond Ford´s mass production. Who would have realized that Sakichi Toyoda, working in the rural hinterlands of what is now Toyoda City in Japan, would have developed a global concept that has changed the face of manufacturing? [5].

Lean production systems do not only help to reduce waste in the production process but also allow the enterprise to focus on customer value [6], [7], and [8]. A lean production system can be defined as an enterprise-specific compilation of rules, standards, methods and tools, as well as the appropriate underlying philosophy and culture for the comprehensive and sustainable design of production. An LPS enables an enterprise to meet the requirements of today´s business environment, taking into account technological, organizational, workforce-related and economic aspects [9]. This definition supports a system approach to lean production [10] and [11], in which LPS are described by hierarchical connections of different elements. The structure of the system has three levels. First of all, the main objectives of the enterprise that directly address the customers´ demands are formulated on the first level of the system (e.g. minimize manufacturing costs). These objectives are then broken down into sub goals (e.g. reduction of downtime) which allow the deduction of operative measures. In order to achieve the sub goals, methods (e.g. analysis of reasons for machine failure) and tools (e.g. failure list) are applied. Methods and tools of similar content are bundled in fields of activities (e.g. total productive maintenance). Altogether, 14 common fields of activities could be identified: among them, visual management, workplace organization, 5S-housekeeping, team work, total quality management, continuous improvement, process standardization, total productive maintenance, leveling and mixed production, just-in-time concepts and kanban [9]. Within these fields of activities the LPS addresses technological and organizational issues, as well as workforce-related aspects.

In addition to these tangible elements, a common vision of the ideal state as well as a philosophy and corporate culture that also reflect the lean ideas are crucial parts of the LPS [5] and [12].

2 IMPLEMENTATION OF A LEAN PRODUCTION SYSTEM IN SME

2.1 Special characteristics of the implementation of a LPS in SMEs

In recent days many large enterprises have developed and successfully implemented a LPS. While large enterprises are able to provide necessary resources like budget, manpower, and time, as well as experts' know-how to configure and implement a LPS, SMEs lack these essential resources it. [13]. It is important to mention that the implementation of an LPS is not just a regular rationalization project, but a fundamental change in the organization and culture of an enterprise [14]. Moreover, many approaches to LPS implementation can not be applied by SMEs for several reasons: specific needs and expectations of SMEs (e.g. lack of essential resources) are not adequately considered, size restrictions and flat hierarchies are neglected, the link to the strategy of the SME is deficient and, referring to one of the most important drivers in SMEs, the entrepreneurs and employees cannot cope with the new challenges without widespread support [15]. In addition, the implementation of a LPS in SME needs the continuous support of a well-structured qualification background.

These specific characteristics explained above affect the size and structure of the implementation teams (e.g. project teams), the time horizon and scope of planning and also the whole sequence of the implementation process of a LPS. For the development and design of a LPS implementation process the following aspects have to be taken into account for SMEs:

- Mainly it is only possible to configure small project teams whose project work partly takes place during off-time.
- The entrepreneur of the SME is in charge of the project management and acts as driver and motivator of the whole process.
- Most SMEs lack experts' know-how but also financial resources to afford external support.
- Missing performance indicators will complicate an analysis of the current state of the organization and the monitoring of the implementation and future benefits of an LPS.
- Pilot projects as performed in large enterprises can hardly be carried out in SMEs. Instead of realizing an overall project, urgent and easy-to-integrate project modules need to be defined and realized.
- The communication of aims and project schedules to the employees as well as the integration of them in the implementation process should occur at an early stage.

2.2 Phases in the implementation process of a lean production system in SMEs

Based on the implementation process of LPS suggested by Dombrowski [14] and taking into account the special characteristics of SMEs an implementation process of a LPS for SMEs has been developed. This process contains seven different phases that are executed consecutively and can also be repeated. Figure 1 visualizes the LPS implementation process in SMEs.

The implementation process typically starts with the awareness (phase 1), when the entrepreneur of the SME learns about success stories of existing LPS. The entrepreneur's frequent contact with entrepreneurs of other SMEs (e.g. in SME networks) provides access to this knowledge. If the entrepreneur decides to pursue the idea of an LPS, the achievable benefits have to be analyzed more in detail. Moreover, the integration of the lean principles into the existing production strategy is necessary and objectives for the LPS have to be formulated [13]. In this regard, the entrepreneur often has to consult external experts. Simultaneously, all employees of the SME need to be informed about these issues at a very early stage. At the end of this assessment and strategic planning (phase 2) the entrepreneur decides whether to commit to the LPS or to abandon this idea.

Next, a central LPS planning and steering team is installed. Generally the team is comprised of the entrepreneur, employees with lead positions (e.g. executive producer) and possibly external experts. This team is responsible for the conceptual design of the LPS and determines the sub goals, fields of activities of the LPS and also the methods and tools to be used. Since many SMEs lack LPS knowledge, the central planning and steering team is, if possible, supported by external experts. At the end of this phase the LPS design is adopted. Once the conceptual design has been agreed on (phase 3), the central planning team also devises a master and detail plan for the implementation and plans necessary organizational changes. The master and detail plan provides milestones, comprises workshops and training courses, specifies the implementation on a local scale, and plans the utilization of resources. These activities are part of the LPS implementation planning (phase 4). In this phase implementation teams are installed. Employees with lead positions and shop-floor employees constitute the implementation teams that account for the implementation of the tangible measures (methods and tools). The decision on the tangible measures marks the end of this phase.

Following these basic planning and set-up activities, which are centralized, the decentralized roll-out starts with a pilot project phase (phase 5). During the pilot phase the implementing teams are testing new methods and tools in selected sectors of the SME. With the experiences gained in these trials the implementation in the whole SME is less risky. In contrast to large enterprises, most SMEs cannot afford to apply trial-and-error procedures by the implementation without endangering their existence. The success of the implementation of the new methods and tools will only be possible once the entrepreneur gets all involved employees on board [16]. The newly implemented methods can only develop their full potential if the employees accept the processes and utilize the implemented methods. Once a method or tool is successfully implemented, the rollout (phase 6) for this element has been completed.

After the transition to the daily operations phase (phase 7) the implemented elements have to be continuously applied and developed in order to ensure continuous improvement. Therefore, during pilot projects, rollout or daily operations, a leap back to the LPS implementation planning phase may occur. Furthermore, if substantial changes in the LPS become necessary changes in the conceptual design might be necessary. This can lead to the repetition of the conceptual design or even the lean assessment phase. Under normal circumstances these iterations also occur, since once in a while it is necessary to review the

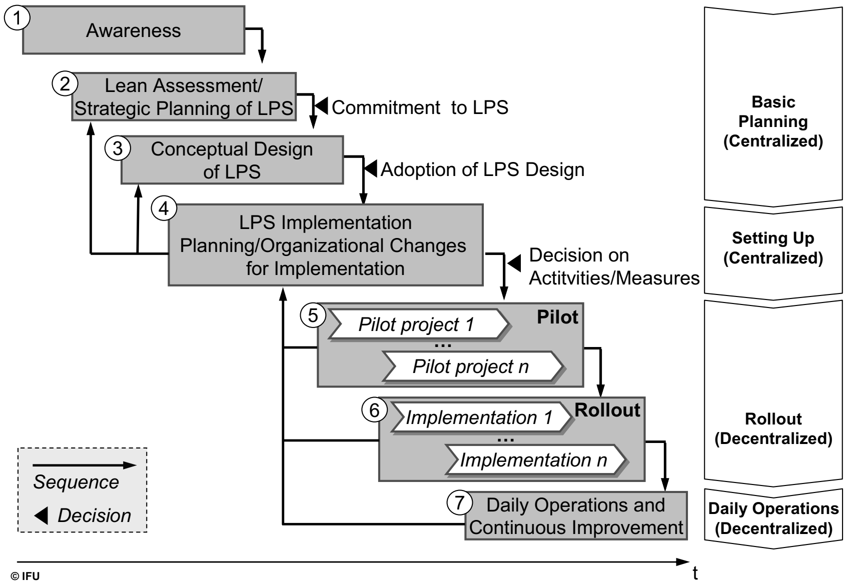


Figure 1: Phases in the implementation process of a Lean Production System in SMEs

implementation process and realign the LPS with the production strategy of the SME.

3 QUALIFICATION FRAMEWORK AS SUPPORTING FUNCTION OF LPS IMPLEMENTATION IN SME

3.1 Need of Qualification by LPS implementation in SME

In the European Union, one quarter of persons employed at SMEs have completed some kind of tertiary education (4% of employees have a postgraduate degree, and 22% possess a university diploma or equivalent). Another 54% have completed a secondary school. The proportion of those employees who did not complete a secondary school is the highest in the manufacturing (30%) industry [17]. Unfortunately, these facts indicate a lack of skilled labour, that constitutes one of the major business constrains in SMEs. Furthermore, SMEs feel reluctant to invest in people very often, as they fear the possibility of skilled labour being “poached” by competitors.

The Toyota Production System (TPS), that represents the paradigm of Lean Production, is based on the philosophy that employees are the greatest asset of an enterprise. Toyota leaders are fond of saying they “build people, not just cars” [5]. Meaning that in the manufacturing process and during the TPS implementation, employees have to learn and develop.

Based on this statement, it can be affirmed, that during the different phases of the LPS implementation process, knowledge is handled and qualification activities are necessary by all means [14]. SME entrepreneurs should be conscious of the advantages derived from employee qualification, such as enhanced staff retention and higher motivation as well as increased competitiveness and productivity [18]. Highly qualified employees in SMEs must be considered as a strategic competitive advantage [19]. Particularly with regard to LPS implementation in SMEs a well-structured qualification background should be developed.

The current situation in SMEs shows that employee qualification is becoming an increasing issue, which needs to be solved utilizing novel approaches. On this account a qualification framework has been developed to support the LPS implementation in SMEs. The following framework is currently being field-tested in close cooperation with six small enterprises as shown in the publicly financed research project «ProfiL» (Production and Organization Flexibility in Life Cycle). This project is funded by the German Federal Ministry of Education and Research (BMBF) within the Framework Concept ”Research for Tomorrow’s Production” and managed by the Project Management Agency Forschungszentrum Karlsruhe, Production and Manufacturing Technologies Division (PTKA-PFT).

3.2 Qualification Framework

The need for qualification by the implementation of a LPS in SMEs depends on two relevant factors. On the one hand it depends on the complexity of the methods and tools to be implemented. For example basic methods and tools, such as work standards or methods of 5S housekeeping can be successfully applied in SMEs without excessive effort [13]. Other methods and tools such as kaizen or total quality management require knowledge and experience about the processes of the SMEs. Implementing these methods requires a specific qualification of the entrepreneur and employees as well as the alignment of these methods with the current production strategy.

On the other hand, the current qualification level of the entrepreneur and employees of the SME determines the need for qualification. This qualification level does not only consist of the knowledge needed for workmanship but also of the specialized knowledge about the lean philosophy and lean methods and tools. The state of specialized knowledge among entrepreneur and employees varies from «unknown method», «known by name or content known» to «successful implementation of the method».

The qualification framework presented in this paper considers the need for specialized knowledge about the lean philosophy and methods as well as the current qualification level of the entrepreneur and employees of the SME. In addition, the framework offers various approaches to impart the required knowledge such as cascade-training, frontal experts-training and method-adoption by the worker. These approaches are currently proved and tested within the research project «ProfiL» and will be explained below.

The qualification framework contains three different modules depicted in figure 2.

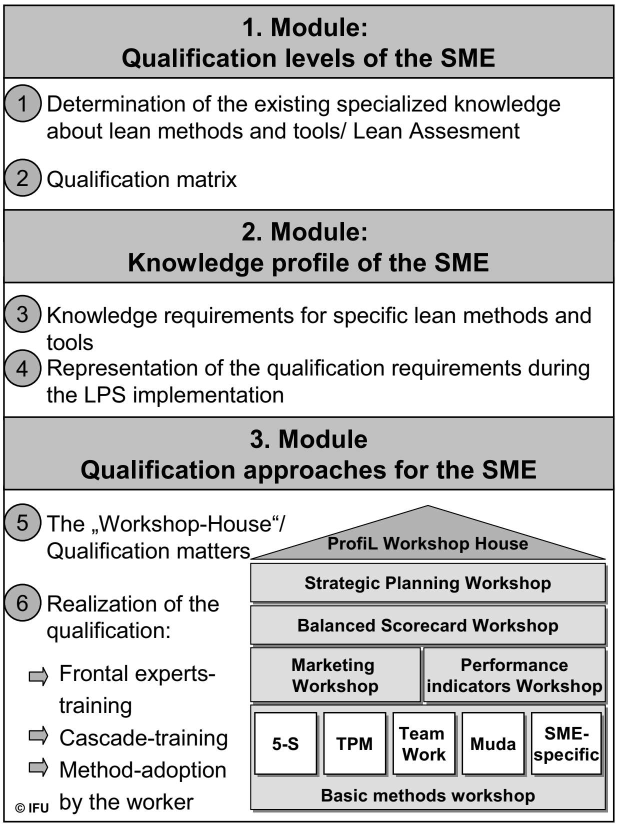


Figure 2: Structure of the Qualification Framework

Module 1: Qualification levels of the SME

Module 1 contains two steps. First of all, the existing specialized knowledge about lean methods and tools currently available in the SME will be examined by using a standardized questionnaire and conducting interviews with the entrepreneur and employees of the SME. Additionally, the state of implementation that has been reached regarding already known methods will be analyzed. The results of the previously conducted Lean Assessment (phase 2) will be incorporated as well.

A successful implementation of LPS requires the transfer of basic knowledge about methods to employees as well as specialized knowledge to the entrepreneur. Furthermore, it is necessary to keep records of recent skills and the current qualification level of all employees.

A qualification matrix can be utilized as an effective tool to evaluate and track the progress and qualification level of each employee. In this matrix each employee’s capabilities are represented and placed. Poor qualification levels will be detected and the entrepreneur will have a vested interest in having exceptionally qualified employees. Also, employees get the sense that they do not matter as individuals if the leader does not place importance on the training effort. On this account a qualification matrix has been also developed and implemented successfully by several SMEs of the research project «ProfiL». The matrix depicts all three components of knowledge required for the execution of day to day business. All three components represent knowledge about activities, tasks, and processes [20].

The result of the first phase of this qualification concept is an effective and well-founded description of the qualification level of the entire enterprise. The aforementioned qualification level is a requirement for the next phase in which a knowledge profile of the SME will be developed.

Module 2: Knowledge profile of the SME

On the one hand the implementation of LPS requires certain methodological skills, on the other hand the corporate strategy and organizational structure needs to be taken into consideration.

In this regard the creation of a knowledge profile seems appropriate. This framework helps to evaluate which knowledge is necessary for the implementation of certain methods. In other words which knowledge has to be available within the enterprise respectively needs to be “procured” if not. The need for specific knowledge regarding different methods constitutes the knowledge requirements for a LPS implementation and therefore gives a hint on the knowledge needs to make available to the enterprise, the entrepreneur and employees respectively. The knowledge profile contains the previously developed qualification level, the knowledge requirements for specific methods, as well as a strategic commitment which links the qualification concept and the strategic planning of the SME.

In case of the research project “ProfiL” the knowledge profile had been developed with the support of external experts. In general SMEs need to be assisted with the creation of its knowledge profile.

Module 3: Qualification approaches for the SME

A «workshop-house», in the style of the house illustration of the Toyota Production System that contains a catalogue of workshops of lean production methods has been developed within the research project «ProfiL». This «workshop-house» should support the transfer of basic knowledge about methods to employees, specialized knowledge to the entrepreneur during the basic planning phase as well as essential knowledge for the operative implementation of these methods to the implementing teams during the roll-out phase. These workshops also contribute to increase the motivation of the employees and entrepreneur.

The fundament of this «workshop-house» is represented by basic methods and tools that are based upon Toyota’s philosophy: «identify and eliminate waste». These basic methods such as teamwork, work standards, methods of 5S housekeeping, and methods of TPM, require minor amount of time and specialized knowledge, support the processes of the enterprise and can be successfully applied in large enterprises but also in SMEs [13].

The second element of the third module presents various approaches (e.g. frontal experts-training, cascade-training, and method-adoption by the worker) that allow the realization of the qualification. These approaches represent various ways to impart the required knowledge.

Experts in the field of Lean Production impart knowledge about LPS implementation, methods and tools in several workshops directly to the employees and entrepreneur. This approach represents the so-called frontal experts-training. The content of the workshops should be adapted to the knowledge profile of the SMEs. The different needs, demands and qualification levels of the entrepreneur and employees have to be considered here. This approached is depicted in figure 3.

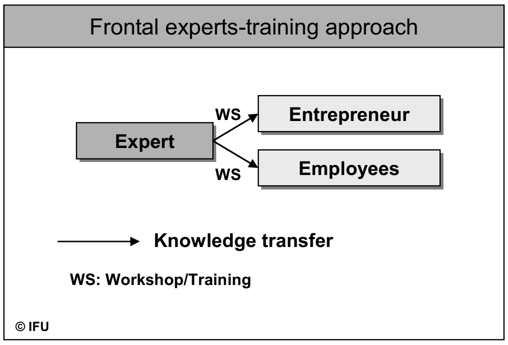


Figure 3: Description of the frontal experts-training approach

This approach can be used at the beginning of the LPS implementation process during the basic planning and setting up phases to start the process right. Since the costs involved are high, SMEs are not able to afford the use of this approach for the whole implementation process [6]. Therefore it would be possible to combine this approach with the cascade-training approach as a possibility to solve this problem. This mixed approach has been considered and successfully used in the research project «ProfiL».

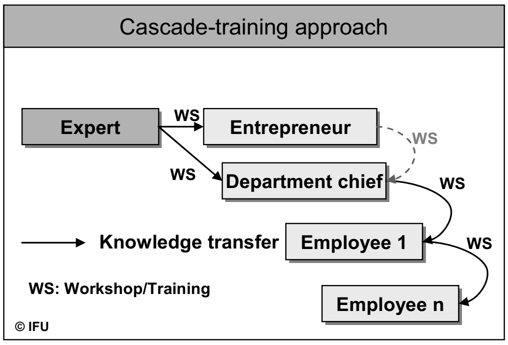


Figure 4: Description of the cascade-training approach

The cascade-training (cascade means in this context that the information transfer and the impartment of knowledge happens in a top down way) consists firstly of a frontal experts-training for the entrepreneur and employees with lead positions (e.g the central LPS planning and steering team by the LPS implementation). Secondly, the entrepreneur and these employees are responsible for the transfer of this acquired knowledge to other employees later on. This transfer will take place within workshops. Accordingly all employees that are involved in the LPS implementation will be qualified. In the figurative sense, each learner can be theoretically considered as a teacher [6].

In addition, the experience gained by using the cascade-training approach in the research project “ProfiL” is extremely positive. This approach depicted in figure 4 should be particularly used during the rollout phase of the LPS implementation. The entrepreneurs, employees with lead positions or selected employees which assume the roll of the teacher have to comprehend and study not only the lean philosophy but also the contents of the methods and tools to be applied. Otherwise the cascade-training approach offers entrepreneurs the possibility to transmit their belief in the LPS implementation to the employees. They also have to provide support and coaching to the workshop members. Furthermore the entrepreneur does not only have to inspire the employees to achieve the LPS goals but also to motivate them to support the implementation [5].

The willingness and desire of the employees to support the LPS implementation can also be increased by using the «method-adoption by a worker» approach. By implementing this approach entrepreneurs or employees will be qualified as experts in specific subject areas (in this case methods and tools of lean production) [20]. They will be named as “godfather” of the method or tool and are also primarily responsible for the implementation of the method or tool in the SME In addition; they are available for any further questions concerning the method and its implementation. They continuously track the level of achievement of the method and, know the correct way each method should be performed. With this ability, the “godfather” of the method can ensure that the LPS implementation is being performed correctly to plans. Furthermore, they should constantly analize the application of the methods during the rollout and daily operations phase, looking for ways to improve and make better use of materials, machines and manpower encouraging the employees to develop continuous improvement in thinking and action.

The «method-adoption by a worker» is not only a qualification approach but also an approach for a successful implementation as wells as further and sustainable development of the lean production methods and tools by the employees. Thus the employees serve here as “multipliers” of strategic knowledge and learning.

4 SUMMARY

For the implementation of a LPS SMEs need continuous support by a well-structured qualification framework. When applying the framework to SMEs following aspects have to be considered:

- The qualification framework must be taken into account within the strategy and organization of the SMEs.
- All processes for the analysis of the qualification level of the SME and the creation of its knowledge profile have to be established during the LPS implementation process. Furthermore, it is essential to run, improve, and monitor them continously.
- The Management of the SME has to demand the skills and willigness of all employees. Further on the content and the mode concerning the impartment of knowledge

needs to be adapted to the specific needs and organizational structure of the SME.

- A systematic and methodical qualification concept holds decisive competitive advantages for SMEs.

5 REFERENCES

[1] Herrmann, C., Bergmann, L., Crespo, I., 2006, Developing of an Integrated Modernization Concept for SMEs, Proceedings of the 39th CIRP Internantional Seminar on Manufacturing Systems, 457-463.

[2] Schneider, H., 2000, Produktionsmanagement in kleinen und mittleren Unternehmen, Schäffer-Poeschel, Stuttgart.

[3] Zahn, E., Huber-Hoffmann, M., 1995, Die Produktion als Wettbewerbskraft, Produktion als Wettbewerbsfaktor, Beiträge zur Wettbewerbs- und Produktionsstrategie, 134-155, Gabler.

[4] Ohno, T., 1998, Toyota Production System – Beyond large scale production, Portland, OR, Productivity Press

[5] Liker, J.K., Meier, D., 2006, The Toyota Way Fieldbook: A Practical Guide for Implementing, Toyota´s 4Ps, McGraw-Hill, New York.

[6] Spath, D., 2003, Ganzheitlich produzieren: Innovative Organisation und Führung, Log_X, Stuttgart.

[7] Womack, J.P., Jones, D.T., 2003, Lean Thinking: Banish Waste and Create Wealth in Your Corporation, Simon & Schuster, New York.

[8] Womack, J.P., Jones, D.T., 2005, Lean solutions: How Companies and Customers Can Create Value and Wealth Together, Simon & Schuster, London.

[9] Dombrowski, U., Palluck, M., Schmidt, S., 2006, Strukturelle Analyse Ganzheitlicher Produktionssysteme, ZWF Zeitschrift für wirtschaftlichen Fabrikbetrieb, 101:113-118.

[10] Cochran, D.S., 1999, The Production System Design and Deployment Framework, SAE.

[11] Korge, A., 2005, Lean Management mit System, wt Werkstattstechnik online, 95:29-34.

[12] Spear, S., Bowen, H.K., 1999, Decoding the DNA of the Toyota Production System, Harvard Business Review, 97-106.

[13] Dombrowski, U., Crespo, I., Schmidt, S., 2007, A Holistic Approach to Production Optimization as an Integral Part of Strategy Development and Implementation in Small-and Medium-sized Enterprises, 2007, Proceedings of the 40th CIRP International Manufacturing Systems Seminar, Liverpool.

[14] Dombrowski, U., Schmidt, S., Crespo, I., 2007, Knowledge Management as a Supporting Function Lean Production System Implementation, Proceedings of the 2nd International Conference on Changeable, Agile, Reconfigurable and Virtual Production, 453-462, Toronto.

[15] Behringer, S., 2004, Unternehmensbewertung der Mittel- und Kleinbetriebe, Erich Schmidt Verlag, Berlin.

[16] Dombrowski, U., Quack, S., 2007, Erfolgreiche Restrukturierungsprojekte durch Mitarbeiterpartizipation, ZWF Zeitschrift für wirtschaftlichen Fabrikbetrieb, 102:568-571.

[17] Enterprise Observatory of European SMEs, 2007, Enterprise Observatory Survey, Flash EB Series 196, Enterprise Publications.

[18] Observatory of European SMEs, 2003, Report No.8, Highlights from the 2003 Observatory, Enterprise publications.

[19] Dombrowski, U., Tiedemann, H., Bothe, T., 2003, Mitarbeiterbildung als strategischer Wettbewerbsfaktor, Further education and training employees as a strategic competitive factor, BFT Betonwerk und Fertigteil-Technik, 4:6-14.

[20] Dombrowski, U., Horatzek, S., Zeisig, M., 2002, Situatives mitarbeiterbezogenes Management der Ressource Wissen, wt Werkstatt Technik online, 92/6:312-316.

An environmental perspective on Lean Production

Christoph Herrmann[1], Sebastian Thiede[1], Julian Stehr[1] and Lars Bergmann[1]
[1]Institute of Machine Tools and Production Technology (IWF), Technical University Braunschweig, Germany

Abstract
To cope with the challenges of the business environment companies strive to reorganize their production with certain principles and methods in the sense of Lean Production. Whereas these measures are naturally primarily motivated by classical production objectives like cost, time, quality or flexibility, companies also face diverse originally environmentally driven challenges which also incorporate a strong economic relevance. However, methods of Lean Production are not necessarily environmentally friendly. Thus, there may exist certain conflicts in goals for companies. Against this background, this paper presents an analysis of the coherences and interdependencies between Lean Production and ecologically oriented variables specifically energy consumption. Besides the discussion of principles and methods regarding their environmental influence, a simulation approach is used to analyze the specific effects of certain Lean Production measures on economic as well as environmental variables of a manufacturing line.

Keywords:
Sustainable Manufacturing; Manufacturing System; Lean Production

1 INTRODUCTION

Today's business environment for producing companies is characterized by shorter product life cycles, rising product variant diversity, increasing production volume fluctuations and rapid changing technologies. These changing external conditions as well as the global business environment are challenging production enterprises to continuously adapt and improve their production systems. In this context, production systems represent the sum of all processes, resources, principles, and methods within a production enterprise. Having in mind the success of the Toyota Production System, Lean Production Systems are widely seen as promising approach to cope with these challenges in manufacturing companies.

Besides classical economical production objectives (e.g. cost, time, quality), environmental driven objectives (e.g. low CO_2 emissions), become increasingly relevant for producing companies. Besides the idealistic vision and strive for sustainable manufacturing, increasing raw material prices, necessary investments for environmental technologies, potential penalties for lacking compliance with environmental regulations as well as certain regulative incentives, the introduction of CO_2 certificates, the rising public awareness on resource consumption and climate change potentially resulting in challenging consequences on the corporate image or especially rising energy costs are just some examples that underline the increasing important economic relevance for companies of originally environmentally driven issues.

Therefore the evaluation criteria of all technological and organizational measures within producing companies are significantly extended nowadays striving for an integrated solution to foster sustainability in manufacturing or in other words to harmonize the requirements of a sustainable development with the needs of manufacturing. Whereas Lean Production is typically aiming at optimizing classical economic objectives like costs, time, quality and flexibility the questions comes up whether there are coherences and eventually conflicts of goals between lean principles or methods and the environmental driven objectives such as low emissions.

2 LEAN PRODUCTION

2.1 The Toyota Production System

To cope with the market pressure under constrained conditions in post-WWII period, the founder of Toyota, Sakichi Toyoda and the engineer Taiichi Ohno developed the so-called Toyota Production System (TPS) [1]. The TPS can be regarded as a general framework and philosophy to organize the manufacturing facilities and processes at Toyota and the interaction of these facilities and processes with the suppliers and customers to provide best quality, lowest cost, and shortest lead time through the elimination of the seven forms of waste under involvement of all employees [2]. Usually, the TPS is illustrated by a house like shown in Figure 1 and comprises of two pillars, Just-in-Time and Jidoka [1] [2].

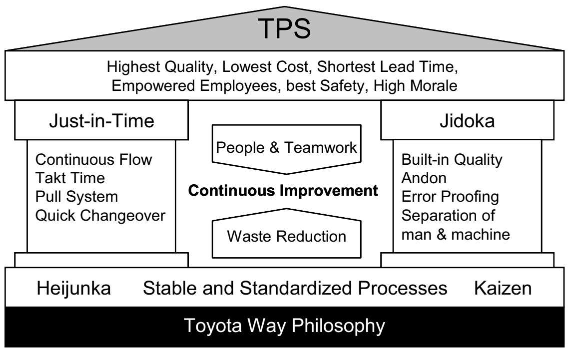


Figure 1: TPS house, adapted from [1] [2]

TPS is maintained and continuously improved through iterations of standardized work and kaizen, following the PDCA (plan, do, check, act) cycle. Toyota was capable of significantly reducing cost and inventory using the TPS, enabling it to become one of the ten largest companies in the world. As the main goal of the TPS is to eliminate waste, TPS is generally known as lean manufacturing.

2.2 Lean Production in Europe

Due to the success of the TPS and the need to continuously improve the own business the TPS framework and several of its inherent methods have been absorbed by many European production enterprises in order to improve productivity and flexibility. Most of them did so by copying the main concept of the TPS and renaming it to take ownership of it as their own system (e.g. Bosch Production System (BPS), Autoliv Production System (APS), Mercedes Benz Production System (MPS), et cetera). A majority of these companies use individual adapted house-like figures to illustrate and communicate their production system structure [3] [4] [5] [6]. As these TPS adaptations show significant accordance, they can be summarized as Lean Production Systems (LPS).

Generic Structure of Lean Production Systems

Sub-systems	Just in Time	Kaizen	Quality	Work Structures	…
Principles	Levelling Pull System One Piece Flow …	Standardization Muda Elimination Suggestion System …	Six Sigma TQM Stable processes …	Teamwork Ergonomics Safety …	…
Methods / Tools	FiFo SMED Kanban …	Kaizen Workshop PDCA Standards 5S …	DMAIC FMEA SPC 5 Why …	Job rotation Workload Analysis Safety Analysis Interview / Survey …	…

Figure 2: Generic structure of Lean Production Systems [7]

Figure 2 shows the principle structure of a LPS illustration. Although some LPS contain subsystems and elements that are not contained in the TPS in exactly the same position or manner, these subsystems and elements can easily be mapped to correlating subsystems and elements of the TPS.

2.3 Lean Production System Design

As mentioned before, production systems consist of all required functions, processes, activities, resources, principles, and methods within an enterprise to produce marketable goods, services, or a combination of both. To achieve strategic objectives (SO) as defined by the strategic management, specific production system capabilities are necessary prerequisites [8]. From the operations management perspective, production system capabilities are the result of the fulfillment of certain functional requirements (FRs) in the production system. To fulfill functional requirements, specific design parameters (DPs) are applied according to the production strategy as shown in Figure 3 oriented on the viable system model (VSM).

Changes in the environment are represented by change drivers (CD), as they increase the pressure on companies to continuously adapt production system capabilities. Moreover, strategic objectives and changing product requirements require adapting production system capabilities. A continuously adaptation of production systems to changing conditions is a challenge for the operations management, as appropriate DPs have to be identified and applied that fulfill multi-dimensional requirements.

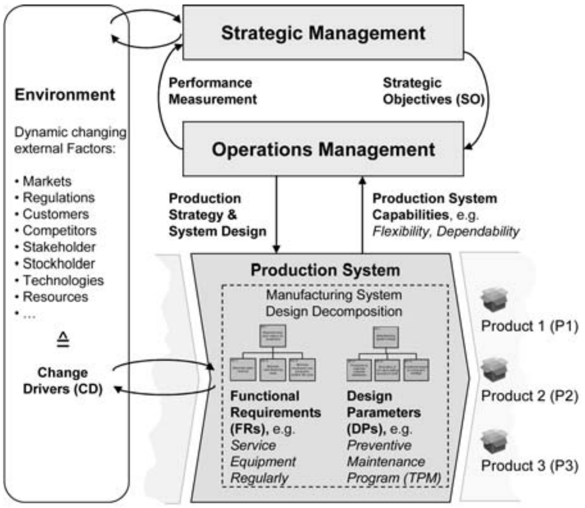


Figure 3: Connection between strategic objectives, functional requirements, design parameters, production system capabilities and change drivers [12]

Approaches to support production system design specifically considering waste avoidance in the sense of lean manufacturing have been extensively studied and a diversity of FRs and DPs for improving quality, time and costs have been discussed [8] [9] [10]. To describe the interrelations of FRs, DPs, and production system capabilities, the manufacturing system design decomposition approach (MSDD) can be used [8] [10]. Based on the axiomatic design methodology [11] and focusing on optimizing return on investment (ROI) as primary objective of a company, the MSDD hierarchically defines objectives for production system design in terms of FRs. To satisfy FRs, appropriate DPs are assigned to the FRs (Figure 3 and 4). As the fulfillment of specific FRs has impact on the achievement of specific capabilities, knowledge about the relation between the level of fulfillment of specific FRs and the expectable contribution to specific production capabilities can be used for lean production system design [12] [13].

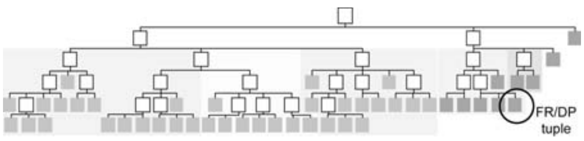


Figure 4: Coherences between strategic objectives and functional requirements in MSDD (e.g. costs) [8] [9] [10]

3 BEYOND LEAN MANUFACTURING

3.1 Sustainable Manufacturing as new paradigm in production

Environmental challenges like global warming, increasing pollution or inadequate resource consumption are getting accelerated due to ongoing global industrialization and will have unforeseeable consequences for current and upcoming generations. Besides these idealistic perspective (long-term) these developments also evoke strong economic effects (short-term) on producing companies e.g. through rising energy or material costs. Additionally social trends like the aging society or the consequences of ongoing globalisation on population also have major impact.

Against this background it is very obvious that an isolated consideration of classical economic variables is not sufficient nowadays. In fact, Sustainable Manufacturing is the new necessary paradigm for producing companies which involves the integration of the economic, environmental and social perspective (known as the triple bottom line) for all technological and organisational measures within the normative, strategic and operative production management (Figure 5). As a holistic approach which strives to avoid problem shifting within producing companies, their supply chain and life cycle phases this involves the consideration of all three basic strategies of sustainability (efficiency, sufficiency, consistency) on different layers beginning from the single (production) process, process chains on a factory layer, strategic decisions on a company layer or activities in closed looped supply chains like utilizing Re-X-options, such as remanufacturing or refurbishment (network layer) [14].

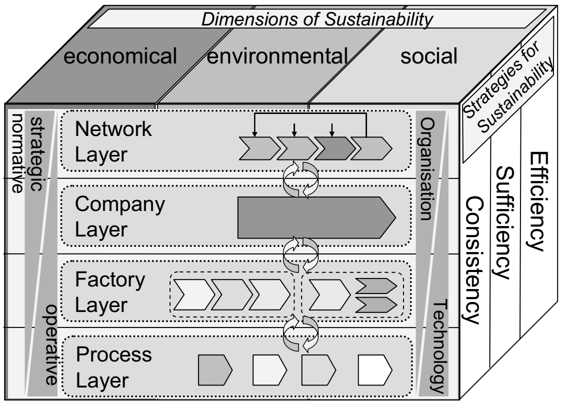


Figure 5: Framework for Sustainable Manufacturing

Resulting from the thoughts above, a holistic evaluation of production companies naturally also involve the simultaneous consideration all relevant dimensions within integrated process models. Thus, traditional economical variables are necessarily getting combined with typical ecological metrics (e.g. scrap, hazardous materials/ materials use, energy use, water use, air emissions, solid and hazardous waste, water pollution/ wastewater) for example [15] [16] [17].

3.2 Lean and Green

Traditionally, lean and green manufacturing were seen as "distinct set of solutions targeting different forms of wastes" [18]. Referring to the framework for sustainable manufacturing (Figure 5) on a company layer traditionally isolated initiatives for lean (e.g. lean production systems) and green (e.g. EMS - environmental management systems like ISO 14000) management were started separately aiming at economic (lean) and environmental objectives (green). On a factory or process layer lean management involves the implementation of principles and methods as described above whereas environmental driven measures consider end-of-pipe technologies or process integrated solutions to reduce environmental impact. However, studies underline the strong coherences and interdependencies between lean and green manufacturing activities [18]. Both perspectives share similar basics with eliminating waste as a major aspect. As shown in Table 1, the "deadly wastes" as defined in lean manufacturing also incorporate strong environmental impacts. From that point of view an implementation of lean management is likely to have also positive effect on a company's environmental performance. Therefore studies and even policy makers (e.g. U.S. Environmental Protection Agency) propose the implementation of lean principles, methods and tools (e.g. value stream mapping) as promising way to improve economic and environmental results [17] [19].

Table 1: Environmental impact of waste [17]

Waste Type	Environmental Impacts
Overproduction	• More raw materials and energy consumed in making the unnecessary products • Extra products may spoil or become obsolete requiring disposal • Extra hazardous materials used result in extra emissions, waste disposals, worker exposue, etc.
Inventory	• More.packaging to store work-in-process (WIP) • Waste from deterioration or damage to stored WIP • More materials needed to replace damaged WIP • More energy used to heat, cool and light inventory space
Transportation and Motion	• More energy use for transport • Emissions from transport • More space required for WIP movement, increasing lighting, heating and cooling demand and energy consumption • More packaging required to protect components during movement • Damage and spills during transport • Transport of hazardous materials requires special shippimg and packaging to prevent risk during accidents
Defects	• Raw materials and energy consumed in making the defective products • Defective components require recycling or disposal • More space required for rework and repair, increasing energy use for heating, cooling and lighting
Over processing	• More parts and raw materials consumed per unit of production • Unnecessary processing increases wastes, energy use, and emissions
Waiting	• Potential material spoilage or component damage causing waste • Wasted energy from heating, cooling and lighting during production downtime

These complementary relations between lean and green activities on a factory respectively process layer lead to a stronger convergence of both perspectives nowadays. The positive effects were also confirmed by studies of North American manufacturing companies regarding their status of lean and green management implementation (implemented management system and waste reduction technologies as well as results). The results even reveal that the existence of green manufacturing initiatives actually foster the implementation of lean management in companies and specifically supports to improve cost performance [18].

However, studies also show that there is a critical lack of integration of lean and green activities within the strategic production system design on the company layer [18]. As previous descriptions underline mainly the double-sided effect of eliminating wastes with rather operational activities is considered. This is just one part of the problem and does not necessarily provide the whole picture – exemplarily it lacks relevant strategies which do not directly depend on quantitative waste reduction. From this perspective, lean production is rather a necessary but not sufficient prerequisite for sustainable production system design. Therefore a holistic perspective including all management alternatives and interdependencies and the integration of lean and green activities within one management perspective is necessary to foster the pursuit of zero emission manufacturing system as next evolutionary step simultaneously respecting economic and environmental objectives [18].

Additionally previous approaches and studies acting on a rather rough and pragmatic level and do not consider the multi-dimensional effects and interdependencies of particular principles and methods. Whereas the relations between these activities are manifold and conflicts of goals may show up, a deeper analysis on this layer is also necessary.

Production System Design

As described above, the MSDD is a common and established approach to support lean production system design on a company layer. Therefore it can serve as starting point for an integrated production system design under consideration of environmental driven issues. A deeper analysis of all FR and DP within the MSDD regarding their contribution to three sustainability dimensions reveals the clear economic orientation of classical lean production system design approaches [12]. While lean production systems per definition include the involvement and commitment of employees, the social dimension is also fairly addressed. In contrast to that environmental objectives came up to be not really relevant for production system design [12]. Additionally, classifying the FR / DP regarding sustainability strategies shows that lean production naturally mainly focuses on increasing efficiency and partly sufficiency (avoiding waste, using / producing only what is actually necessary, quantitatively improving material flows). Consistency as important strategy to change material flow qualitatively through substitution or circuitry is not included in this approach. Thus, to use the MSDD as base for sustainable production system design, an extension towards the missing dimensions and strategies is essential [12].

Principles and Methods

Principles, methods and tools are used actually operationally to implement lean production strategies on the factory respectively process layer. Eliminating waste is the major objective of these methods - as Table 1 underlines, this can naturally provoke positive effects in both economic as well as environmental variables. Focusing on an isolated, output oriented evaluation this is certainly true for all methods. However, the utilization of certain methods to simultaneously optimize economic and environmental performance requires the consideration of two essential aspects:

- The interdependencies and conflicts of goals between methods and their uncertain effects on all target dimensions whereas combined implementation strategies may lead to different results as expected (e.g. production downtime of uninvolved machines in case of andon line stop).
- Implementing and maintaining methods often involves certain effort in economic (e.g. labour, time, investment) or environmental (e.g. energy) terms. Thus, it is essentially important to make sure that this effort actually pays off over time in all dimensions. For example, for lean production methods like Just-in-Time / KANBAN simulation based studies suggest that the additional environmental impact from increasing transportation does not necessarily pay off. Thus, implementing this method is not advantageous compared to other strategies in considered cases with given restrictions from an ecological standpoint [20].

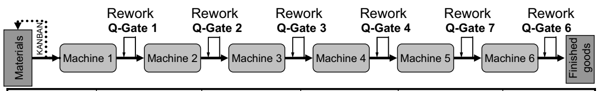
4 SIMULATION BASED ANALYSIS

4.1 Simulation Structure and Methodology

A simulation model shall help to gain deeper insight regarding the interdependencies of lean principles and methods and specifically their environmental impact. As a case study a fictive and transparent reference production system was chosen (Figure 6). It consists of six serially linked production machines and needs material input at the first station. Within this production system diverse principles and methods oriented on lean management can be implemented. Based on the previous discussions and their relevance for lean production systems five methods were analysed:

- Pull-System for material input: KANBAN instead of supply for storage to lower material inventory
- Standards/Qualification: increasing standardisation or qualification of employees incorporate similar effects and are likely to increase production efficiency
- Total Quality Management / Zero Defects Strategy: when activated, quality gates behind each station early recognize product quality problems and redirect to rework areas. The standard case is just one quality gate at the very end of the production line.
- Andon Line / Jidoka: in case of quality problems the employee uses the andon line to stop the whole production line until the actual reason is found – thus, avoiding repeating failures or defects.
- Continuous Flow: in contrast to lot size oriented transportation between machines, a continuous flow, idealistically one piece flow, can be integrated through transportation system like conveyors.

As shown in Figure 6 different combination of these methods lead to different scenarios that were considered whereas all other production parameters (e.g. cycle time, availability, production amount) are naturally hold constant to enable comparisons.



Scenario	Pull-System Material Input	Standards Qualification	Quality Strategy	Andon Line (JIDOKA)	Continuous Flow
Scenario 1	⊘	⊘	⊘	⊘	⊘
Scenario 2	⊘	✓	⊘	⊘	⊘
Scenario 3	⊘	⊘	⊘	✓	⊘
Scenario 4	⊘	⊘	✓	⊘	⊘
Scenario 5	⊘	⊘	✓	✓	⊘
Scenario 6	✓	⊘	⊘	⊘	⊘
Scenario 7	⊘	⊘	⊘	⊘	✓
Scenario 8	✓	⊘	✓	✓	✓

✓ implemented ⊘ not considered

Figure 6: Reference Production Line and Simulation Scenarios

Due to the high relevance and good comparability as well as applicability energy consumption is used here as metric to measure environmental performance. In this model, energy gets consumed by all production machines, transportation equipment for material supply and internal logistics (e.g. conveyor, fork lifter), quality gates and rework stations. Therefore for all this factors energy consumption rates per hour for different states (e.g. producing, idle) were assigned on a scale from 1 to 10. Whereas effects through transportation were a major focus of this study the energy consumption rate of the equipment for internal transportation were varied (1-9) to conduct a simple sensitivity analysis, thereby naturally holding all other parameters constant. While the standard value for the energy consumption rate of the production machines was 7 (producing) and 3 (idle) a value of 1 for the transportation equipment literally would mean that it consumes 7 times less energy in one hour of full usage.

4.2 Simulation Results

Figure 7 shows the results of the simulation runs with all scenarios and the parameter variation regarding the energy consumption rate of transportation equipment. Metrics are environmentally the total energy consumption for whole production line and from an economic respectively lean perspective the throughput time and work in progress (WIP).

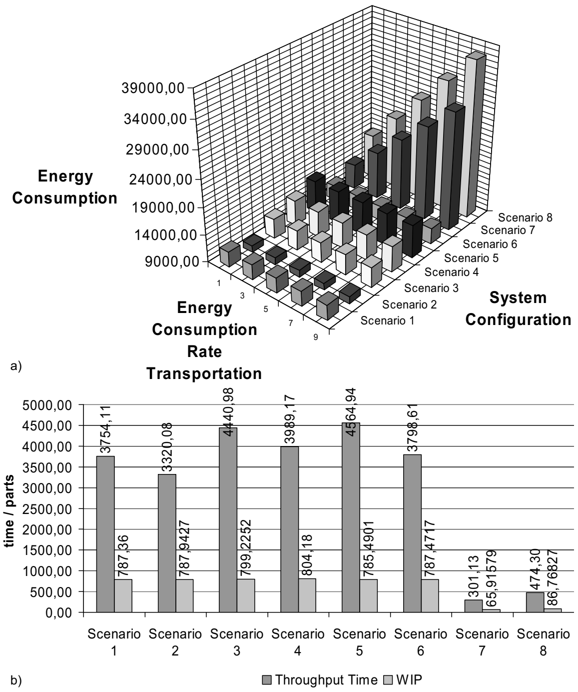


Figure 7: Simulation Results – Energy Consumption (a) and Throughput Time / Work in Progress (b)

The results show some interesting findings. Focusing on the effect of lean methods on energy consumption just the standardization respectively qualification of employees actually lowers the energy consumption. This is quite logic because without relevant energy input (e.g. some indirect but low effort through workshops or training) production efficiency is directly positively influenced leading to less time and energy needed to produce the fixed production program. All other methods lead to at least slightly higher total energy consumption.

- Quality Strategy: More quality gates automatically consume more energy (e.g. automated quality checks). On the positive side, energy muda as energy wasted on processing defect parts is totally prevented.
- Andon Line: The total stoppage of the production line in case of defects lead to a strong increase of production and throughput time and therefore energy consumption. This also provokes high energy consumption through idle states.
- Pull System for Material Supply: the KANBAN-System involves more supply runs to provide the materials. While reducing inventory this lead to a slight increase of energy consumption naturally strongly depending of the actual order sizes.

The consideration of the interlinking between the production machines certainly reveals the most interesting results. A non-continuous flow with lot-wise transportation (e.g. realized by fork lift) is not critical regarding energy consumption under these circumstances even with higher energy consumption rates. In contrast to that the usage of a conveyor to realize a continuous flow enfolds a conflict of goals while it significantly increases the energy consumption but also dramatically decreases the average throughput time and inventory in production respectively work in progress. The continuous running conveyor that ensures the continuous transportation of parts steadily consumes energy which scales up strongly even with quite low consumption rates whereas non-continuous solutions are energetically advantageous in this case. Naturally besides the used technology (e.g. electric or diesel fork lift, truck) this very much depends on factors like the effort of the transportation process (e.g. distance, transportability) and lot sizes respectively amount of runs.

5 CONCLUSION

This paper provides an environmental perspective on Lean Production whereas thereby both the strategic as well as the rather operational perspective was considered. On the strategic side a further integration of management systems is necessary to support sustainable lean production system design under exploitation of all sustainability strategies.

The discussion regarding the operational implementation through principles and methods shows that lean production activities are not necessarily and per definition optimal from an environmental perspective while interdependencies and effects of different activities and dimensions have to be considered. To gain deeper insight and sensitize for these problems a simulation study for a specific case was conducted. The implementation of a continuous flow was found to be critical regarding energy consumption and therefore conflicts of goals with classical economic objectives show up. Again it is important to mention that the simulation approach is for a specific case with a defined production line and distinctive parameters. However it shows first interesting results and can serve as starting point for further research.

6 REFERENCES

[1] Ohno, T., 1998, Toyota Production System - Beyond large Scale Production, Portland, OR, Productivity Press.

[2] Liker, J. K., 2004, The Toyota Way, New York, NY, McGraw-Hill.

[3] Ohno, T., 1993, Das Toyota-Produktionssystem, Campus Verlag, Frankfurt/Main, New York.

[4] Stühmeier, W., Stauch, V., 2002, Mercedes-Benz-Produktionssystem - Implementierung und Controlling in der Produktion der A-Klasse-Motoren, in: Barthel, J. (ed.), 2002: Ganzheitliche Produktionssysteme (GPS) in der Automobilindustrie, Dokumentation einer Workshopreihe, Institut für angewandte Arbeitswissenschaft e.V., Köln, Wirtschaftsverlag Bachem.

[5] Feggeler, A., Neuhaus, R., 2002, Was ist neu an Ganzheitlichen Produktionssystemen?, in: IfaA (ed.): Ganzheitliche Produktionssysteme, Gestaltungsprinzipien und deren Verknüpfung, Köln, Bachem.

[6] Spath, D., 2003, Ganzheitlich Produzieren - Innovative Organisation und Führung, Stuttgart, Log_X.

[7] Herrmann, C.; Bergmann, L.; König, C., 2006: "An inter-organizational Network Approach for Lean Production System Improvement", in: 39th International CIRP Seminar on Manufacturing Systems, June 7-9, 2006 Ljubljana, Slovenia, 2006, pp.113-121.

[8] Duda, J. W., 2000, A Decomposition-Based Approach to Linking Strategy, Performance Measurement, and Manu-facturing System Design, Ph.D. Thesis, MIT.

[9] Womack, J. P., Jones, D. T., Roos, D., 1990, The machine that changed the world, Rawson Associates.

[10] Cochran D. S., Arinez, J. F., Duda, J.W., Linck, J., 2001, A Decomposition Approach for Manufacturing System Design. In: CIRP - Journal of Manufacturing Systems, 20(6), , pp. 371-389.

[11] Suh, N., 1997, Design of Systems, Annals of the CIRP, Vol. 46, No. 1, pp.75-80.

[12] Herrmann, C.; Bergmann, L.; Thiede, S., 2008, Methodology for Sustainable Production System Design. In: 15th CIRP International Conference on Life Cycle Engineering, Sydney, 17-18 March 2008.

[13] Herrmann, C.; Bergmann, L.; Thiede, S., 2007, Life Cycle Oriented Design of Lean Production Systems. In: Proceeding of the 3rd VIDA Conference, Poznan, June 2007.

[14] Herrmann, C., Bergmann, L., Thiede, S., Halubek, P., 2007, Total Life Cycle Management - An Integrated Approach Towards Sustainability, 3rd International Conference on Life Cycle Management, Zurich.

[15] Schultz, A., 2002, Methode zur integrierten ökologischen und ökonomischen Bewertung von Produktionsprozessen und -technologien, Dissertation, Magdeburg.

[16] Herrmann, C.; Bergmann, L.; Thiede, S.; Zein, A., 2007, Framework for Integrated Analysis of Production Systems. In: Takata, S.; Umeda, Y. (Hrsg.): Advances in Life Cycle Engineering for Sustainable Manufacturing Businesses - Proceedings of the 14th CIRP Conference on Life Cycle Engineering.

[17] EPA, 2005, "The Lean and Environment Toolkit, Version 1.0," The U.S. Environmental Protection Agency, Washington, DC.

[18] Bergmiller, G. G., 2006, Lean Manufacturers Transcendence to Green Manufacturing, USF, USA,.

[19] Simons, D. and Mason, R., 2003, Lean and Green: Doing More with Less, ECR Journal, 3(1):84-91.

[20] Venkat, K. ; Wakeland, W., 2006, Is Lean Necessarily Green?, In : Proceedings of the 50th Annual Meeting of the International Society for the Systems Sciences, Sonoma.

Quantified Interdependencies between Lean Methods and Production Figures in the Small Series Production

Jürgen Fleischer[1], Gisela Lanza[1] and Kathrin Peter[1]

[1] Institute of Production Science (wbk), Universität Karlsruhe (TH), Germany

Abstract

Small and medium-sized companies encounter enormous difficulties when trying to implement so-called *lean methods* according to the role model of the Toyota Production System. This is caused by the varying effects of lean methods on production figures depending on the production conditions concerning product variety and volumes, variation of process and set-up times etc. Thus, there are no general guidelines for the "best possible" implementation of these methods in a small series production. This article shows an approach to analyze and evaluate the influence of lean methods in small series productions based on quantified interdependencies with the relevant production figures.

Keywords

Lean methods; production figures; small series production

1 INTRODUCTION

Since the 90s, many car manufacturers, their suppliers and major enterprises from other branches of industry have followed Toyota's example by introducing lean methods to their production system and have considerably improved its performance [1]. Common lean methods are for example those with an immediate effect on the production process, i.e. pull production, just-in-time principle, flow production, production leveling, reduction of set-up times (SMED), Total Productive Maintenance (TPM) or Total Quality Management (TQM), and so-called "soft" methods such as standardization, visualization and the continuous improvement process [2]. The performance of a production system is generally rated according to the key performance indicators time, costs, quality, flexibility and productivity [3]. The increase of the system performance achieved by lean methods is caused by their effects on relevant production figures such as inventory, process and set-up time, machine availability, scrap rate etc. [4]. A survey about the introduction of lean methods in 1305 companies reveals that the turnover has increased by an average of 30% with a simultaneously reduced lead time by 30% and inventory by 20% [5]. In addition to these quantitative improvements there are also qualitative impacts like increased staff motivation and satisfaction and greater process transparency. It is difficult, however, to predict such potential improvements as they only become noticeable after the methods have been applied for a certain period of time.

Experience shows that lean methods have different effects on the production figures and thus on the key performance indicators of a production system [6]. The combination of several methods generally yields better results than the implementation of "stand-alone solutions" [5], but their impact on the system behavior may differ according to the type and characteristics of the method. These interdependencies are mostly known in a qualitative manner. Besides, the interrelations between lean methods and the basic influencing factors of a production system which considerably determine its structure and organization, such as market demand, number and type of products and the vertical manufacturing integration, have not been researched in detail yet. Figure 1 illustrates the complexity of the different hitherto nonpredictable interrelations between lean methods and production systems.

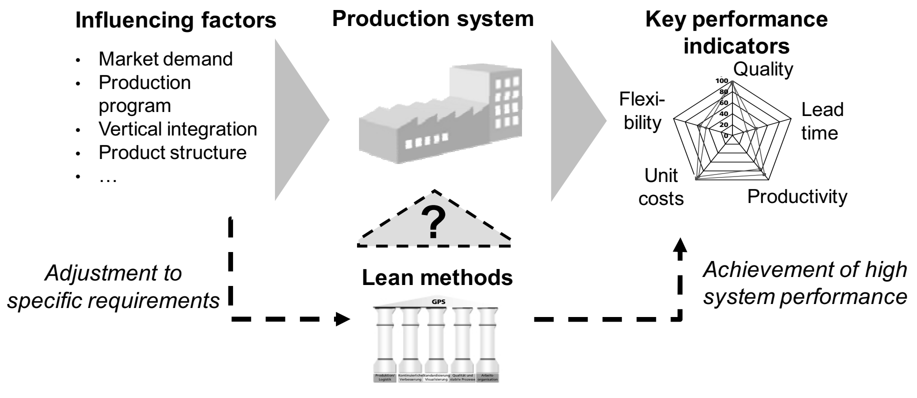


Figure 1: Interrelation between lean methods and production systems.

2 IMPLEMENTATION OF LEAN METHODS IN SMALL SERIES PRODUCTION

The rising competition caused by globalization does not stop at companies with small series production which do not have a choice but to improve their production performance. It is not sufficient, however, to just copy the lean methods that are successfully introduced yet in companies with large series production. They have to be adjusted to match the different influencing factors from the production environment and, accordingly, the specific process requirements [7]. Some of the most common factors distinguishing small series production include increased product variety, smaller production volumes, more variation regarding process scope and time required for each product variant, greater deviation of the product demand, higher degree of vertical integration and comparatively less capacity of resources (i.e. employees, means of production, space) [8]. The relevance of the key performance indicators may also differ, as small series manufacturers are required to be very flexible in terms of customer requests and order changes at fairly short notice before start of production. They need to ensure a certain level of flexibility, which often causes increased buffer sizes and longer lead times. Thus, flexibility as key performance indicator takes priority over the lead time as opposed to the targets of large companies.

This gives rise to the question as to how "lean" companies with small series production should or can be and what are their limitations? Which of the lean methods will help to reach the desired key performance indicators against the background of specific influencing factors?

A recent survey states that up to date only few small and medium sized companies with small series productions use lean methods such as just-in-time, Kanban, TPM or TQM, although considering them to be of major importance [9]. Therefore, more than 80% of companies consider research on the selection and introduction of lean methods in small and medium-sized companies with single unit production or small series production necessary [10]. These companies lack a pragmatic tool for the prognosis of the quantitative effects of lean methods in a given production setting in order to justify the appropriate measures and investments. Their generally limited resources do not allow any reorganization strategies with the risk of not achieving the intended increase of system performance. Besides, these companies rarely apply simulation tools for the analysis and evaluation of the production system as there are very cost and time consuming [11].

3 IDENTIFICATION OF TYPES OF SMALL SERIES PRODUCTION

There is a huge variety of small series production systems differing in the attributes of the above-mentioned influencing factors. These attributes can be categorized as illustrated in Figure 2 according to their impact on the structure and organization of the system. *Product types* differ either in terms of process times for identical process steps or in terms of number and order of the process steps as such (product variants with only minor differences concerning the production process such as color are listed under the same product type). *Volume* refers to the number of products manufactured per year. *Demand behavior* is used to describe the distribution of the required quantity per product and the temporal distribution of orders (taking seasonal peaks into consideration). *Vertical integration* refers to the share of in-house production and refinement processes in the value chain. The term *product structure* relates to the quantity, geometry and complexity of the individual product components. Both last mentioned factors are closely interrelated with the effort necessary to achieve the required product quality. The *variation of process time* as an indicator of process stability represents the extent to which process times required for individual process steps may deviate from a reference value (i.e. depending on the automation level). The rather qualitative factor, *prognosis accuracy*, is an indicator for the degree to which demand predictions regarding volume and time of request match the real order situation.

The object of current research activities at the Institute of Production Science (wbk) sponsored by the ministry of Science, Research and Art of Baden-Württemberg represent four companies with small series production. Based on the different attributes of the respective production system four typical types (A-D) of small series production as shown in Figure 2 were defined.

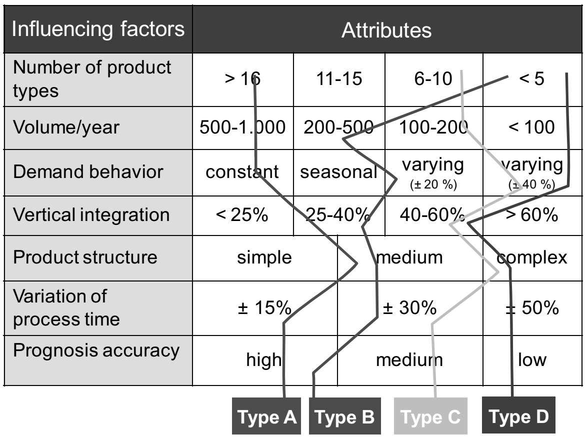


Figure 2: Types of small series production.

4 INTERDEPENDENCIES OF LEAN METHODS IN SMALL SERIES PRODUCTION

The research work at the Institute of Production Science (wbk) aims at the quantification of interdependencies between selected lean methods, relevant production figures and the resulting key performance indicators for each of the four defined types of production systems. Subsequently, recommendations will be derived on how to implement these methods efficiently. The interdependencies are identified by sensitivity analyses with the simulation software Plant Simulation® based on real data from the companies.

The investigations are focused on lean methods which have an immediate impact on production and logistic processes and include the **push or pull system**, the **Kanban principle**, **production leveling**, **TPM**, **TQM** and **SMED**. Soft methods such as standardization, 5S or the continuous improvement process are difficult to simulate by means of software tools and have, therefore, not been included in the analysis.

The simulation studies consider the lead time, productivity and quality as key performance indicators of the production systems. Costs are currently not included and need to be evaluated separately. (For example that the same set-up time may lead to different costs due to the type of tools or the operator's qualification required at the particular machines.) The relevant production figures for the study with significant effects on the key performance indicators are specified respectively in the following examples of some initial analysis results. The production period within the simulation study is limited to one year.

4.1 Quantification of interdependencies between lean methods and production figures of type B

This chapter illustrates some of the analysis results for the type B production system (see Figure 2) and explains them in detail. The small series production of the considered company is structured according to the job shop production principle and includes seven work stations: six conventional production machines, such as milling, turning and drilling machines, and one manual assembly station. In front of each work station, a buffer is located. Within the system three different product types with a medium degree of complexity are produced. Two of them are machined on the same machines with the same process order but differ with regard to times required for the turning and drilling process. The overall process time of the high-quality products of the third product type exceeds the time required for the other two types as the former is joined by means of adhesion instead of welding and then polished on another machine. The overall annual production volume amounts to approximately 500 products.

The effect **Total Productive Maintenance (TPM)** has on the production system and in particular the breakdown behavior of machines was looked into with regard to three different maintenance concepts: 1. *Autonomous maintenance:* standard maintenance activities at short intervals and simple to medium complex repair work done by the machine operator, 2. *Preventive maintenance (internal):* all maintenance work at longer intervals and repair carried out by an internal maintenance operator, and 3. *Preventive maintenance (external):* all maintenance activities at longer intervals and repair work performed by an external service provider with profound know-how about the machines. These TPM concepts primarily affect the two production figures Mean-Time-To-Repair (MTTR) and Mean-Time-Between-Failure (MTBF) which are characteristic for machine breakdowns.

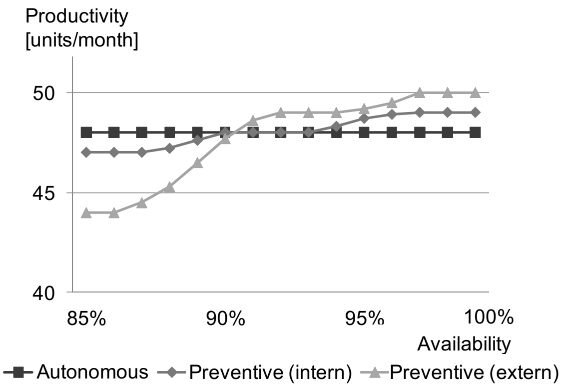


Figure 3: TPM: Correlation of productivity and availability.

Figure 3 shows that for the above-described production system combined with a preventive external maintenance concept, availability of the bottleneck machine is required to be at least 92% to reach a higher productivity as for the other two concepts. This originates from the longer idle times in case of machine breakdowns lasting several hours as opposed to the duration of less than one hour if the operator is able to perform small repair jobs himself.

To evaluate the effects of **Total Quality Management (TQM)** on the production system two different concepts were compared in this example: 1. *operator self-checks* at all work stations according to an inspection schedule, 2. *final inspection* after the manufacturing stations before the assembly station done by a specific operator. In case of an occurring failure based on statistics about the failure rate, the product is either immediately removed from the production process as scrap part or it is registered as rework. With the operator self-checks the product is reworked by the operator on the same machine or, if necessary, at a proceeding station within the process. In the scenario with the final inspection, the defective part is reintroduced into the process at the work station which caused the defect and reworked.

By the example of the product with the highest production volume Figure 4 displays the average values of the key performance indicators lead time (LT) and productivity (P) of the production system against the failure rate of the overall production volume.

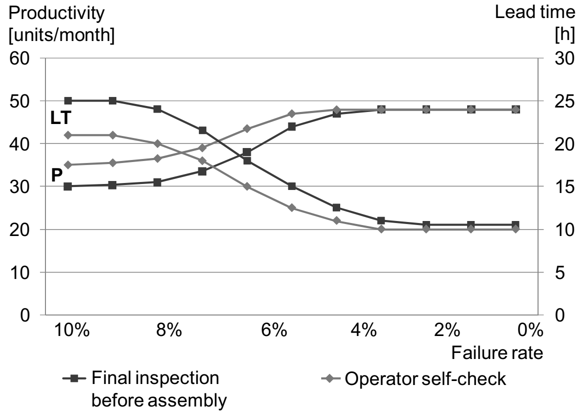


Figure 4: TQM: Correlation of productivity, lead time and failure rate.

The graphs in the figure demonstrate that a very poor product quality (failure rate ≥ 8.0%) results in a higher average lead time for the final inspection, since the defects are only noticed after the entire manufacturing process and the parts have to be transferred to the appropriate work station. At the same time, the productivity is less than in case of the operator self-checks. As due to the fact that the failure is not immediately detected, subsequent failures are occurring probably leading to more scrap parts. From the failure rate of less than 8.0% both figures show a considerable improvement, whereas the values remain relatively constant reaching a rate of 3.4%. From this failure rate, however, there are no major differences in terms of time or units per month for neither of the two TQM concepts.

The interdependencies of the lean method **Single Minute Exchange of Die (SMED)** aiming to reduce set-up times can be illustrated by the example of the bottleneck machine process (turning) (see Figure 5).

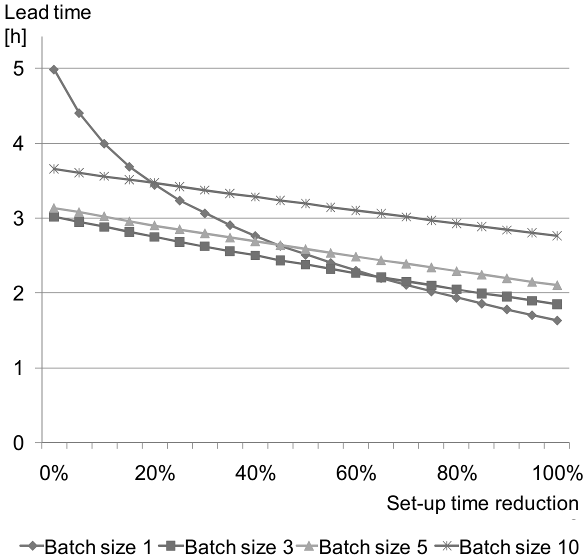


Figure 5: SMED: Correlation of lead time, set-up time and batch size.

Up to a batch size of 3, the lead time of a product manufactured on this machine clearly decreases arithmetically. And the smaller the batch size, the smaller is the lead time. Though, the lead time of a one-piece-flow production will only be less than that with the batch size of 3, if the set-up time can be reduced by approximately 64%.

4.2 Evaluation and recommendations

By means of these quantified interdependencies between selected lean methods and the key performance indicators of small series production systems, guidelines for the appropriate combination of lean methods and their respective characteristics for each type of production system will be derived in order to achieve a high performance.

These recommendations are also meant to facilitate the evaluation of an effort-benefit ratio with regard to the implementation of lean methods. This allows companies to ascertain prior to implementation how much effort will be required and whether the potential benefit is enough to justify this effort. As far as the introduction of TQM measures is concerned, Figure 4 shows that actions taken to raise the product quality, like operator self-checks, successfully improve the performance in terms of lead time and thus productivity. Figure 5 illustrates an example of how an increase of flexibility of the production system can be achieved by changing batch production to one-piece-flow but may require an excessively huge effort to reach the required set-up time reduction, for example through set-up work-shops with operators.

5 SUMMARY

Lean methods cannot easily be transferred from large series productions like in the automotive industry to small series productions. This comes from the different correlations between the influencing factors of a production system such as production program or market demand and the production figures like inventory, set-up time, availability, failure rate etc. Up to now, except of a general approach for the introduction order of lean methods [12], no rules or guidelines exist to efficiently implement lean methods in small series productions, which means that there is no decision-making support regarding essential measures and investments.

Research activities at the Institute of Production Science (wbk) look into the quantification of interdependencies between selected lean methods, production figures and the resulting key performance indicators such as lead time, productivity and quality with regard to small series production. The results can be used as a basis to determine the type and characteristics of lean methods best suited for the specific requirements of a certain type of production system to achieve a high system performance.

6 REFERENCES

[1] Womack, J.P., Jones, T., Roos, D., 1990, The machine that changed the world, New York, Rawson Assoc. 323.

[2] Spath, D., 2003, Ganzheitlich produzieren: Innovative Organisation und Führung, Stuttgart, LOGX Verlag.

[3] Hon, K.K.B., 2005, Performance and Evaluation of Manufacturing Systems, Annals of CIRP, 52/2: 675-690.

[4] Syska, A., 2006, Lean Production Controlling, Industrie Management, 22/4: 33-36.

[5] Lay, G., Dreher, C., Kinkel, S., 1996, Neue Produktionskonzepte leisten einen Beitrag zur Sicherung des Standorts Deutschland, Mitteilung aus der Produktionsinnovationserhebung, Fraunhofer ISI, Karlsruhe.

[6] Zäh, M.F., Aull, F., 2006, Lean Production-Methoden und Interdependenzen, wt online, Jg 96, H.9: 683-687.

[7] Dombrowski, U., Palluck, M., Schmidt, S., 2005, Ganzheitliche Produktionssysteme im Fokus der Fabrikplanung, 6. Deutsche Fachkonferenz Fabrikplanung, Ludwigsburg.

[8] Aurich, J.C., Drews, O., Fuchs, C., Wagenknecht, C., 2006, Produktionssysteme im Mittelstand, wt online Jg 96, H.5: 302-307.

[9] Herrmann, C., Bergmann, L., Thiede, S., 2007, Gestaltungselemente und Erfolgsfaktoren - Ergebnisse einer empirischen Umfrage unter produzierenden Unternehmen, Intelligenter Produzieren, VDMA Verlag, H.4: 20-22.

[10] Korge, A., Becker, M., Scholtz, O., 2002, Ganzheitliche Produktionssysteme – Erhebung zur Verbreitung und zum Forschungsbedarf, Fraunhofer IAO, Stuttgart.

[11] Blumenau, J.-C., Fritz, J., 2006, Modellierung und Auswertung in Sekundenschnelle – Kombination analytischer und simulationsbasierter Verfahren, wt online, Jg 96, H.1: 30-35.

[12] Reinhart, G., Zäh, M., Habicht, C., Neise, P., 2003, Einführung schlanker Produktionssysteme, wt online, Jg 93, H.9: 571-574.

Further Potentials of Smart Logistics

Günther Schuh[1], Sebastian Gottschalk[1], Tim Höhne[1], Philipp Attig[1]
[1]Laboratory for Machine Tools and Production Engineering WZL, RWTH Aachen University, Germany

Abstract
Within the research project 'Smart Logistics' dynamic control principles of logistic systems have been developed. Real-time information, allocated by RFID-supported Kanban boards, allows a dynamic tour planning of trolley trains. The implementation of the 'Smart Logistics' components and software systems shows savings of individual transportation tours and in the result savings of staff requirements. The availability of detailed information of actual production status avoids empty runs of the logistic resources. Furthermore, the research has revealed potentials beyond the realized effects of the project. With the high resolution information available, a dynamic redirection of the logistic train during its tour is feasible. This will lead to more efficient transports and an increased flexibility.

Keywords:
Production; Flexibility; Logistics

1 INTRODUCTION

The increased turbulence of corporate environments for production enterprises is characterized by shortened product lifecycles and cumulative variants. Customers demand smaller delivery lot-sizes with higher variance and shorter delivery times [1]. Most enterprises react by improving the flexibility of their production systems to shorten the reaction time [2, 3].

Therefore, the buffers at the work stations have to be reduced and the production lot-sizes have to be adapted to the customer demands. This leads to a challenge for the logistic system. Recently, the development of flexible production systems has been focused on the flexibility of value-adding resources, whereas the contribution of logistics has been underestimated so far [4, 5, 6].

For today's production systems a logistic concept, which is both economic and flexible, is necessary. Generally, standardized processes and resources lead to efficient systems. Forklifts for example provide a very flexible transport system but are inefficient due to many empty runs. This problem is strengthened by smaller lot-sizes because of low carrier load. Typical objects to standardize are carries, for example trolleys. They provide an efficient equipment management and save costs due to scale effects. For transportation of these trolleys in spacious shop floors, logistic trains are a good solution. They can supply different workstations with the standard trolleys and are efficient, thanks to the load of up to 12 trolleys. Trolley trains are typically used with static routes on the shop floor to supply the work station in fixed time slots. Kanban mechanisms are able to visualize and deploy the material demands. Instead of using driver-less transport devices, personnel guide the vehicles. The trolleys are also able to be moved by hand from the train to the workstation nearby.

The demand of shorter reaction times in the production also challenges the proposed trolley trains due to the shorter time slots and therefore means an increased quantity demand of trains and personnel. A dynamic tour planning for the trolley trains can be a solution and will be described in this paper.

2 DYNAMIC TOUR PLANNING FOR INBOUND LOGISTICS

2.1 Challenges

To realize a dynamic tour planning, a concept of high resolution mapping of the production status is necessary. The actual production status is the basis for individual material demands and therefore transportation orders. Events like material requirements have to be generated directly and electronically on the shop floor. Especially the conventional Kanban mechanism is too slow to react to the individual material demands.

Logistic systems have to guarantee a stable production in every situation. Therefore, a safety concept in case of system failure is necessary. Kanban boards equipped with RFID-Sensors (RFID = radio frequency identification) for automated reading of the Kanban cards can provide this fall-back solution. In case of an electronic problem, the boards can be used in the common way: With static train-routes and visual management. Due to the demanded adaptability of production systems to changing product lines, the logistic system also has to be adaptable regarding the layout and place of the RFID-Kanban boards [7].

The tour planning of the described inbound logistic systems has to face several challenges: Several trolley trains have to be planned simultaneously. Each of the trains can carry up to 12 different trolleys and therefore can combine the same amount of orders in one route. For each of these orders a distinct time slot is defined by the individual buffer size at the workstation. In addition, very short calculation times for the tour planning are necessary due to the permanent dynamic material demands.

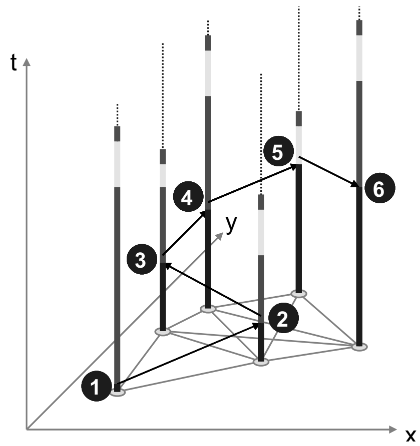


Figure 1: Time slots and tour planning.

Figure 1 shows the model of several material demands inside a shop-floor production. The x-y-area represents the production layout whereas the timeline is arranged vertical. Each demand has an individual time limit for the delivery (green: buffer size still sufficient; yellow: last trolley volume in progress; red: production hold-up). The tour planning has to consider these limits when calculating a path-efficient route (here: sequence one to six).

2.2 State of the art

Auto-ID systems

Wireless communication technologies are emerging in many use cases. However, the use of wireless technologies is a great challenge to production enterprises. Indeed, WLAN, Bluetooth and RFID-technologies have been used professionally for a long time in office communication, in stores and warehouses and also in traffic engineering. Although the fundamental research of RFID-Technologies delivers more and more opportunities, wireless communication in production is not stable in varying circumstances, yet. The multiplicity of possible disturbances inside the production environment necessitates the use of a constant environment for an automated identification. Therefore, the use of passive RFID-Tags directly on the logistic resources (for example on the trolleys) is difficult and error-prone [8].

Tour planning algorithms

An efficient tour planning algorithm is necessary for the depicted logistic system. Many solutions for tour planning exercises under different circumstances are available in science and as IT-tools at the market. The logistics of express delivery companies are supported by planning systems for effective tour planning. From a logistics point of view, many similarities can be identified with respect to production logistics. Therefore, these systems may be adapted and used in a production logistics context. However, in the past this transfer has hardly taken place. The main reason is the requirement of solutions at short-term intervals, while day-to-day planning is usually sufficient for the logistics in express companies [9]. Planning solutions based on the Travelling Salesman Problem (TSP) and Vehicle Routing Problem (VRP) are insufficient. The carriage of several orders at the same time prohibits the use of several existing fork-lift control systems.

3 SMART LOGICTICS APPROACH

'Smart Logistics' is a project funded by the German Federal Ministry of Economics and Technology. Within this project a flexible material supply system was developed using innovative information and communication technologies.

3.1 Production monitoring and information framework

RFID-Kanban Board

The 'Smart Logistics' approach demands that information or requisitions are generated instantly and electronically. Further, the system has to be easily operable by staff and also provide safety concepts, in which a non-electronic control mode is usable.

From this point the idea of a semi-electronic Kanban system was developed and advanced. In opposite to common electronic Kanban systems working with a push-button instead of Kanban cards, the 'Smart Logistics' concept uses ordinary Kanban cards equipped with RFID-tags. This composition ensures safety in case of a system failure. First implementations of this idea have been built up in practice, but as yet they use the electronic information only for visualization purpose instead for event-driven triggering of further processes.

The realization of these boards is a technological challenge. On the one hand, the boards must be adaptable both in different functional areas of the factory and in the size of these areas, in order to guarantee cost-efficient installation. On the other hand, the boards have to be able to work reliably in real-life production environments (dust, magnetism, temperature etc.). Therefore, specially matched antennas and control components have been developed [10].

Information framework

The 'Smart Logistics' Kanban-Boards are integrated into a communication infrastructure of a wireless network (WLAN). The technological challenge of this solution is the reliability, in particular the avoidance of disturbances by machines in terms of shielding and electromagnetic fields [11].

To compensate connection problems and assure secure data exchange a network manager was implemented. It coordinates and prioritizes all data flows within network components thus critical data reaches its destination and is not thwarted by secondary information. Additionally, the network use for additional external data is enabled by the network manager.

3.2 Negotiation based solution determination

The route planning for multiple operators is solved by a negotiation based approach in which a route planning algorithm schedules every train tour. Due to this negotiation based planning quick decisions can be made. The approach is more robust regarding modifications and variations, as stalled resources do not take part in negotiations. In addition, optimal configurations of available resources are determined on a short-term basis, hence employees and equipment can be combined flexible. The necessary rules are much less complex than those in similar hierarchical control systems. They are simply formulated for individual resources from their individual perspective. Diversity and complexity become controllable and robust operations are enabled [12, 13].

During the negotiation period of the occurring task an optimization algorithm based on the TSP solves the routing problem for every transportation route. To ensure a high computing performance the algorithm applies dynamic programming. Dynamic programming divides complex optimization issues into sub problems which are solved not in one single optimization model but in a logic sequence of smaller sub models which certify an optimal solution.

Because of the time frame-based type of the standard TSP, the applied algorithm of Simonetti and Balas was modified in order to model the negotiation process as stated above. Furthermore, it supports the minimization of the necessary computing time [14].

Applying the Simonetti and Balas algorithm to flexible material supply processes requires several adjustments. During a single tour a vehicle can pass more than one marketplace; this leads to certain restrictions in the sequence of material distribution. Some delivery orders have to wait until the vehicle has passed a marketplace where the ordered material is picked up. This restriction had to be implemented in the original timeframe algorithm.

However, the model still has two major restrictions: Only closed loops with the same starting and ending point can be planned in the current solution. After the tour of one train has started, it is fixed and cannot be adjusted. All incoming orders have to wait until the next tour is planned, even if integration into an existing route would be efficient and therefore recommendable.

4 CLASSIFICATION OF SMART LOGISTICS SCENARIOS

In order to classify innovative logistics systems, the interactions between the enablers and their individual effects must be understood and categorized. To rank logistics systems by their potential, an evaluation model has been developed and is illustrated in the following paragraphs.

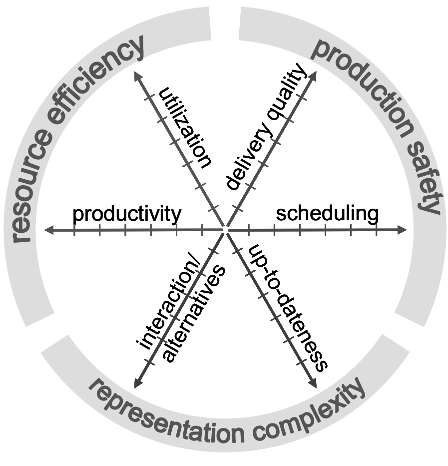


Figure 2: Classification concept of logistic potentials.

'Smart Logistics' systems can roughly be classified according to the following criteria: Resource efficiency, production safety and representation complexity. This classification system is helpful to identify the potential of a 'Smart Logistics' solution as well as its technological focus.

The level of resource efficiency is determined by the parameters load and productivity. Load is defined as the ratio of logistical tasks to working hours; productivity rates the qualitative work content within logistical tasks. A quantitative improvement can be achieved by increasing the process efficiency in logistics and modifying the process environment inside the production. For internal improvements the load and productivity of logistics resources has to be optimized by reducing or avoiding empty runs and high efficient bundling. The external potentials can mainly be exploited by reducing the number of production defects: A lower rate of scrap and rework for example implies fewer bottle necks and downtimes in supply. In addition, the number of mistakes due to incomplete information is reduced.

The evaluation of production safety is based on the quality and adherence of delivery dates in supply. The supply quality relates to the correctness of placing the right material at the right place. The example of semiconductor production shows the effect to the scrap rate. Silicon wafers look identically at different process stages leading to a high risk of confusion and creating expensive scrap. It also affects the process stability in production. Incorrect materials must be exchanged so that time gets lost when hold-ups occur due to undersupply. Adherence to delivery dates in supply relates to the right delivery time. It effects the stability of production processes regarding downtimes due to undersupply caused by tardy delivery and stock level.

The complexity in representation or mapping is another important element beside efficiency and safety and has to be evaluated. It can be determined by 'planning alternatives' and 'up-to-dateness'. High mapping complexity poses the main challenge in exploiting these potentials and indicates the benefit of digitalization. In general, production logistics systems can be enhanced in terms of planning and control. Mapping is enabled and limited by physical structures including digitalization hardware. In a system such as 'Smart Logistics', structures and processes have to be configured and initialized at the beginning. As long as the utilized technologies cannot be considered mature, a fallback concept for the dynamic steering system should be foreseen. Such 'non-electronic' fallback systems like fly-by-wire or electronic gas pedals without a mechanical backup system will very probably become obsolete. In day-to-day use the capabilities of 'Smart Logistics' should be reviewed at regular intervals and reconfigured if necessary.

Major improvements can be achieved in case of benchmarking a large number of planning alternatives, as this type of complexity involves increased system flexibility. If the focus is set on topicality, manually assessing of the logistics quality can be challenging. The identification of the technological focus of a 'Smart Logistics' solution is benefited and justified by the interaction model. Thus the efficiency of many planning alternatives can be influenced mainly by planning and optimization measures (reference scenario: multi-variant series production), permanent control is not required due to low claims concerning 'up-to-dateness'. If production safety is dominant, for example in semi-conductor manufacturing, permanent control is mandatory and can imply the need for a localization technology.

'Smart Logistics' is not a suitable solution for all needs. For instance, the technology is too expensive for forklift control systems at present. This application can be characterized by high requirements concerning 'up-to-dateness', but improvements in efficiency can hardly be achieved, as quick decisions must be made in most cases. Otherwise the number of planning alternatives is limited and negotiation based planning approaches are therefore not required.

5 FURTHER POTENTIALS

Fixing a tour after it has been started diminishes potential flexibility and avoids dynamic expansions of tours in the event of new orders appearing. The replanning of running tours by the possibility of task dispatching and rescheduling leads to lower response times. Reduced response times lead to smaller lot sizes and thus to new potentials in production, so that improvements and increases in production workload, flexibility in variants and structure can be achieved. This enables even more efficient reactions to disturbances in production.

Another modification of the tour planning algorithms focuses the tour design. Every tour has to end at the marketplace where the tour began. Regarding the trolley load of a train, this limitation is useful due of the stock balancing of each marketplace. But in combination with the continuing replanning of running tours this limitation is not longer essential. Trains can take up additional orders before ending their tour at the marketplace and go on with the tour after having reached the marketplace just as an intermediate stop. A standardized trolley concept supports this type of continuing tour planning. The only restriction remaining is the stock balancing of empty trolley between the marketplaces.

For this concept additional research is necessary. A dynamic planning and rescheduling process during the delivery runs demands assistance systems for the logistic staff. This assistance system also delivers the possibility of role-specific recommendations concerning different alternatives of execution. Wireless networks inside production are being analyzed within the 'Smart Logistics' project, but they show unreliable results in unstable environments.

6 SUMMARY

The growing demand of flexible production challenges the logistic processes. Lot sizes are decreasing and shorter reaction times of material supply are necessary. Accordingly, innovative solutions for flexible and adaptive logistic systems have to be developed.

Within the 'Smart Logistics' project an innovative approach was developed to provide flexible and dynamic logistic processes. Therefore, communication technologies like RFID and wireless LAN are used for information gathering inside the production system and for data transmission. Additionally, the data is processed across several software levels to secure a robust information handling and interpretation.

Manually operated logistics trains are used for efficient material transport in spacious shop floors. The routing problem is solved by individual agents in a negotiation-based planning system. The algorithm in use is able to calculate an optimal solution within seconds by recursion. Also, sideline jobs can be handled by negotiation of the agents. All developed modules and solutions have been verified experimentally by an industry partner and first results show the profitability of the system in use.

The research of potentials of this and other innovative logistic systems has designed a classification scheme which distinguishes the three dimensions 'resource efficiency', 'production safety' and 'representation complexity'. The classification model also shows further potentials for even more dynamic systems. Therein, the handling of even shorter reaction times with smaller buffers is possible.

7 REFERENCES

[1] Jones, D., 2005, Creating Lean Solutions. In: Jones, D.; Schuh, G.; Wiegand, B.; Womack, J. (eds.), 2. Lean Management Summit, Aachen: 17-28.

[2] Wiendahl, H.-P., 2002, Changeability – a New Challenge for Production, CIRP General Assembly, Scientific Technical Committees Meetings, San Sebastian, 51/1: 15-19.

[3] Nyhuis, P.; Wiendahl, H.P. 2004, 3-Sigma PPC – A Holistic Approach for Managing the Logistic Performance of Production Systems, CIRP General Assembly, Krakow, 53/1: 371-376.

[4] Valckenaers, P.; Heikkilä, T.; Baumgärtel, H.; McFarlane, D.; Courtois, J., 1999, Towards a Novel Manufacturing Control Principle. In: Van Brussel, H.; Valckenaers, P. (eds), Proceedings of the 2. International Workshop on Intelligent Manufacturing Systems, Leuven: 871-875.

[5] Pritschow, G.; Wiendahl, H.P., 1995, Application of Control Theory for Production Logistics – Results of a Joint Project. In: Annals of the CIRP - Manufacturing Technology, 44/1: 421-424.

[6] Schuh, G.; van Brussel, H.; Boer, C.; Valckenaers, P.; Sacco, M.; Bergholz, M.; Harre, J., 2003, A Model-Based Approach to Design Modular Plant Architectures. In: 36. CIRP International Seminar on Manufacturing Systems, Saarbrücken: 396-373.

[7] Tönshoff, H.K.; Schnülle, A., 2001, Customer Specific Design of Highly Productive and Reconfigurable Manufacturing Systems, Production Engineering, 13/2: 55-58.

[8] Lodewijks, G.; Veeke, H.M.P.; De La Cruz, A.M.L., 2006, Reliability of RFID in Logistic Systems, In: Proc. IEEE International Conference of Service Operations and Logistics, and Informatics, Shanghai: 971-976.

[9] Wildemann, H., 1999, Logistikstrategien, In: Eversheim, W.; Schuh, G. (eds.): Betriebshütte - Produktion und Management, 7. Ed., Berlin: 15_1-15_109.

[10] Bärwald, W.; Baumann, S.; Fuss, T.; Keil, R.; Leitzke, C.; Richter, K., 2007, 'Smart Logistics' RFID-Equipment for Production Logistics, 1. Annual Conference RFID Eurasia, Istanbul: 1-6.

[11] De La Cruz, A.M.L.; Veeke, H.P.M.; Lodewijks, G., 2007, Applicability of RFID in the Prognostics of Logistic Systems, IEEE International Conference on Service Operations and Logistics and Informatics, Philadelphia: 1-5.

[12] Montalvo, A.C.; Higuera, A.G., 2007, RFID enhanced Multi-Agent based Control for a Machining Cell, In: International Journal of Flexible Manufacturing Systems, 19/1: 41-61.

[13] Ritter, A.; Richter, H.; Westkämper, E., 2003, Agent Life Cycle Management for Manufacturing Systems, CIRP 36. International Seminar on Manufacturing Systems, Saarbrücken.

[14] Balas, E.; Simonetti, N., 2001, Linear Time Dynamic Programming Algorithms for new Classes of Restricted TSPs, A Computational Study, Informs Journal on Computing, 13/1: 56-75.

Production Management in SME Networks - Evaluation of Potentials and Achievements

Christoph Herrmann[1], Lars Bergmann[1], Philipp Halubek[1]

[1]Institute of Machine Tools and Production Technology (IWF), Technical University Braunschweig, Germany

Abstract

During the last years, the attention on SME networks has been continuously increased. Especially in the production business, SME more and more organize SME networks in order to extend their production capabilities. While a lot of potentials have been tapped by sharing production resources and knowledge, or by collaborative purchasing, the current achievements should not deceive the idle lying potentials within SME networks. Based on a case study of a SME network(s) in Germany, this paper presents a methodology that allows the evaluation of the potentials and the achievements in SME networks.

Keywords:

Production Management; Interorganizational Network; Small and medium-sized Companies

1 INTRODUCTION

To achieve competitive advantages, the formation of interorganizational co-operations such as strategic alliances, joint ventures and networks have become a common element of small and medium-sized companies' (SME) strategies in all kinds of different industry sectors [1] [2] [3]. This trend is emphasized by the fact that more than 20,000 interorganizational networks have been reported world-wide over a period of only two years [4]. Motives for founding interorganizational networks are advantages such as mutual learning or sharing of production capabilities. Moreover, production networks allow to level and balance the workload and capacities of the partners in terms of virtual factories [5]. These advantages are of increasing importance for realizing strategic goals [4] [6].

While a lot of achievements in SME networks could be realized within the last years, most of the founded SME networks fall apart after a certain time and were canceled [7]. As rather obvious fields for cooperation and improvement are exploited right at the startup phase of most networks, it becomes harder over time to take advantage of new fields of cooperation. As a consequence, a stagnation of cooperation activities impedes the ongoing creation of mutual benefits for the network partners. The maintaining of introduced network activities as well as the ongoing identification of new network activities require efficient information and communication structures in the network and inside the SME. But the interconnection of SME processes on the network level remains a weak spot that impedes ongoing network activities.

2 LIFE CYCLE PHASES AND ORGANIZATIONAL STRUCTURES OF NETWORKS

Empirical studies show evidence that rather general recommendations for the development and management of SME networks can only hardly be given due to the complexity and multitude of influencing factors [8]. To reduce and cope with the complexity of SME networks, many authors propose analytical approaches and focus on structuring and classifying interorganizational networks regarding certain characteristics, such as the purpose of the network or the way of leadership [7] [9] [10] [11].

Another important criteria to classify interorganizational networks are the consecutive life cycle phases, networks pass through. The classification of networks according to their life cycle phases allows to derive and assign network life cycle phase specific management tasks. By that, consecutive network management tasks can be identified along the life cycle phases of interorganizational networks.

2.1 Life Cycle Phases of Interorganizational Network

A life cycle oriented perspective on interorganizational networks allows to depict a wide diversity of networks with different objectives, different compositions of network partners, or different organizational structures. According to Killich and Luczak, four consecutive life cycle phases of interorganizational networks can be distinguished [12] [13] (Figure 1). The four life cycle phases of interorganizational networks should not be considered as strict sequentially sequences, as they tend to overlap and repeat themselves [14].

To support the management of interorganizational network, a wide diversity of network management approaches have been developed. These approaches can be classified according their focus on specific life cycle phases of networks [5] [15] [16] [17] [18] [19] [20] [21] (Figure 1). While most network management approaches focus on the initiation and formation phase, a lack of approaches for the main phase, the leadership phase, can be identified. As a consequence of lacking general network management guidelines, the strategic and operations management in SME networks generally show significant weaknesses.

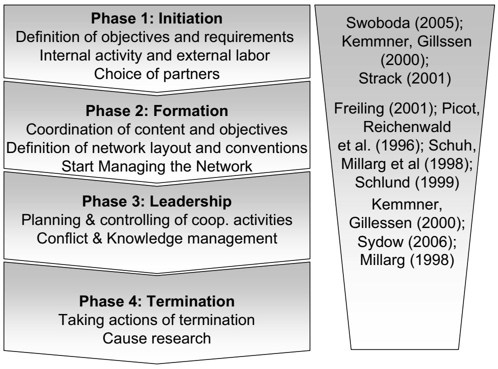


Figure 1: Life cycle phases and related management approaches for interorganizational networks.

During the initiation phase, decisions on network partners and the definition of network objectives and commonly accepted rules and conditions are of special concern. Thus, within this first network life cycle phase, normative and strategic management decisions of the founding network partners have to be taken in order to constitute the purpose, the legal form, and the organizational structure of the new network. By that, the normative and strategic management of the network is constituted. Within the second phase, the formation phase, the operations management has to be set up in order to operationalize the network layout and coordination functions.

After starting the network, the leadership phase requires to constitute a planning and coordination function. By that, the leadership phase incorporates the operations management activities that fulfill the purpose of the network including for example, conflict and knowledge management activities. Moreover, the leadership phase requires to constitute control mechanisms that provide an action framework and information channels between the strategic and operations management. The close connection of strategic and operations management activities is a necessary prerequisite for continuous adaptation processes and for ensuring the viability of a network. Nevertheless, after a certain time of cooperation, interorganizational networks can end. If the defined conditions for the termination of a network are fulfilled, the network enters its last life cycle phase.

2.2 Organizational Structures of Interorganizational Networks

Interorganizational networks are organizations that can be regarded as social-technological systems [22]. From a systems perspective, the network partners represent the elements of the system and the relations between the network partners are for example information and coordination channels that build the basis for transactions like business processes. By that, the systems' elements and relations form the structure of the network system. To gain insights into the potentials and achievements of networks over different life cycle phases, a structural systems perspective of networks can be drawn on.

If an organization, such as a network, wants to maintain it's identity over time, it needs to develop a structure and certain functions that enable the viability of the organization. The Viable System Model (VSM) developed by Beer [23] [24] comprises of a sufficient and necessary set of organizational functions for the viable organization, whereas viable means maintaining identity within a particular chosen environment through time (Figure 2). As the VSM is a recursive system, each VSM contains and is contained within other viable systems [23]. By that, the VSM has a high potential to support the design and diagnosis of complex organizational interactions at the structural level [25]. Based on the structure of the VSM, interorganizational networks can be depicted and analyzed regarding their capability to maintain viability over time.

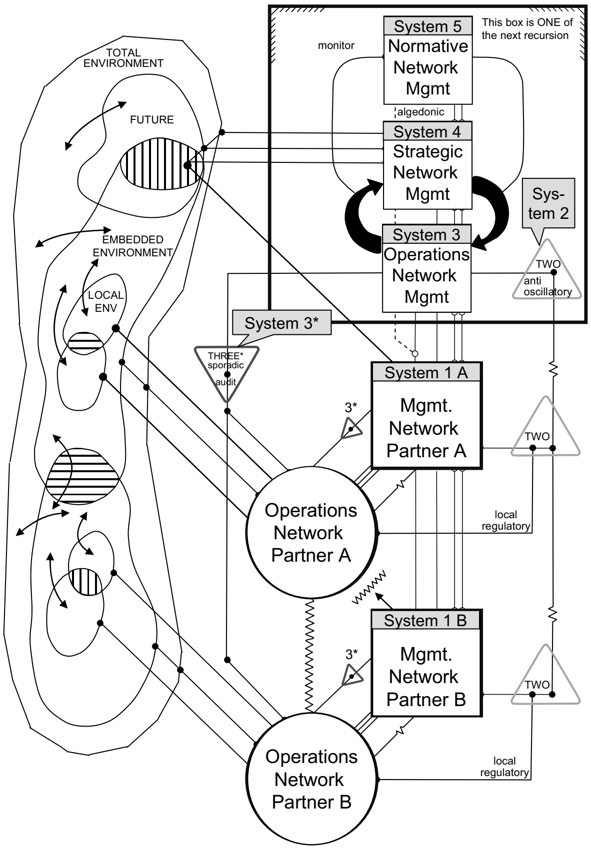


Figure 2: The Viable System Model, network oriented adaption according to [22].

In the VSM, each sub-system is responsible to fulfill specific system functions.

The Systems 1 (S1) stand for everything that is done in the organization, for example for its operations. Considering interorganizational networks from the VSM perspective, the S1 represent the network partners within the network that constitute the purpose of the network system. Between the network partners transactions are applied, for example in terms of joint order processing. Each network partner is connected to his local environment, while the network as a whole is connected to a broader environment that comprises of the local environments. The management of each S1 network partner is fulfilled by managing directors of the SME.

System 2 (S2) comprises all activities and resources involved in the coordination between the operative S1 units. S2 represents the coordination mechanism within the network and, thus, coordinates the network activities between the partners. According to the recursive structure of the VSM, the network partners comprise of the five subsystems, too. Therefore, both the network and the SME comprise of S2 functions. Consequently, the S2 of the network needs to be connected to the S2 within the network partners. In practice, this may require to connect several enterprise resource planning tools over different companies and to uplink the information to the network level for planning and optimization reasons. Other examples for realizing S2 functions are the information exchange via emails, meetings, or workshops.

System 3 and System 3* are responsible for all activities and resources that bring about the optimizing of the operations of the individual S1 systems. For that, the S3 needs information on the status of the S1 that are provided by S2 and S3*. The S3 and S3* of the network represent the operations network management that is responsible for the present-day network activities. In practice, the S3 function often is institutionalized by a central network manager.

System 4 stands for all the activities and resources that serve to observe the environment and to gain experience from it and to derive and develop strategies to be developed for the future. In a network, the S4 is responsible for the strategic network management and needs to develop plans for future network activities and thereby for the adaptation of the network, too. For example, the S4 needs to develop plans to affiliate new network partners or to initiate new fields of network activities. For that, the S4 function can be institutionalized by a network manager or a strategic network management board.

System 5 represents all the normative rules and regulations that apply in the organization, such for example ethical attitudes and normative rules. Within the network, the S5 function comprises of network rules and shared network values that constitute the network purpose. These rules have to be set up by the managing directors of the SME partner.

2.3 Network Development and the Challenge of Network Viability

In order to adapt an interorganizational network to changes within the environment, the network requires the capability to change its structures, operations and the way its network-related activities and the individual network partner driven operations are linked together. As a consequence, networks needs to comprise of organizational structures that provide mechanisms for adaptation to both its external environment and the internal environment on all recursive levels.

The five subsystems of the VSM link together their functions to provide a configuration of interconnected, information loops that convey information generated from and by management activities. By that, the VSM provides structures of organizations including key processes, communications and information flows [23] [24] [26]. Within interorganizational networks, these structures and information flows are a prerequisite for the viability of the network as they are required to constitute and adaptation process of the network. Thus, to maintain viability, interorganizational networks need to develop communication and information flows that allow adaptation processes in all sub-systems.

To realize adaptation processes in networks, specific structures and information exchange mechanisms have to be developed. This requires the active involvement of network partners in terms of normative, strategic, and operation management functions. Moreover, information and communication mechanisms need to be established according to the recursive structure of the VSM [22]. The recursive structure of the VSM and its interaction mechanisms emphasize that organizational structures, activities and behaviors at all recursions of interorganizational networks have to be considered when viability issues are discussed. Consequently, adaptation of interorganizational network requires to establish information channels and interaction mechanism from the network partner level to the network level (Figure 3).

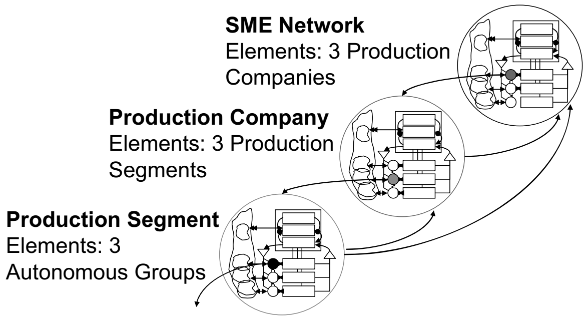


Figure 3: Recursive structure of the VSM, adapted figure according to [23] [27].

The interconnections of system functions and information channels across the network recursion and the network partner recursion become of special concern for the network development [26] [27] [24]. According to the S5 function, the normative values of the network should be in alignment with the normative values of the network partners over all levels of recursion [22] [26]. From the perspective of the VSM, this capability represents the balancing of autonomy and cohesion of the network partners within a network. The strategic network management (S4) needs to match the strategic plans of network with those of the network partners of the next lower recursion. For that, the provision of information on the ongoing S1 operations, capabilities, and plan (the network partners) up to the network level is necessary. This information enables to take decisions on changes for adaptation purposes such as the initiation of new network activities or the integration of new network partners. The realization of this information mechanisms becomes a key factor for the viability of networks.

3 CONCEPTUAL APPROACH FOR THE MANAGEMENT IN PRODUCTION SME NETWORKS

To support the management in SME networks, a conceptual approach has been developed. The approach focuses on minimizing network related transaction costs as well as on maximizing the overall mutual benefits of the partners by implementing information flows for the strategic and operations management. For that, the way information is acquired, provided, and exchanged within a network and especially between the central strategic and operations network management, is investigated

3.1 The Role of Information Exchange as a Prerequisite for the Operations Management in SME Networks

Insufficient information exchange can lead to lacking planning, control, or adaptation capabilities and may result in the termination of a network. Thus, establishing strategic management (S4) functions and operations management (S3) functions in SME networks requires routines for the daily planning, control, and adaptation processes on the network recursion. To establish information and coordination mechanisms in SME networks, certain representatives of the network partners have to be assigned as information exchange interfaces. As a direct personal and trustful information exchange between the SME network partner is an important factor, the connection between SME generally is establish by meetings of the managing directors of SME.

To develop future plans as well as to plan joint operations within a production SME network, information on the production capabilities, schedules, exchange plans, and costs have to be considered. Thereby, the provision of production capability and capacity related information becomes of vital importance. Rather heterogeneous IT-systems often impede a network-wide connection of enterprise resource planning (ERP) or production planning and control (PPC) systems.

Additional tools are required to provide, exchange and process information for planning, control, and adaptation routines on the network level. For that, a specific set of information on the production capabilities, capacities, schedules, costs, and conditions need to be exchanged between the value creating systems S1 of the network partners and the S2, S3, and S4 function on the network recursion. This paper presents an example for an instrument that enables SME networks to systematically exchange strategic and operational information between the SME recursion and the network recursion. By that, achievements and potentials can systematically be evaluated.

3.2 Acquisition and Provision of Production relevant Information

To take case-by-case decisions on sharing or exchanging production processes between SME network partners, an efficient exchange of production process relevant information and production capability relevant information is required. Short reaction times in the turbulent business environment require to provide valid and actual information on machine/process capabilities and schedules. For that, information assessment, processing and providing by adequate exchange processes become a basis for sharing and integration of production machines/processes.

Long-term and short-term related types of information need to be distinguished in order to effectively structure and provide required information. On the one side, long-term information are considered as mainly static values that describe production capabilities such as the number of drilling machines, their maximal drilling diameter, or a required coolant. Considering very specialized production SME characteristics, these process and technology information will not change rapidly and remain static over long-time periods. As to this, these long-term information can be implemented into a shared technology database that need comprises of all production capabilities of all network partners. Moreover, technology and capability related information can be used for joint strategic technology development such as the shared acquisition of new machines.

On the other hand there exist rapidly changing information like information the order processing or the production schedule and current capacity information. To enable a case-by-case oriented interaction between SME, operational information need to be provided in terms that can easily be integrated into the business processes of the partners. While the identification, collection and processing of schedule and capacity relevant information is subject to the daily business order process, the consideration and integration of information contents from other SME partners requires standardized compact forms for information representation.

To enable an easy assessment and provision of production relevant information for both, the case-based cooperation as well as for the strategic technology development, information have to be compacted to increase their significance. To minimize the transaction costs of cooperation, standardized conditions for spontaneous cooperation can be establishes by a database system. Moreover, the mutual consideration of production capacities and capabilities of network partners required to filter information according to specific process parameters and conditions. This requires the overall collection and comparative representation of production capabilities on the network recursion.

4 USE CASE - APPLICATION OF THE CONCEPTUAL APPROACH WITHIN A SME NETWORK

The presented conceptual approach has been developed and applied in close cooperation of the existing German KIM Network (Kooperationsinitiative Maschinenbau, www.-made-in-braunschweig.de) and the Institute of Machine Tools and Production Technology at the Technical University Braunschweig, Germany.

4.1 The SME Network KIM

The KIM Network is an interorganizational network that consists of 22 SME network partners with 12 up to 1000 employee each. All network partners are engaged in the fields mechanical, machine, and plant engineering. The network is clustered into 2 independent sub-network that consist of 10 and 12 network partners (Figure 4).

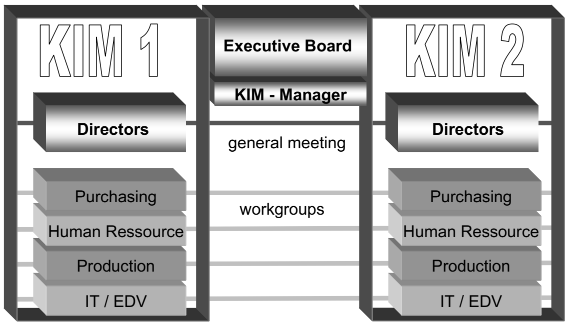


Figure 4: Structure of the KIM Network.

This splitting has been applied in agreement with all network partners in order to create small circles of persons as a prerequisite for mutual trust. The KIM network has the legal status of a registered association. The executive board of the association defines the strategic orientation of the KIM network and also provides general normative guidelines. Both sub-networks share the normative KIM values and the strategic function of the executive board. The KIM network passed through several different life cycle phases. During each life cycle phase, different activities and exchanges were applied and the focus on the main purpose of the network experienced some adaptation. At the startup phase in the year 2000, the network was founded by 11 SME in order to save jobs in the local region of Braunschweig by mutually leveling employee capacities. Based on a collective agreement, a network-internal employee pool was founded for rapid and easy exchange of employee between the network partners. Another field for cooperation that was identified right in the startup phase of the network was the cooperative purchase of raw materials, the cooperative disposal and the cooperative product development. Each field of cooperation is institutionalized by a working group that shares information and plans to control the network activities.

After the formation phase of the KIM network, the demand for a professional network manager was identified at the beginning of the leadership phase. Next, the operations network management was institutionalized by a central network operations manager. The central KIM network manager is financed by all network partners of both sub-networks. Nevertheless, the network operations manager allows the sub-networks to act autonomously and enables the sub-networks to separately plan and carry out network activities. Based on the insights into the capabilities and demands of the two sub-networks, the operations network manager develops future plans in cooperation with the executive board in order to adapt the objectives and fields of cooperation of the KIM network.

4.2 Use Case A: Information Exchange for the Operations Network Management

Within the growth phase of the KIM network, the demand for a closer connection and harmonization of the production systems had been identified. Due to the high number of similar or congruent production processes and process parameters between the network partners, the chance to substitute or outsource specific production processes within the network had been identified. By that, the overall production efficiency or quality aspects can be maximized, costs can be minimized or the delivery reliability can be increased. Nevertheless, the main bottleneck for the case-based connection or substitution of production processes has bee indentified as the high efforts for acquisition of the required information.

To provide information channels between the operational management of the network partners' (S1) that are also linked to the operations network management (S3), an information platform has been developed (see Figure 5). This internet-based machine/process database has been developed and introduced to serve as a prerequisite for mutual planning purposes. All network partners regularly provide their actual production capabilities in terms of rough schedules, possible process parameters, costs, etc. into the database.

Idem ID	Machine/ Process	Machine/Process Description
247	welding	Inverter 400 V WIG- Star 450 i
248	milling	CNC-milling machine DMU 125P
249	sawing	CNC-metal-powder flame cutting machine

Machine/Process Details	
Idem ID	249
machine/process	Sawing
slitter	to 150 mm
cutting area	1000mm x 2500mm
blowpipes	3
speed of operation	50 – 6000 mm/min

Figure 5: Screenshot of the machine/process database.

To further optimize the operational cooperation and connection of production processes, more detailed information cluster systems have been added to the database as a second step. The additional information cluster provides information on detailed capacities, schedules, order demands, technology potentials and changeover times in a compacted form to support the operational management of the network partners (Figure 6).

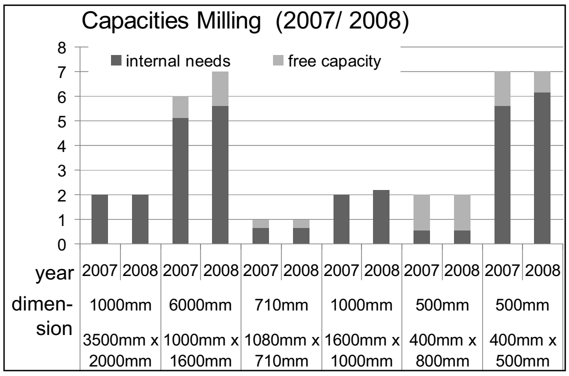


Figure 6: Information sheet for the provision of scheduled and free production process capacities.

By providing detailed machine/process related information to all network partners by the database system, all network partners can easily access and consider information on detailed production resources. Based on this information, production bottlenecks can be prevented by leveling the overall production utilization. Moreover, optimal network partners in terms of experienced experts can be identified for specific production processes. By providing detailed S1 information in-between the network partners, the network enables the partners to consider all production process capabilities of all network partners like their own. Thereby, the coordinating S2 and the planning S3 functions of the KIM network could be significantly improved.

4.3 Use Case B: Information Exchange for the Strategic Network Management

To foster the strategic long-term planning and adaptation process of the KIM network, possible strategic directions and activities regularly have to be assessed, evaluated, and matched. For that, information of machine/process database are complemented with strategic business information and represented in terms of technology portfolios (Figure 7).

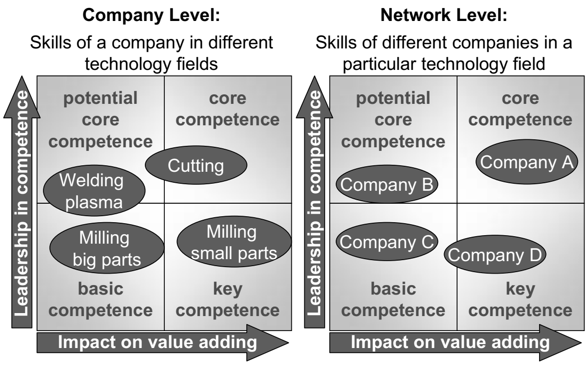


Figure 7: Technology portfolios.

Based on the portfolios, actual information on the operational S1 capabilities of the network partner recursion is provided to the S4 on the network recursion and the S4 on the network partner recursion. Thus, the KIM network partners use the portfolio to support decisions on joint future technology investments on the network recursion. For that, the technology attractiveness of certain production technology or process is considered with respect to the current strength and experience of the network partners. This procedure fosters a continuous and synergetic adaptation process of the KIM partners and of the KIM network.

5 CONCLUSION

The paper has investigated organizational structures and life cycle phases of interorganizational SME networks as a basis for the development of an approach for the operations and strategic management of SME networks in the production business. Based on the insights of the VSM, the important role of information exchange mechanisms between the network partners and the network recursion has been identified. The information exchange of S1 related production capabilities and processes between the network partners and to the operations and strategic network management were identified as key factors for the viability of networks.

In order to constitute an ongoing information exchange, an approach has been developed that allows to continuously evaluate achievements and potential field for future cooperation in SME networks. For that, a machine/process database and technology portfolios have been developed that enable to provide information on the S1 production capabilities and capacities up to the network recursion. By that, the strategic and operations network management is provided with information for planning, controlling, and adapting the network activities. The developed approach is demonstrated based on the German KIM network.

6 REFERENCES

[1] Aloysius, J. A., 2002, Research Joint Ventures: A. Cooperative Game for Competitors, European Journal of Operational Research:591-602.

[2] Zahn, E., Tilebein, M., 2000, Lernprozesse in Organisationen. In: Welge, M., Häring, K., Voss, A. (eds.), Management-Development: Praxis, Trends und Perspektiven, Schäffer-Poeschel, Stuttgart:117-137.

[3] Gulati, R., Gargiulo, M., 1999, Where do interorganizational networks come from?, American Journal of Sociology (104):1439-1493.

[4] Kale, P., Dyer, J. H., Singh, H., 2002, Alliance Capability, Stock Market Response, And Long Term Alliance Success: The Role of the Alliance Function., Strategic Management Journal:747-767.

[5] Schuh, G., Millarg, K., Göransson, Å., 1998, Virtuelle Fabrik. Neue Marktchancen durch dynamische Netzwerke, Hanser, München.

[6] Morath, F. A., 1996, Interorganisationale Netzwerke: Dimensions - Determinants - Dynamics. Nr. 15, University Konstanz.

[7] Das, T. K., Teng, B. S., 2003, Partner analysis and alliance performance, Scandinavian Journal of Management 19:279-308.

[8] Rautenstrauch, T., Generotzky, L., Bigalke, T., 2003, Kooperationen und Netzwerke. Grundlagen und empirische Ergebnisse, Eul, Lohmar.

[9] Dibbern, F. F., 2002, Mittelstand im Wandel des Wettbewerbs – Möglichkeiten und Grenzen kooperativer Unternehmensnetzwerke als strategische Option, Albert- Ludwigs-Universität Freiburg, Freiburg.

[10] Thielemann, F., 1996, Die Gestaltung von Kooperationen kleiner und mittlerer Unternehmen, Inst. für Angewandte Innovationsforschung, Bochum.

[11] Sydow, J., 2006, Management von Netzwerkorganisationen, Beiträge aus der "Managementforschung", Gabler, Wiesbaden.

[12] Killich, S., Luczak, H., 2000, Aufbau erfolgreicher Unternehmenskooperationen – Ein Leitfaden für mittelständische Automobilzuliefer, Forschungsinstitut für Rationalisierungs, Aachen. Sonderdruck 06.

[13] Killich, S., 2005, Kooperationsformen. In: Becker, T., Dammer, I., Howaldt, J., Killich, S., Loose, A. (eds.), Netzwerkmanagement. Mit Kooperation zum Unternehmenserfolg, Springer-Verlag, Berlin.

[14] Bronder, C., 1993, Kooperationsmanagement – Unternehmensdynamik durch strategische Allianzen, Campus Verlag, Frankfurt am Main.

[15] Swoboda, B., 2005, Metaanalyse und Kurzabriss. In: Zentes, J., Swoboda, B., Moschett, D. (eds.), Kooperationen, Allianzen und Netzwerke. Grundlagen - Ansätze - Perspektiven, Gabler, Wiesbaden:3-31.

[16] Kemmner, G.-A., Gillessen, A., 2000, Virtuelle Unternehmen. Ein Leitfaden zum Aufbau und zur Organisation einer mittelständischen Unternehmenskooperation, Physica-Verl., Heidelberg.

[17] Strack, J., 2001, Controlling virtueller Unternehmen. Konzept zur Flexibilisierung und Steuerung dezentraler Netzwerkstrukturen, Shaker, Aachen.

[18] Freiling, J., 2001, Resource-based view und ökonomische Theorie. Grundlagen und Positionierung des Ressourcenansatzes, Dt. Univ.-Verl., Wiesbaden.

[19] Picot, A., Reichwald, R., Wigand, R. T., 1996, Die grenzenlose Unternehmung. Information, Organisation und Management ; Lehrbuch zur Unternehmensführung im Informationszeitalter, Gabler, Wiesbaden.

[20] Schlund, M., 1999, Planungsinseln - Handbuch für Teamarbeit in indirekten Unternehmensbereichen als Transformationsprozeß zur lernenden Organisation, Fraunhofer IAO, Stuttgart.

[21] Millarg, K., 1998, Virtuelle Fabrik. Gestaltungsansätze für eine neue Organisationsform in der produzierenden Industrie, Transfer-Verl., Regensburg.

[22] Beer, S., 1985, Diagnosing the system for organizations, Wiley, Chichester.

[23] Beer, S., 1979, The Heart of Enterprise, Wiley, Chichester.

[24] Beer, S., 1981, Brain of the Firm, Wiley, Chichester.

[25] Espejo, R., Harnden, R. (eds.), 1989, The viable system model. Interpretations and applications of Stafford Beer's VSM, Wiley, Chichester.

[26] Malik, F., 2006, Strategie des Managements komplexer Systeme, Haupt, Bern.

[27] Gomez, P., 1978, Die kybernetische Gestaltung des Operations Managements. Eine Systemmethodik zur Entwicklung anpassungsfähiger Organisationsstrukturen, Haupt, Bern, Stuttgart.

Digital Virtual Holons – An Approach to Digital Manufacturing Systems

Hasse Nylund[1], Kai Salminen[1] and Paul Andersson[1]

[1] Department of Production Engineering, Faculty of Automation, Mechanical and Materials Engineering, Tampere University of Technology, Finland

Abstract

This paper presents an approach to digital manufacturing systems. The system consists of real and virtual systems as well as the digitally presented information and knowledge needed in the system. The virtual manufacturing system can be a replica of the real manufacturing system, or an idea of a new possible system visualized at operational, tactical and vision levels. The approach is based on principles of Holonic Manufacturing Systems (HMS), Fractal Manufacturing Systems (FrMS) and Service-Oriented Architecture (SOA).

Keywords:

Digital; Holonic; Manufacturing System

1 INTRODUCTION

Today's manufacturing enterprises face the challenges of globally distributed markets. They are forced to be context aware and adaptable to changes. A digitally presented manufacturing system is one of the key enablers. This, at best, shortens the time used in decision processes by offering a wider view (compared to that offered by the knowledge of an individual party) of the information and knowledge.

The research on digital and virtual manufacturing, factories and enterprises has no commonly agreed definitions, but they share some characteristics often found in the literature, for example [1-6]:

- An emerging and integrated approach to improve product and production engineering processes and technology.
- Computer-aided tools, such as modelling and simulation, for planning and analyzing real manufacturing processes.
- A framework for new technologies, including the collection of systems and methods.

This paper proposes an approach to digital manufacturing systems where the role of humans plays an important part. The digital manufacturing system is defined as:

'An integrated environment for design and development of products, production systems and business processes.'

The digital manufacturing system provides the necessary computer tools as well as digitally presented information and knowledge, in conjunction with human knowledge and skills. The digital information and knowledge exists only once in a formal and up-to-date form. It can be distributed, but is accessible to all related parties regardless of time and location.

The approach is based on the principles of Holonic Manufacturing Systems (HMS), Service-Oriented Architecture (SOA) and Fractal Manufacturing Systems (FrMS). The approach is part of a research project that aims to create a concept of a new, more adaptive, efficient and autonomous manufacturing system.

2 HOLONIC STRUCTURE OF DIGITAL MANUFACTURING SYSTEMS

HMS describes the manufacturing system by reference to holons, which are autonomous and co-operative units of a manufacturing system [7]. Autonomy means that the units can perform their tasks without the help of other units. The co-operation of the holons enables them to communicate with each other to achieve their shared objectives.

A fractal is an independently acting holon that can be precisely described [8]. The FrMS approach explains the self-similarity of the holons at different structuring levels. A manufacturing system holon is part of a manufacturing enterprise holon, i.e. it is a sub-holon in relation to the enterprise. At the same time, a manufacturing system holon consists of manufacturing unit holons; it is therefore a holarchy of manufacturing units.

SOA is a form of distributed systems architecture [9]. The communication in the manufacturing system and its context is based on SOA. It is seen as services, which hold all the information needed for autonomous operations and co-operation. This provides an environment for networked distributed operations between manufacturing holons. The communicating manufacturing holons can be units within a system or systems in an enterprise network communicating with each other.

2.1 Structure of Manufacturing Systems

The structure of the proposed digital manufacturing system consists of three types of manufacturing holons and corresponding manufacturing domains, each of them having a specific role in the system.

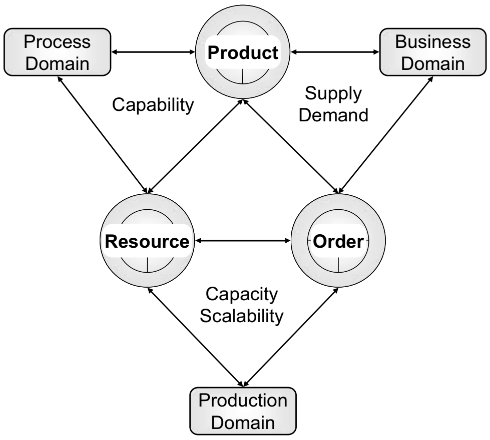


Figure 1: The structure of the proposed Digital Manufacturing System [10].

The structure of the proposed digital manufacturing system is presented in Figure 1. It is based on HMS reference architecture Product-Resource-Order-Staff Architecture, (PROSA). It explains the structure with three types of basic holons: product holons, resource holons, and order holons [11][12]. The holons are connected with the process domain, production domain, and business domain.

Products are what the customers see as supply; at the same time, they define the capabilities needed to make all of the products. *The process* domain describes the capabilities of the system. The system needs to be able to manufacture all of the features of the products, i.e. *the resources* should be associated with corresponding methods.

The resources, having the needed capabilities, also define the capacity of the system in *the production* domain. The production domain is responsible for manufacturing orders at the right time. It should have enough capacity to manufacture the volume and scalability needed to handle the variation in orders. *The business* domain makes sure that the system offers the right products, so that there will be enough *orders*, i.e. so supply matches demand.

2.2 Structure of Manufacturing Entities

Figure 2 represents the structure of a manufacturing entity, an autonomous and co-operative holon. The autonomy of the holon consists of digital, virtual and real parts, and the co-operation is seen as the communication part of the holon.

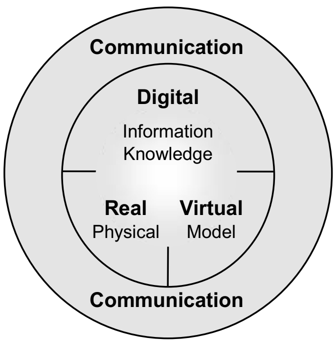


Figure 2: Digital, real, virtual and communication parts of a manufacturing holon.

The different roles and descriptions of the parts are briefly given as follows:

- *The digital part* includes all the digitally presented information and knowledge.
- *The real part* represents what exists physically in the real system.
- *The virtual part* is a representation of the physical part as a computer model.
- *The communications part* is the language and content of the information that is transferred within the system.

The digital part is common for the real and virtual parts in the current operating system, as the information and knowledge should exist only once. In future development cases, copy of the knowledge is used to avoid inaccurate information updated from the failed ideas of future development cases.

3 THE COMPLETION OF A GENUINE HOLON

The completion of a holon can be seen in terms of how easily it can become part of holarchy. A genuine holon has an information part, the digital part of a holon, which holds the information and knowledge of the holon as well as the optional physical part representing real life or a model of it.

- *A level 1* holon is a genuine '*plug and produce*' holon, which automatically connects as a part of the system. It is capable of communicating at both physical and information levels.
- *A level 2* holon can be connected to the system as a '*plug and play*' holon. It is capable of operating when the communication interfaces are implemented.
- *A level 3* holon does not have the ability to communicate with the system. It has autonomy, but the communication capabilities still need to be built. This can be achieved using, for example, sensors, machine vision, barcode readers etc.
- *A level 4* holon lacks both autonomy and communication capabilities. The autonomy should be implemented into the digital part of the holon as models, simulations and algorithms. This means that the autonomy of the holon is constructed in the digital manufacturing system.

By building the autonomy and communication capabilities into the digital part of the manufacturing system for the holons lacking the capabilities themselves, all of the holons can be used as a level 1 holon.

4 SERVICE-ORIENTED AND LEARNING HOLONS

The basic conceptual model of SOA architecture consists of service providers, service requestors and service brokers [13]. The holons in a digital manufacturing system based on SOA have the following roles:

- Service provider holons: typically the resource holons.
- Service requester holons: the order holons.
- Service broker holons, rules of the holons, i.e. the autonomy of the upper holon, holarchy.

Figure 3 presents the idea of a service provider holon. It has a certain role in the context in which it is operating. It has the role of a service provider. As it acts as a servicing holon, it is constantly learning from its operations.

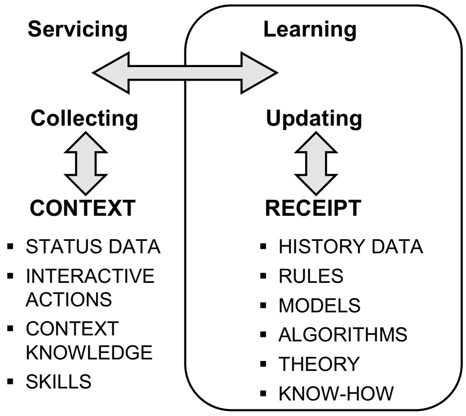


Figure 3: A servicing and learning resource holon.

Figure 3 is an example of a resource holon in a learning service-oriented environment. While the holon is servicing, it is aware of its context. It uses its skills and knowledge to operate, interact and collect information in the real manufacturing system. As it operates, it is constantly updating the digital manufacturing system. Information collected from the real manufacturing system is used to improve the digital part so it more precisely represents the real part.

5 FROM HISTORY ANALYSIS TO FUTURE VISIONS

One viewpoint from which to consider digital manufacturing systems is the time span in which the holons exist. It can be seen as, for example, history, current operations, tactical decisions or future vision.

In history analysis, the digital manufacturing system is used to analyze the real manufacturing system. As the system is operating, all the events occurring in the system are logged. The logged data can be examined and analyzed to find out what happened and why it happened. In finding the root causes for the phenomena, the system can learn from its past.

In real time operation (or in operations that are as real time as they can be), the digital and real manufacturing systems co-exist, constantly updating each other. The state of the real manufacturing system can be seen in the digital manufacturing system and actions can be made based on the state of the real system.

Tactical decisions consider the near future to which the manufacturing system is heading. In this case, the operating process occurs mostly in the digital manufacturing system. The future manufacturing system is developed digitally using the existing digital information and knowledge as well as possible new solutions. Possible benefits are, for example [14]:

- In the case of new products, the capability of the system to manufacture the products can be evaluated.
- Different solutions to make needed changes can be evaluated and the best solution selected.
- Criteria for selecting the best solution, based on requirements such as cost, quality, time.
- In the early steps, at a conceptual stage, the change requirements can be detected in advance.

The last of the benefits above is more in the area of future visions. They are similar to tactical decisions, the difference being the time horizon, where the outcome is more obscure but where more possibilities can be investigated.

6 AN EXAMPLE OF DIGITAL MANUFACTURING SYSTEMS LIFECYCLE

The lifecycle phases presented in Figure 4 include design, planning, implementation and operation. All of the phases rely on the existing information and knowledge.

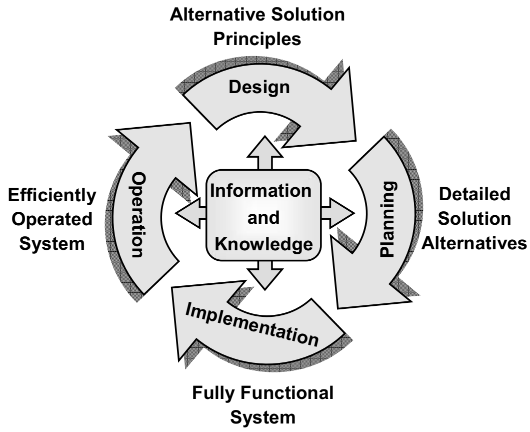


Figure 4: Proposed lifecycle phases of digital manufacturing systems [10].

At the end of the design phase, the goal is to have several alternative solution principles. The alternatives rely on the existing knowledge and whatever new knowledge should be acquired. The holons in the design phase evolve from unclear visions to more precise alternative solution principles. The holons, at this point, have only a digital part.

The planning phase aims to have detailed solution alternatives. The old system at the start of the design phase is changed to new system alternatives at the end of the planning phase.

The result of the implementation phase is a fully functional digital manufacturing system. The operation phase mimics the real manufacturing system and aims to efficiently operate the system. This phase verifies and validates previous phases and can result in iterations back to previous phases.

There can be several independent design and development cases in a real manufacturing system, each of them having a different focus. The congruency of them all is based on the fact that all of the cases use the same information and knowledge. When something new is learned, the information and knowledge is updated, and is accessible to the system to be used in future tasks.

7 ROLES OF DIGITAL MANUFACTURING SYSTEMS

Example viewpoints to explain the role of information in digital manufacturing systems are: capability descriptions, decision support system, open communications and education [15].

Capability descriptions are intended to describe the system capabilities. The descriptions are needed to be presented formally so that both humans and machines can use the information and communicate in the system.

Decisions made in a company can be seen from, for example, the respective viewpoints of the current operational system, short-term forecasting and long-term forecasting, where digital simulation models offer efficient tools for decision making.

An open communications system enables an efficient exchange of information and ideas between entities inside a company, in the enterprise network with suppliers and customers, as well as with experts in research centers and universities.

In education digital simulation models provide an efficient way to demonstrate events that are happening in the manufacturing system. Effects of individual decisions on the whole system are easier to understand with animation and realistic-looking graphics.

8 SUMMARY

One of the key enablers in global competition is a digitally presented manufacturing system. It enables a wider outlook towards the relevant information and knowledge, shortening the decision processes. Some of the characteristics of digital manufacturing systems commonly found in the literature are: an integrated approach to improve manufacturing engineering, the use of computer-aided tools and a framework for new technologies.

In this paper, an approach to digital manufacturing systems has been explained. It is based on the principles of HMS, FrMS and SOA. The system consists of basic building blocks: product holons, resource holons and order holons. The holons are connected with process, production and business domains.

The structure of individual holons consists of digital, real and virtual parts, as well as a communication part. The real part represents everything existing physically, while the virtual part is a representation of the physical part in a computer model. The digital part holds the information and knowledge of the holons for both real and virtual parts. The communication part enables the co-operation of different holons to achieve their mutual plans and goals.

Different viewpoints from which digital virtual holons can be viewed were discussed. The completion of a genuine holon offers a way of interpreting how holons in a manufacturing system can be truly holons, whilst the roles of the holons were examined as actors in a service-oriented environment. As the holon operates in the real world, it constantly learns from its actions and updates the digital information and knowledge to become more precise for future tasks.

The usage of the digital manufacturing system was discussed in terms of the time-horizon, the viewpoints of the lifecycle of digital manufacturing systems and examples of different roles of digital information and knowledge in digital manufacturing systems.

9 ACKNOWLEDGMENTS

This paper is part of a scientific research project FMS 2010. It is co-financed by Tekes, the Finnish Funding Agency for Technology and Innovation, and several major companies in Finland.

10 REFERENCES

[1] Souza, M.C.F., Sacco, M., Porto, A.J.V., 2006, Virtual manufacturing as a way for the factory of the future, Journal of Intelligent Manufacturing, 17:725–735.

[2] Reiter, W.F., 2003, Collaborative engineering in the digital enterprise, International Journal of Computer Integrated Manufacturing, 16/7–8:586–589.

[3] Kühn, W., 2006, Digital Factory - Integration of simulation enchanging the product and production process towards operative control and optimization, International Journal of Simulation, 7/7:27-39.

[4] Offodile, O.F., Abdel-Malek L.L., 2002, The virtual manufacturing paradigm: The impact of IT/IS outsourcing on manufacturing strategy, International Journal of Production Economics, 75:147–159.

[5] Bracht, U., Masurat, T., 2005, The Digital Factory between vision and reality, Computers in Industry, 56:325-333.

[6] Maropoulos, P.G., 2003, Digital enterprise technology-defining perspectives and research priorities, International Journal of Computer Integrated Manufacturing, 16/7-8:467-478.

[7] Valckenaers, P., Van Brussel, H., Bongaerts, L., Wyns, J., 1994, Results of the Holonic Control System Benchmark at the KULeuven, Proceedings of the CIMAT Conference, Troy, NY, USA, pp.128-133.

[8] Warnecke, H.J., 1993, The Fractal Company: A Revolution in Corporate Culture, Springer-Verlag, Berlin, Germany.

[9] W3C, 2004, Web Services Architecture, http://www.w3.org/TR/ws-arch/, accessed 27th January 2008.

[10] Nylund, H., Hokkanen, M., Salminen, K., Andersson, P., 2007, Lifecycle of Digital Manufacturing Supporting Design and Development of Adaptive Manufacturing Systems, Proceedings of The 12th Annual International Conference on Industrial Engineering - Theory, Applications & Practice, Cancun, Mexico, pp.544-549.

[11] Van Brussel, H., Wyns, J., Valckenaers, P., Bongaerts, L., Peeters, P., 1998, Reference architecture for holonic manufacturing systems: PROSA, Computers in Industry, 37:255-274.

[12] Wyns, J., 1999, Architecture for Holonic Manufacturing Systems - The Key to Support Evolution and Reconfiguration, PhD thesis, K.U.Leuven, Belgium.

[13] Gottschalk, K., 2000, Web Services architecture overview, IBM developerWorks, Whitepaper.

[14] Nylund, H., Andersson, P., Hokkanen, M., Salminen, K., 2007, Virtual Reconfiguration of Manufacturing Systems, Proceeding of the 2nd International Conference on Changeable, Agile, Reconfigurable and Virtual Production, Toronto, Canada, pp.801-808.

[15] Nylund, H., 2007, Role of Digital Information and Knowledge in Adaptive Manufacturing Systems, Annals of DAAAM for 2007 & Proceedings of the 18th International DAAAM Symposium, Vienna, Austria, pp.515-516.

Complex Manufacturing Space: An Integrated Description Model of Real and Virtual System Concepts for Information Enhanced Manufacturing

Masahiko Onosato, Fumiki Tanaka, Hiroaki Date and Ryoji Kawagishi
Division of Systems Science and Informatics,
Graduate School of Information Science and Technology, Hokkaido University, Japan

Abstract
Manufacturing in the digital age is deeply dependent on the information and communication technologies (ICT) and the manufacturing activities have been changing from direct operations by human operators to digital data exchanges by computers. It is fairly difficult for engineers and workers to understand the entire image of digital manufacturing activities. A concept of complex manufacturing space (CMS) introduces the duality of the physical world and the information world to manufacturing design, planning, control, and other activities. This paper explains the complex manufacturing space concept and some technical challenges for its implementation.

Keywords:
Virtual Manufacturing; Integrated Modeling; Complex Manufacturing Space

1 INTRODUCTION

Digitalization of objects and processes has been spreading quickly in various activities in manufacturing such as three dimensional product data preparation with CAD, computer-controlled machining and assembling with NC machine tools and robots, automated quality inspection and control with intelligent sensing systems, and so on. In advanced manufacturing systems, most functions are connected via communication networks, and now huge amount of manufacturing information is generated, distributed, stored and consumed in the information world.

Information and communication technology (ICT) has been changing manufacturing in terms of not only its productivity but also its fundamental style. In the conventional scheme of manufacturing, humans had faced real entities such as materials, products, tools and machines in design and production processes. Direct interactions between humans and real objects had formed a closed loop that connects the mental world to the physical world in manufacturing. Emergence of computers and advanced ICT has been separating these two worlds by inserting the virtual world between them. The virtual world by ICT, such as a virtual manufacturing system [1], offers interaction loops of necessary information services both to humans in the conceptual world and to machines in the real world.

In this new manufacturing scheme, humans might carry out entire process of product design and production with using only the virtual world, unless interacting with objects in the real world such as product mock-ups and machine tools. The rapid change of manufacturing style from direct interactions to indirect ones via the virtual world requests us to extend the notion of manufacturing models so as to understand the its real-virtual duality.

In this paper the concept of complex manufacturing space (CMS) is proposed as an integrated description model of real and virtual concepts for information-enhanced manufacturing, followed by implementation issues and prototyping of CMS.

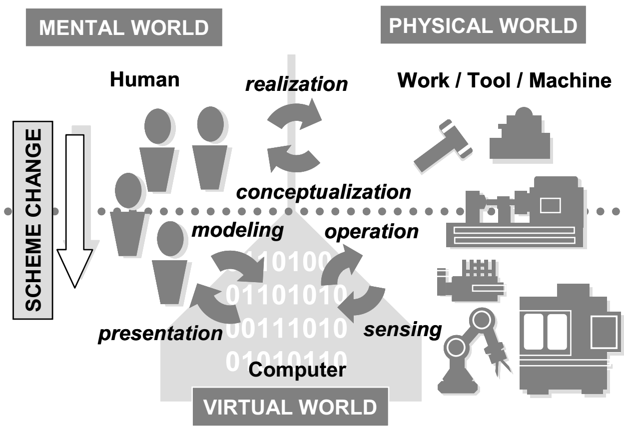


Figure 1 : Manufacturing scheme change by virtual world.

2 COMPLEX MANUFACTURING SPACE (CMS)

2.1 Digitalization in manufacturing systems

Every manufacturing system is understood as a function transforming described requirements to realized products. In this sense, information and substance have been always the two prime elements processed in manufacturing activities. Therefore production engineers have always made much effort to maintain consistent association between real elements and information about them. Old-style management methods for paper-based information have been replaced with computer-based ones for digital information step-by-step. For examples, drawings in design cabinets are replaced with

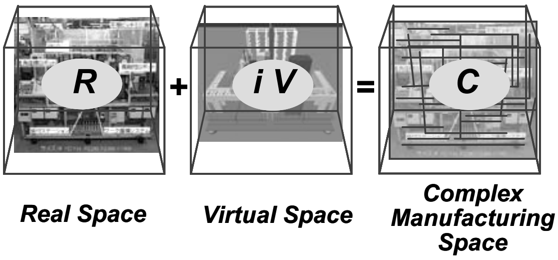


Figure 2 : Concept of complex manufacturing space.

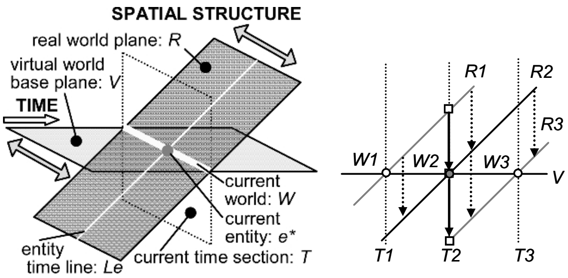


Figure 3 : Space-time structure of complex manufacturing space.

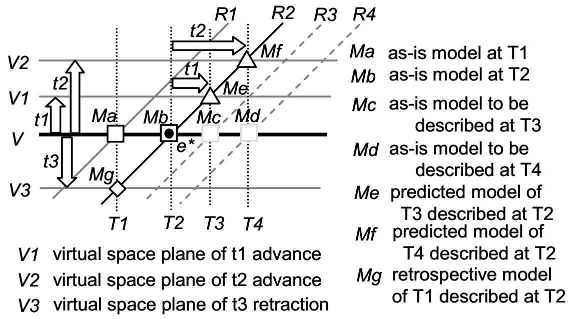


Figure 4 : Model desciptions on different virtual space planes in complex manufacturing space.

CAD models stored at file directories in computers, work instruction documents are changed to prompt instructions on an on-site display, and BOM descriptions of an assembly are now stored in IC data tags attached to a work pallet.

These efforts for digitalization of manufacturing functions have been still separated from others and made within the old-style scheme. Most digital data prepared in various manufacturing activities are used as just one-time information and not organized in the new manufacturing scheme with the virtual world. Of course you can use many effective ways developed for associating manufacturing information from difference functions, but knowledge accumulation by such problem-dependent solutions has made the virtual world more complicated and inconsistent, and it will be almost impossible for designers and engineers to recognize an entire image of information-enhanced manufacturing systems in a simple way

In the new manufacturing scheme shown in Figure 1, one should recognize a manufacturing system as a double-layered system with the virtual world and the physical world and this duality of these two worlds is the core concept in information-enhanced manufacturing.

2.2 Concept of complex manufacturing space

To get a comprehensive image of information-enhanced manufacturing, the authors propose a concept of complex manufacturing space. The term ‘complex’ of ‘complex manufacturing space’ has the same meaning as ‘complex number’ in mathematics. By introducing imaginary part ***i*** to real number system, we get a broader number system with another dimension such as $a + \boldsymbol{i}\, b$. This complex number system gives us more universal views and operations than the original real number system. In complex manufacturing space, the authors use the letter ***i*** for representing information space, not for imaginary part in the original notation.

Complex manufacturing space *C* is a composite space of real space *R* and virtual space *V* as shown in Figure 2. Every element in CMS, $c \in C$ has a real space part $r \in R$ and a virtual space part $v \in V$ and it can be denoted as $c = r + \boldsymbol{i}\, v$. Of course, it might be just an analogy of the mathematical notation and we cannot introduce the same operations as the complex number system to our CMS since *R* and *V* are not so simple as number systems.

The main purpose of introducing the concept of CMS to information-enhanced manufacturing is to declare that every entity in a manufacturing system has its dual structure bridging both real and virtual spaces. In the next section, CMS is explained from the viewpoint of an integrated description model.

3 CMS AS AN INTEGRATED DESCRIPTION MODEL

3.1 Space-time structure of CMS

As mentioned in the last section, CMS consists of two sub spaces: real space and virtual space. Let us discuss about the spatial structure of CMS based on Figure 3.

First we introduce two prime planes: a real world plane *R* and a virtual world base plane *V*. Although we use the term “plane” for easy explanation, both planes are basically four-dimensional space with three dimensions for spatial structures and one for time axis. The real world plane *V* represents the past, present and future states of the real world, and it includes entity time lines *Le* for each entity’s life in the real world. The virtual world base plane *V* is a description space for ‘as-is’ models. These two planes intersect at a line *W* which means the current world in three-dimensional space. The current world contains a set of current entities *e** such as machines, works, tools and so on. The important point of the real world plane is that you can observe and operate entities in the real world only at this current world *W* and you cannot access to other place *R* – *W* on the real world plane. In other words, what you can see and touch are only entities at present time, not ones in past or in future.

As time advances from *T1* to *T3* in Figure 3, the real world plane is continuously shifting downward from *R1* to *R3* and the intersecting line of the current world is moving rightward on the virtual world base plane from *W1* to *W3*.

3.2 Model descriptions in CMS

Complex manufacturing space is intended to offer an integrated description model in real-virtual mixed environment of information-enhanced manufacturing. Every model in CMS has its temporal aspect determined by its position. Figure 4 shows us a side view of CMS. At time *T1*, the virtual space base plane *V* intersects with real world space plane *R1*, and a model *Ma* is defined by using sensing data from the real world. This Ma model is an 'as-is' model at *T1*. Then, at time *T2*, we get a new 'as-is' model *Mb* from current entity *e** and *Ma* becomes an old 'as-is' model of the entity. At future time *T3* and *T4*, you will get new 'as-is' models *Mc* and *Md*, but you never get these models at the time point of *T2*. An important point is that every 'as-is' model is defined on the virtual space base plane and we cannot get any 'as-is' model of future because no future entity exists at present.

In addition to the virtual space base plane, you can define new virtual space planes in CMS. In Figure 4, new virtual space planes *V1* and *V2* are located above *V* in CMS with distance *t1* and *t2* from *V* respectively. These virtual space planes with offsets from *V* mean future model description spaces. The other virtual space plane *V3* placed under *V* means past model description space. For example, a predicted model *Me* on *V1* at *T3* (= *T2* + *t1*) is defined at *T2*, and a retrospective model *Mg* on *V3* at T1 (= T2 – t3) is defined at *T2*. These models are not 'as-is' models. When the time is just at *T3*, you can get a new 'as-is' model *Mc* and compare it with the predicted model *Me*. In some cases, a predicted model *Me* may be defined as a 'to-be' model which requests the real entity to be *Me* at *T3*.

4 INTERACTION BETWEEN HUMANS AND CMS

4.1 Real-virtual duality in CMS

In the new scheme for information-enhanced manufacturing shown in Figure 1, designers, engineers and workers should often face the abstract virtual world instead of the intuitive physical world in order to produce and control real entities in the physical world indirectly. The construction of human-friendly interface to the virtual world is the most important issue in the new manufacturing scheme [3].

The real-virtual duality of entities in CMS is a key for intuitive management of various descriptions about manufacturing objects. This duality requires the bidirectional accessibility as follows.

From virtual space to real space

Each model of an entity *e* described in CMS is related to the entity's time line *Le* on the real world plane and it is expected to keep a reference to a current entity *e** in the current world *W*. For example, when you choice some entity model in CMS, the entity referred by the model can be appropriately pointed in the real world.

From real space to virtual space

Each current entity in the real world works as an entry point to model descriptions in CMS. When you point some entity in the real world, you can get an access point to retrieve related models stored in CMS.

In order to realize the real-virtual duality above mentioned, the structural equivalence between the real world and the virtual world should be appropriately maintained by constructing and revising "as-is" models of entities contained in the real world.

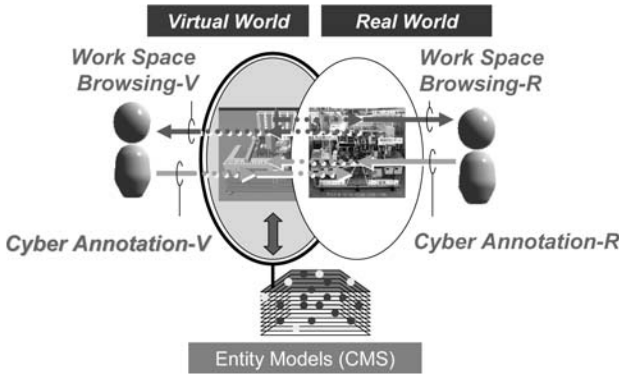


Figure 5 : Four types of user interactive services in CMS.

4.2 Real-virtual interaction services in CMS

When we can develop an information system which realizes the CMS structure and functions, the system will provide next four types of interactive services to human users as shown in Figure 5.

Work Space Browsing-V

This interactive service enables users to browse the current state of the real entities in anyplace via GUI for the virtual world models. This browsing shrinks spatial distances in the real world.

Work Space Browsing-R

With this browsing service, users can get the past-present-future information from CMS about an entity pointed in real working environment, and retrieved information is presented to the users by using a mobile information terminal or augmented reality device. This service gives users an easy entry point to CMS unless typing database commands.

Cyber Annotation-V

Cyber annotation-V is an indirect communication with a real entity which exits at present, or sometimes will exist in future, via the virtual world model. For example, if you want to send a message for some real entity, you can put it to the virtual description model in CMS.

Cyber Annotation-R

Cyber annotation-R is the reverse directional service of work space browsing-R. With this service, you can add information related to a pointed entity in real working environment and the information is properly assigned to description models in CMS. As well as work space browsing-R, each real entity plays a role of an entry point to CMS.

Although most function of each service described here is partially developed by other research groups [4][5] or already provided in business [6], an integrated system of these user interface services in a consistent way is not realized at present. The authors have been developing an experimental system of CMS with which can offer all these four services in

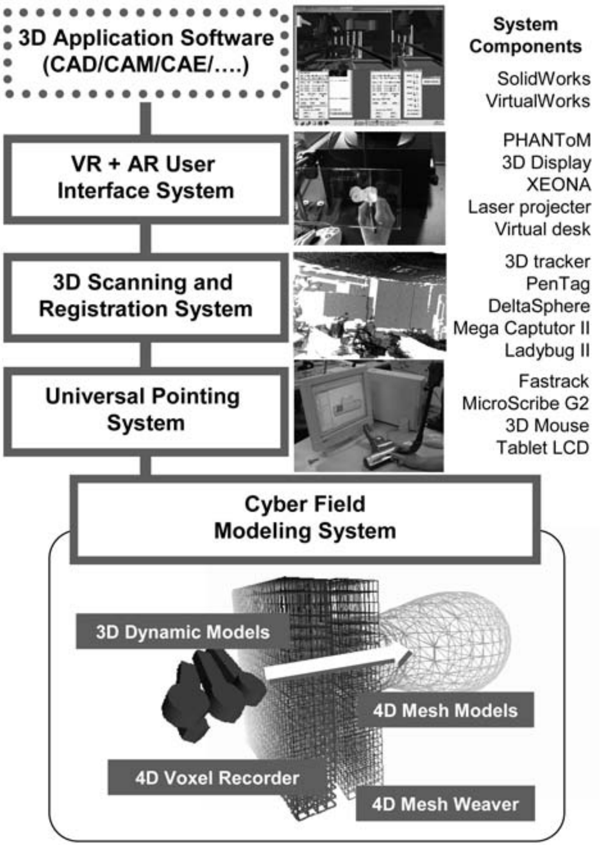


Figure 6 : Configuration of CMS experimental system.

an integrated way. In the next section our research approach to CMS and the configuration of our experimental system are summarized.

5 EXPERIMENTAL SYSTEM OF CMS

For implementing CMS, the authors designed an experimental system configuration of CMS shown in Figure 6. The system consists of four prime subsystems as follows;

Cyber Field Modeling System

This is the core part of CMS implementation and records 3D dynamic models as 4D continuous mesh models by way of the 4D voxel recorder and the 4D mesh weaver. This system enables space-time shape modeling in 4D virtual space, and it can extract 3D geometric models at requested time points.

Universal Pointing System (UPS)

This system offers universal style of object pointing both in the real world and the virtual world. For example, you can use various pointing devices in different application software systems and specify either a real object or virtual models on screens or other display devices.

3D Scanning and Registration System

Working environment in information-enhanced manufacturing, such as a intelligent design room or a computer-supported assembly cell, is spatially digitized by using 3D scanners. Measured point sets are analyzed for virtual world registration of real entities expected in the working environment.

VR+AR User Interface System

Virtual models in CMS are demonstrated to users with high reality with VR interface XEONA, which presents glassless stereo images and a haptic display [7]. AR interface such as information projection onto real entities is also supported for work space browsing-R.

In addition to these subsystems, 3D application software will be connected to the system for getting CMS services. For example, virtual manufacturing system and CAD system are expected clients of CMS services.

The prime four subsystems have developed as a prototype system and the authors is working on the integration of these prototypes for realizing CMS services.

6 CONCLUSIONS

The concept and application cases of complex manufacturing space were introduced for the real-virtual integrated description models in information-enhanced manufacturing. Our research laboratory at Hokkaido University advances the CMS research project both from its theoretical framework and practical system implementation.

ACKNOWLEDGMENTS

This research was partially supported by the Japan Society for the Promotion of Science (JSPS), Grant-in-Aid for Scientific Research (B) No. 18360063, 2006-2007.

REFERENCES

[1] Onosato, M., Iwata, K., 1993, Development of a Virtual Manufacturing System by Integrating Product Models and Factory Models, Annals of the CIRP, 42/1, 475-478.

[2] Onosato, M., 2005, Manufacturing Knowledge in the Digital Age: Recording Expert's Activities of Problem Solving in Digital Working Environment for Knowledge Intensive Manufacturing, CIRP Journal of Manufacturing Systems, 34/5, 427-436.

[3] Kapp, P, Weimer, T, Westkämper, 2006, Factory Cokpit : A Factory Simulator for the Entire Factory Life Cycle, CIRP Journal of Manufacturing Systems, 35/6, 543-550..

[4] Ishii, H., Ullmer, B., 1997, Tangible Bits: Towards Seamless Interfaces between People, Bits and Atoms, Proceedings of CHI'97, 234-241.

[5] Rekimoto, J., Ullmar, B., Oba, H., 2001, DataTiles: A Modular Platform for Mixed Physical and Graphical Interactions, Proceedings of CHI2001, 269-276.

[6] Google, Inc., Google Labs Ride Finder - Taxi and Shuttle location, URL: http://labs.google.com/ridefinder.

[7] Onosato, M., 2006, Digital Archive of Manufacturing Knowledge based on Virtual Manufacturing, Proc. of the 1st Virtual Manufacturing Workshop (VirMan'06), 31-35.

Grid Engineering for Networked and Multi-scale Manufacturing

Carmen Constantinescu[1, 2] and Engelbert Westkämper[1, 2]

[1] Institut für Industrielle Fertigung und Fabrikbetrieb, IFF - Universität Stuttgart, Germany

[2] Institut für Produktionstechnik und Automatisierung, IPA - Fraunhofer Gesellschaft, Germany

Abstract

Networked and multi-scale manufacturing systems are expected to have much more flexibility to respond to dramatic changes in the world market. A real responsiveness might come from dynamic and unlimited resource accessibility rather than from rigid factory structures and boundaries. New challenges arising from these requirements, as Reference Model of Factory Planning, Reference Factory Data Model and continuously and multi-scale integration and synchronization of Engineering and Manufacturing Worlds guide our research on Grid Engineering for Manufacturing. A concept with a holistic approach and a software infrastructure framework are proposed as our solution for the rapid prototyping of factories based on integrating grid technology with digital manufacturing technologies.

Keywords:

Production systems, modelling, simulation, grid technology

1 PROBLEM STATEMENT

The manufacturing enterprises, called factories, are confronted today with new models of competition and new modes of operation. They have to provide competitive industrial goods and support services at decreased prices, high quality, by overcoming the customer expectations. In order to remain competitive or to survive, the factories have to be adapted permanently to the needs and requirements of markets and economic efficiency. In achieving these objectives new and innovative methods, technologies and tools have to be employed in planning and permanently optimisation of the factory operation and its corresponding manufacturing processes [1]. This leads on one side to the requirement of development and mastering highly efficient modelling, simulation and optimisation algorithms and tools, which support all scales of manufacturing engineering, from network of manufacturing systems to technical processes and along all Factory Life Cycle phases, from investment planning to the dismissal or refurbishment [2, 3]. On the other perspective, the application of Life Cycle thinking to modern manufacturing engineering practice aims to manage the total and comprehensive Life Cycle of the factory and its products towards more sustainable consumption and production leaded to another new industrial paradigm, the “Sustainable Life Cycles Management” [4]. Its implementation is supported by integration and synchronisation of Information and Communication Technologies (ICT), the digital manufacturing technologies, collaboration models and tools, which are used to trace factories, products, processes and technologies during their Life Cycle from engineering to the end of their lives. The Sustainable Life Cycles Management approach is founded on several successful research results and current projects developed at the Universität Stuttgart, Institute of Industrial Manufacturing and Management (IFF) in cooperation with Fraunhofer – Institute of Production Automation and Management (IPA) [4, 5, 6, 7, 8]. Our current research focuses on the Factory Life Cycle dimension by enhancing the first prototype of a Virtual Environment for Factory Planning which supports the facility layout modelling, planning and optimisation activities by facilitating the collaboration, immersion and presence in VR-based manufacturing environments. The paper presents the new challenges arriving from the current research and the possible solution, the Grid Engineering for Manufacturing. Several theoretical considerations on Grid Computing and the first concept of the Grid-based collaborative environment for factory and process planning are afterword shortly introduced. The conclusion and future work are closing the paper contribution.

2 NEW CHALLLENGES IN MANUFACTURING ENGINEERING

The networked and multi-scale manufacturing engineering and its stated problem have to face new challenges arriving from the migration of the “Rapid prototyping” concepts, methodologies and tools to the factory as a whole, approached under the new paradigm as a special type of product. The “rapid design and virtual prototyping of factories” emerged as a new challenge for the manufacturing community aiming at dramatically reducing the time of setting-up and ramp-up the factory to half of the current required time. These challenges are highlighted in the following, as a base and motivation for the possible solution envisioned and shortly presented in this paper.

2.1 Motivation: factory functional architecture

The main phases of the Factory Life Cycle are investment planning, engineering, process planning, construction & ramp-up, production, service and maintenance, and finally, dismantling or refurbishment [4]. The origin of these phases resides in fulfilling the core factory functions, as presented in Figure 1. For satisfying the customer orders and deliver the required products and services, several planning functionalities, e.g. planning of investment, site and building, infrastructure, media, logistics, processes, machines, equipments and layout, have to be performed in order the factory to be ramped-up, monitored and maintained. Critical activities as those ranged between investment planning and ramp-up are grouped under the so-called factory and processes planning. Models, methods, technologies and corresponding tools supporting all these activities are creating, storing, accessing and updating relevant data managed local in several data bases, as shown in Figure 1. The data related to the factory is complemented by the product data, stored and managed in specific and local data bases, as well. These data bases are divided according the two worlds: product – engineering world and factory – manufacturing world. Three challenges are arising from the above functional architecture of the factory, which are close related each other.

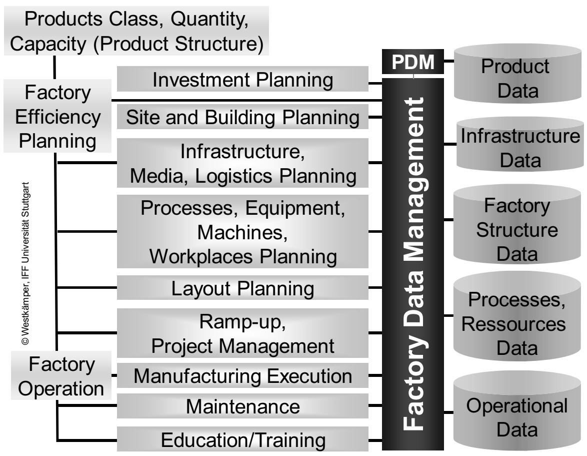


Figure 1: Functional architecture of a manufacturing enterprise (PDM: Product Data Management).

2.2 Reference Model for Factory and Process Planning

The first one represents the development of a “Reference Model of Factory and Process Planning”, aiming at identifying all activities and steps required as standard and mandatory for the purposes of the factory planning. These represent the base of conceiving the business model of factory planning. This Reference Model will represent the main instrument for the planners and interdisciplinary teams involved in factory planning activities. The concepts, purposes and the tasks of factory planning have been approached in the research works of Kettner [9] and Aggteleky [10]. According these fundamentals, the factory planning represents a top-down or a bottom-up process addressing the following:

- Site planning: identifies the location by analyzing the characteristics of the economic and social environment;
- Space planning: disposition of terrains and buildings;
- Infrastructure planning: assures the continuous supply of different required medium, e.g. air, energy, water;
- Logistics planning: design and configuration of information and material flow;
- Production facilities planning: machines and equipment planning, strongly related to the logistics planning;
- Additional and support activities planning: other support role activities, e.g. equipment maintenance;
- Human resource and organization planning: planning of the organizational structure, position requirements, recruitment and continuing education.

These main activities and steps have to be enhanced, refined, classified and strictly formalized in our so-called “*Reference Model of Factory and Process Planning*”.

2.3 Reference Factory Data Model

As a consequence deriving from the first challenge, a “*Reference Factory Data Model*” has to be conceived and developed. This model has to collect all data needed in performing the factory and process planning activities and to define all relationships between this data, as well. The functional architecture is so completed with the corresponding Factory Data Management, additionally to the Product Data Management which is coming from the product engineering world. There is no available at the moment such a reference or generic model required as a standard by the scientific manufacturing community. This development constitutes one of the main topics of our current research works.

2.4 Continuously and multi-scale integration and synchronization of Engineering and Manufacturing Worlds

The third challenge represents the innovative integration of the two worlds and of all corresponding data managements and tools, in order to give reality to the synchronization between the Product and Factory Life Cycle and implementation of the Sustainable Manufacturing Management. Not only the typical problems of Enterprise Application Integration (EAI) arise here, furthermore several critical aspects related to the rapid prototyping of factories have to be mentioned.

The particular challenge of modelling and simulation of networked and multi-scale manufacturing systems consists not only in determining the interactions of the different parameters and influencing factors but also in representing them in real-time simulation models. In the context of the different usages, the term "real-time” is understood here as a fast reaction to arising events as well as the time-deterministic calculation of factory behaviour for the control-coupled simulation.

The multi-scale simulation of manufacturing systems aims to connect different computation and simulation data models approaches through a continuous information flow, higher order equation systems and universal interfaces. The goal of this connection is the bridging of the discrete and numerical simulation, and of the areas of application in the manufacturing domain, as well.

In our approach, under the notion “multi-scale” not only the spatial and temporal scales within separate manufacturing processes is understood, but also the different scales of all running processes in the whole factory, as well as the different scales in the model itself. Therefore, for the purposes of the virtual representation of the factory as a whole [3], at different levels of abstraction called *scales* and along all planning phases in a *continuously/comprehensive* way, several *heterogeneous modelling approaches* have to be coupled with each other.

In order to approach comprehensive the modelling and simulation at all scales, the following challenge has to be overcome. If one regards the today's conditions in the simulation of technical processes and factory organizational operational sequence, then it can be stated that this can be based on completely different, decoupled computation and data models. There is no constant information flow between these scales. Therefore, the main goal is to solve the integration between the heterogeneous simulation models. This heterogeneity consists on the migration of simulation models from *numeric simulation* for the process and material modelling towards the *discrete simulation* for the purposes of logistics simulation. The complexity arises when approaching the modelling of the horizontal scales of the factory as a whole, beginning with the technical manufacturing processes, through equipments, robots, production systems, segments and network of production systems. As a conclusion, the following aspects and scales have to be regarded:

- all levels of factory structures: manufacturing processes, machines, manufacturing system and network;
- several areas and concepts related to: mathematics, physics, chemistry, engineering, economic science;
- several simulation methods: molecular dynamics, finite elements and event-oriented simulation;
- spatial expansion: atomically, microscopically, mesoscopically or macroscopically;
- temporal expansion: nanoseconds, seconds, hours, days, weeks or years.

All here regarded processes are based on different numeric computation and run off on different time scales. This to interconnect and thus a constant "elbow" of the technical processes up to the total expirations in the production system stretch, is the central setting of tasks for the next years. *Different temporal scaling* of the simulation beginning on the different levels of the factory structure can be exemplified. While the physical procedures during the technical processes take place usually within a range from milliseconds up to few minutes, the factory organizational operational sequence in the entire production system extend up to several weeks and years. *The heterogeneity of different numeric solutions* like the molecular dynamic simulation (MD), to Monte Carlo (MC), finite element method (FEM), knowledge-based models as well as the kinetics and logistic simulation [2] for the different levels of the factory structure and/or the operational areas indicated represent a not at the moment solved challenge.

3 POSSIBLE SOLUTION: GRID ENGINEERING FOR MANUFACTURING

3.1 New approach of Grid engineering for manufacturing

Motivated by the three main challenges shortly presented above, our research focuses on searching a possible solution to support the complexity of continuously, multi-scale modelling and simulation of factory for the purposes of its rapid prototyping. The envisioned approach represents the employment of Grid technology and engineering in manufacturing under the so-called Grid Manufacturing. Several considerations and advantages of using Grid, as premises of our work are in the following presented.

3.2 Grid computing and its value proposition

The last few years proved the evolving of the grid computing from a niche technology associated with scientific and technical computing, into a business-innovative technology which is driving increased commercial adoption. Grid deployments accelerate application performance, improve productivity and collaboration, and optimise the resiliency of the IT infrastructure by the ability to store, share and analyse large volumes of data. It ensures that people and applications have access to information at right time which can improve decision making, employee productivity and collaboration. The main advantage represents the virtualisation, of both information and workload. In non-grid infrastructures, the information is “siloed”, resources are dedicated to applications and information. Several grid focus areas of high relevance in the last years have to be highlighted: a) business analytics grid, b) engineering and design grid, c) research and development grid, d) government development grid and e) enterprise optimisation grid. The second and the last areas are of high interest for the manufacturing engineering.

3.3 Grid Manufacturing: integrated and continuously management of the entire Product and Factory Life Cycle

Grid Manufacturing represents in our work a new approach aiming at enabling the collaborative sharing of resources and competences in manufacturing engineering respectively supports the integrated and continuously management of the entire Product and Factory Life Cycle. It facilitates the use of networked and distributed manufacturing resources in order to a) support the collaborative planning, operation and management of manufacturing in close relation with product design and development and to b) respond to the emerging challenges: innovation, speed and flexibility. This represents a possible solution for the third challenge presented in the Section 2, the continuously and multi-scale integration and synchronization of Engineering and Manufacturing Worlds, the orchestration of the Product and Factory Life Cycle.

Based on our former results and experience in developing collaborative environments for manufacturing [4] we began a new development, the Grid Engineering for Manufacturing Laboratory – GEMLab. It will implement all theoretical aspects presented in this paper, with the support of several strategic ICT key players offering valuable solutions for virtual products and digital manufacturing and the expertise in Grid Computing coming from the High-Performance Computing Centre of the Universität Stuttgart.

The main concept of the architecture, structured in three tiers, with additional extension for other existing independent proprietary or commercial application systems is presented in Figure 2.

The first layer, named *Grid Engineering Infrastructure*, networks by employing Grid computing the distributed product design and manufacturing planning resources and activities that are interconnected for fulfilling the Factory Planning Life Cycle phases. These are represented mainly by the specific applications for product design and factory planning, e.g. ProEngineer (PTC), Tecnomatics components (Siemens PLM Software) and RFID applications. The Grid services comprise here all activities which have to be fulfilled e.g. product design, product assessment, factory layout and manufacturing execution, which access and use data from their corresponding data bases, as shown in Figure 2. The RFID information is used e.g. for synchronizing the digital factory with context and situation-aware data of production resources and tools coming from the real production on the shop-floor.

The *Models Infrastructure* is the layer which integrates the Factory Data Model, the Product Data Model and the reusable simulation models already developed for supporting the product design and factory planning activities, e.g. the material flow simulation, the acoustic simulation, work-place and collision detection simulation, etc. These models are integrated with the corresponding Product and Factory Data Models in order to access and use the required data.

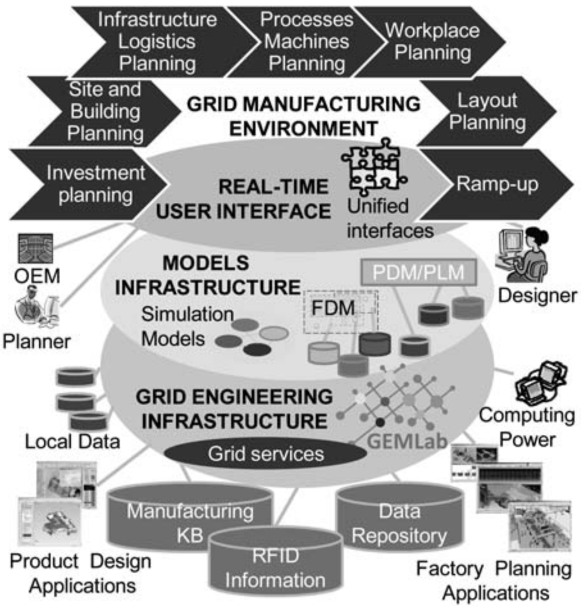


Figure 2: Concept for Grid-based collaborative environment and platform for factory and process planning (FDM: Factory Data Management, PDM: Product Data Management, PLM: Product Life Cycle).

The *On-line Real-time User Interface* layer enables the user to access the design and planning environment through the web portal, to accept request on the services and to display results, respectively the online monitoring.

The conceptual work and the prototyping of the environment is still undergoing, facing the challenge of selecting the suitable Grid technology for realising the Grid Engineering Infrastructure Layer.

The expected benefits coming from implementing the presented approach are as follows: a) continuously design of products and planning of factory by orchestrating the simulation activities in both worlds, product and factory, b) feedback from the later stages can be communicated back to design and planning stages in order to reduce the future implementing errors and the required time for product launching and factory ramp-up.

4 CONCLUSIONS AND FUTURE WORK

The network of multi-scale manufacturing systems are facing new challenges related to the rapid prototyping of the factory in order to reduce the time-to-market of required products and the shortening of the ramp-up of the factory to a half of usual duration. New areas of research and development have been identified, as arising from the stated challenges. These represent core requirements as the development of a Reference Model for Factory Planning which has to support the interdisciplinary planner teams to optimise their planning activities. As a consequence, the conceiving and the development of a Reference Factory Data Model which collect, structure and interlink all data needed in planning of factories and processes is stringently required. The continuously and multi-scale integration and synchronization of Product - Engineering and Factory - Manufacturing Worlds represent another challenge which is possible to be overcome by the shortly presented solution, Grid Engineering for Manufacturing. The first thinking on implementing this approach is based on our current projects in the field of Digital Factory and on the development of the laboratory for Grid Manufacturing, GEMLab. The next steps of our work concern the technical employment of the Grid Computing in our lab and actual collaborative platform, with the support of knowledge and expertise of the High-Performance Computing Centre of the Universität Stuttgart.

5 REFERENCES

[1] Westkämper, E., Winkler, R., 2002, The Use of System Modelling for the Intelligent Planning, Scheduling and Control of Agile Manufacturing. In: The 35th CIRP-ISMS „Manufacturing technology in the information age", 12-15 May, Seoul, South Korea, p. 644 – 655.

[2] Scheibe, A., Constantinescu, C., Decker, M., Ye, Q., Westkämper, E., 2006, A comprehensive and holistic approach of multi-scale modelling and simulation of manufacturing systems. In: The 6th International Workshop on Emergent Synthesis, IWES06, 18-19 August, Kashiva, Japan, p. 205-211.

[3] Westkämper, E., Constantinescu, C., Hummel, V., 2006, New paradigms in Manufacturing Engineering: Factory Life Cycle. In: Annals of the Academic Society for Production Engineering. Research and Development, XIII/1, Volume XIII, Issue 1, p.143-147.

[4] Constantinescu, C., Dürr, M., Decker, F., Westkämper E., 2007, Virtual Environment for Collaborative Factory Planning, In: The 40th CIRP International Seminar in Manufacturing Systems, 30 May-1 June, Liverpool, United Kingdom.

[5] Constantinescu, C., Heinkel, U., Le Blond, J., Schreiber, S., Mitschang, B., Westkämper, E., 2005, Flexible Integration of Layout Planning and Adaptive Assembly Systems In: The 38th International Seminar on Manufacturing Systems, 16-18 May, Florianopolis, Brazil, p.10-18.

[6] Constantinescu, C., Hummel, V., Westkämper E., 2006, Fabrik Life Cycle Management. Collaborative standardisierte Umgebung für die Fabrikplanung (KOSIFA). In: wt Werkstattstechnik online, 96/4, Springer-VDI-Verlag, Düsseldorf, Germany, p.178-182.

[8] Rantzau, R., Constantinescu, C., Heinkel, U., Meinecke, H., 2002, Champagne: Data Change Propagation for Heterogeneous Information Systems, in Proceedings of the 28th International Conference on Very Large Data Bases, 20-23 August, Hong Kong, China, p.1099-1102.

[7] Mitschang B., Westkämper E., Constantinescu C., Heinkel U., Löffler B., Rantzau R., Winkler. R., 2003, Divide et Impera: A Flexible Integration of Layout Planning and Logistics Simulation through Data Change Propagation. In: The 36th CIRP-International Seminar on Manufacturing Systems, 03-05 June, Saarbruecken, Germany, p.411-418.

[9] Kettner, H., Schmidt, J., Greim, H.-R., 1984, Leitfaden der systematischen Fabrikplanung. Hanser, München, Germany.

[10] Aggteleky, B., Fabrikplanung, 1987, Werksentwicklung und Betriebsrationalisierung, Hanser, München, Germany.

Smart Factory - A Step towards the Next Generation of Manufacturing

Dominik Lucke[1], Carmen Constantinescu[1, 2], Engelbert Westkämper[1, 2]
[1] Institut für Industrielle Fertigung und Fabrikbetrieb, IFF - Universität Stuttgart, Germany
[2] Fraunhofer-Institut für Produktionstechnik und Automatisierung, IPA - Stuttgart, Germany

Abstract
The Stuttgart Model of adaptive, transformable and virtual factories, already implemented in German basic research performed at the Universität Stuttgart has been extended with a new perspective, the so-called "Smart Factory". The Smart Factory approach is a new dimension of multi-scale manufacturing by using the state-of-the-art ubiquitous/pervasive computing technologies and tools. The Smart Factory represents a context-sensitive manufacturing environment that can handle turbulences in real-time production using decentralized information and communication structures for an optimum management of production processes. This paper presents our research steps and future work in giving reality to the envisioned Smart Factory at the Universität Stuttgart.

Keywords:
Smart Factory; Real-time Factory; Ubiquitous computing

1 PROBLEM STATEMENT

In recent years manufacturing engineering experienced a dramatic change through different parallel running developments. The globalisation and the wish to produce highly customized products lead to a higher proliferation of variants, shorter product life cycles and closer enterprise networks. The short planning horizons and product life cycles induce the decrease of batch sizes and do therefore require a high dimension of manufacturing flexibility. In order to make the right management decisions, real-time information and the direct realisation of the decisions are indispensable.

Dynamical changes in the factory, caused through internal or external turbulences like a machine breakdown or an order fluctuation in the market, can often not be taken into account and therefore lead to false decisions on different planning levels. Due to complex interactions of the different functions and departments of the factory and their task-oriented specific data formats, the causes of the dynamic changes exponentiate themselves and their consequences for the factory as well. Thus, high flexibility demands are posed to the manufacturing resources, their planning and control [1].

The management control of the complex processes inside and outside of the factory is even today performed through Enterprise Resource Planning (ERP) and Manufacturing Execution Systems (MES) and applications. The increased market turbulences and therefore increased flexibility of manufacturing require complex manufacturing sequences, which are difficult to realise with the today's solutions. Outdated information in the different information systems are leading directly to problems in planning and production.

The information management is responsible for the allocation of the job and process specific information like NC programs or machine properties, however the material management is responsible for the supply of components. Tool and device systems are additional mobile manufacturing resources to conduct a manufacturing process. The coordination of many heterogeneous subsystems providing all required resources, materials and information at the work place is necessary to ensure a constant resource load. To achieve this synchronization, many different specialized software systems and applications are used, like resource management systems, MES or ERP systems. Any failure in a subsystem would result in a significant reduction of the productivity of the whole system. While a MES plans and controls the manufacturing level, the ERP level plans and controls the synchronization of the subsidiary planning entities. Even the smallest difference between the real and digital saved data, e.g. initiated by a malfunction, leads to planning discrepancies and a miscalculation of the optimal working point.

A parallel development was enabled by the integration of electronic components like microchips or sensors into objects due to their decline of price. This development enables a decentralized control in a more economical way.

2 SMART FACTORY: DEFINITION, CHALLENGES AND ENABLING TECHNOLOGIES

The design and development of the Smart Factory require as a first step the definition of the concept. Mark Weiser's vision of the smart environment describes a physical world, which is closely and invisibly interwoven with sensors, actuators, displays and computer elements, which are seamlessly embedded into daily life objects. They are connected with each other by a network [2]. The Mark Weiser's approach of smart environments is transferred to manufacturing issues. After the development of digital and virtual factories next step in evolution of factories is the fusion of physical and virtual world [3] under a so-called Smart Factory.

2.1 Definition

The basic research in the field of Smart Factory at Institute of Industrial Manufacturing and Management (IFF) is performed within the Center of Excellence Nexus (SFB 627) [4]. This interdisciplinary research is funded by the German Research

Foundation (DFG). The so-called *Smart Factory* is defined as a Factory that context-aware assists people and machines in execution of their tasks. This is achieved by systems working in background, so-called *Calm-systems* and *context-aware applications*. In our case, context-aware means that the systems can take into consideration context information like the position and status of an object. These systems accomplish their tasks based on information coming from physical and virtual world. Information of the physical world is e.g. position or condition of a tool, in contrast to information of the virtual world like electronic documents, drawings and simulation models. These systems are working on different levels of the factory, like context-aware information systems in the shop floor (workers cockpit) or advanced manufacturing execution systems that can act context-aware for the shop floor manager. Calm systems are referring in this context the hardware of a Smart Factory. The main difference between *calm* and other types of systems is the ability to communicate and interact with its environment.

2.2 Challenges

The Smart Factory concept enables the real-time collection, distribution and access of manufacturing relevant information anytime and anywhere. The Smart Factory represents a real-time, context-sensitive manufacturing environment that can handle turbulences in production using decentralized information and communication structures for an optimum management of production processes. Premises for further assistance than today are the horizontal and vertical integration of information systems, the assignment of material and flow of information within an enterprise. For acting context-aware, the applications in the Smart factory have to answer the following three questions, from those deriving more challenges:

1. How is an object identified? ➔Identification phase
2. Where is an object located into the factory? ➔Positioning phase
3. What is the situation or status of an object? ➔Status knowledge

These and further challenges are shortly presented:

Identification:

The Identification of objects, as one of the basic challenges in a factory, assigns information of the virtual world like process steps to real world objects. Therefore suitable identification methods, tags, sensors, sensor readers and communication facilities have to be found and chosen, specific to their task in a rough industrial environment.

Localization:

For improving the processes and reducing idle times within the Smart Factory the localization is required to have an actual knowledge about the position of the objects like tools or materials. Depending on the purpose, the accuracy of a positioning system has to be a range within 0,15m - 1m. Furthermore a positioning system used in a manufacturing environment has to work on a large scale and has to be robust against environmental influences, electromagnetic fields, noise of dust, etc. [3].

Status knowledge:

The assistance systems in a Smart Factory have to know the status or situation of the objects in order to provide users the context-aware information.

Update of smart management systems:

Current object information like the status or location has to be communicated to the systems of the Smart Factory. As an example the highly dynamical data like the position of an object should be updated every 10 to 30 seconds [3].

Support for different queries:

Assistance systems in the Smart Factory have to support different kinds of queries [3]. We can differentiate into object-based, location-based/spatial, temporal and combinations of the previous types of queries.

Integration of heterogeneous information:

The integration challenge of different systems in an enterprise is caused by different information models, interfaces and data formats. In order to provide other systems easy accessible information, different systems have to be integrated into a common synchronizing platform.

Real-time characterized reaction:

For supporting people and machines information has to be provided within seconds. This challenge addresses mainly communications technologies and database management.

2.3 Enabling Technologies

The implementation of the Smart Factory is enabled by several technologies, in the following shortly presented (Figure 1). An application of the Smart Factory consists of a Calm-system (referring more to hardware components) and a context-aware application (referring more to software components).

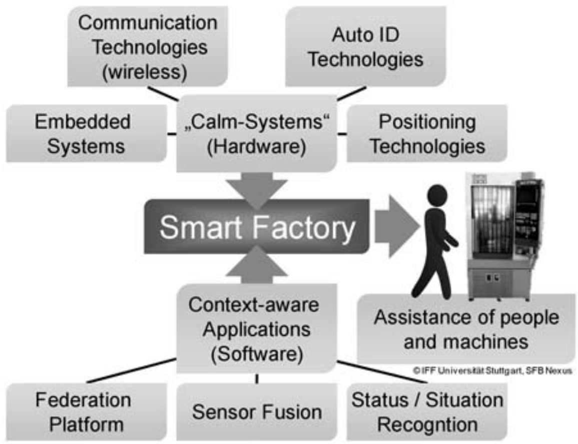


Figure 1: Smart Factory.

Embedded Systems:

To provide small computing power for decentralized intelligent functions, microcontrollers have been developed in the last few years. They are optimized to low energy consumption for working in mobile devices. Today a wide range of "easy-to-use" microcontrollers are on the market with different computing power ranges, like the AVR or Pic microcontroller or the Xscale processors families.

(Wireless) Communication Technology:

Calm systems have the ability to communicate with its environment. Today low mobile resources like a milling machine have got built in LAN or Fieldbus technology to communicate with higher-level enterprise systems like a MES. Recent developments tend to use wireless data transfers for low mobile resources as well as for mobile objects like tools or materials. These technologies are enabling ad-hoc networking of two or more devices and are usually classified after their cover range. Telecommunications technologies for second (e.g. EDGE/GPRS) and third generation (e.g. UMTS) cell phones and WiMAX/WiBro cover long distances up to several kilometers, whereas for short distances e.g. in a factory building, Wireless-LAN, Bluetooth/WIBREE and ZIGBEE are used.

Automatic Identification (Auto-ID) Technologies:

Automatic identification technologies are other essential technologies for a Smart Factory for assigning virtual data to real objects, especially Barcode and RFID (Radio Frequency Identification) technologies are important for manufacturing purposes. The advantages of RFID compared to Barcode-technology are its non-contact, non-line-of-sight nature, the possibilities to store more than just a number and integrating additional sensors e.g. for temperature and humidity. But barcodes are still wide spread in industry, due to its low costs.

Positioning Technologies:

Location information is probably one of the most important factors in ubiquitous computing. Methods like triangulation, scene analysis and proximity are mainly used to determine a position of an object [5].

Federation Platform (Nexus Platform):

The federation technology provides an easy access for applications to heterogeneous data. The Nexus Platform is developed within the Center of Excellence Nexus (SFB 627) provides context-aware's applications a homogenous interface for querying context information of different heterogeneous databases. The context information is stored the in so-called context servers. Currently two types of context server are existing, one for spatial data and one for highly dynamic location data of objects.

Situation Recognition:

In order to provide people in their tasks the right information at the right moment the assistance systems have to know the situation of their users. Therefore situation recognition becomes an important enabling technology for context-aware applications.

Sensor Fusion:

Today's applications often use hard coded systems specifically designed for their task, giving no information to other systems. Sensor fusion technology is used in the Smart Factory for combining different raw sensor values for an aggregation or better measurements. Furthermore having all kinds of sensor data easy accessible via the Nexus Platform in the Smart Factory environment new kinds of measurements can be created for an easy implementation of new context-aware applications.

3 POSITIONIING OF THE SMART FACTORY INTO THE FUNCTIONAL ARCHITECURE OF A MANUFACTURING ENTERPRISE

The manufacturing enterprise has to fulfill along its processes several main functions, graphically represented in Figure 2. As shown, the main phases of factory life cycle are investment planning, site and building planning, the planning of infrastructure and processes, layout planning, ramp-up and project management, manufacturing execution, maintenance, education and training. The architecture introduces the complexity of managing the data's diversity and heterogeneity, which comes from different applications and information systems. Due to the real-time informing and controlling character, the Smart Factory focuses mainly the manufacturing execution, maintenance, education and training functions of a manufacturing enterprise, as shown in Figure 2.

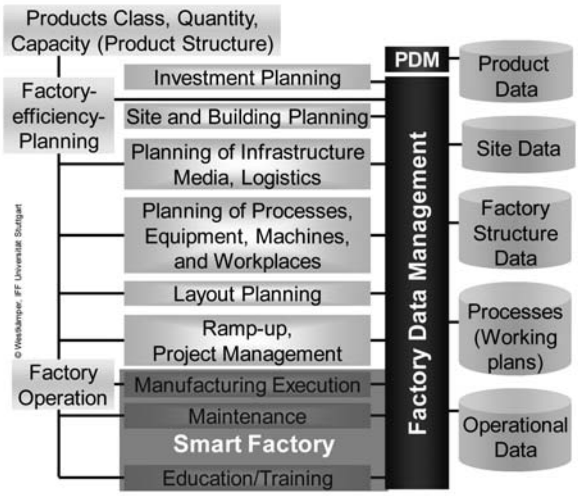


Figure 2: Functional architecture of a manufacturing enterprise (PDM: Product Data Management).

But also other functions in a manufacturing enterprise gain advantages from the up-to-date data of the Smart Factory. Especially functions from factory life cycle like the planning of processes and equipment can benefit from the information coming from the Smart Factory.

Opposite to other existing applications, the information of the Smart Factory is context-aware. The Smart Factory has the potential to canalize information overflow and provide people decision support in production activities. With context-aware applications the manufacturing processes can be optimized and times e.g. for tool searching reduced.

As an example, a situation of a maintenance task is shortly presented in the following. The scenario represents a breakdown of a milling machine and shows the advantages of the Smart Factory usage. As a first step the machine is sending an error message after breakdown. With this information the maintenance worker and the shop floor manager are notified. From here the error information will be analyzed and assigned to a suitable workflow. If it is already documented as an error the worker gets the repairing steps and required tools. Furthermore the locations of the tools are available and displayed in the application. Then the worker can order the tools or grab them itself on the way to the machine.

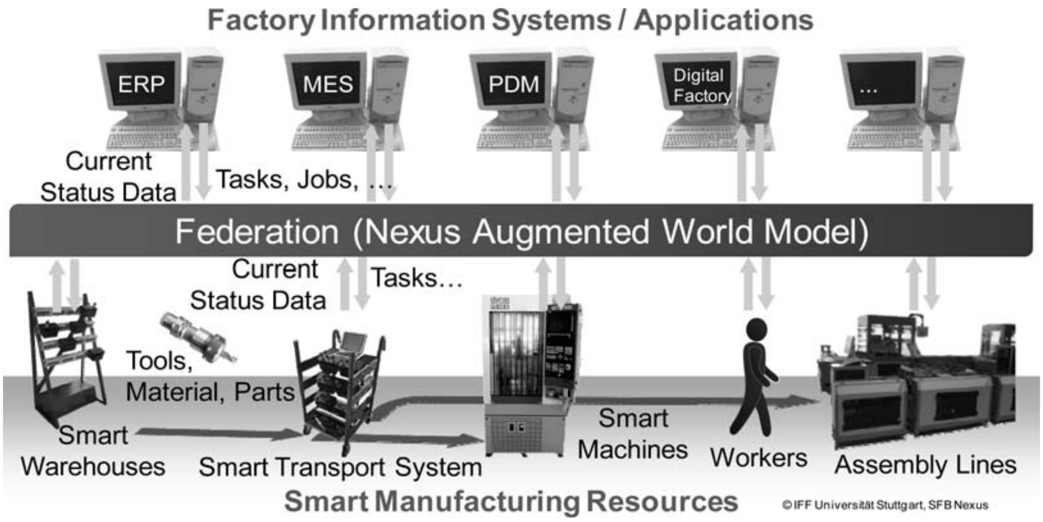


Figure 3: Architecture of the Smart Factory.

Arrived at the machine he can immediately start the repairing. So the times of tooling search are minimized. The workflow has to be updated with the time, actions and used tools in order that the repairing knowledge is saved for further usage. Furthermore with the time information the MES can reschedule the jobs and minimize the effects of the breakdown.

Figure 3 gives an overview of the envisioned Smart Factory. The modules like a MES or a Smart Machine are connected via the Nexus Platform with the Augmented World Model that provides the exchange of sensor, process, job, geometric and further information in the factory.

4 RELATED WORK

In the research field of Smart Factories several other related work, relevant for the scientific community, have to be mentioned. These are in the following shortly presented.

Smart FactoryKL

The Smart FactoryKL is a project located in Kaiserslautern, Germany. Current projects of Smart Factory KL are focused on the development of applications with Indoor Positioning Systems, universal Human Machine Interfaces, web-based information services and the integration of heterogeneous information systems [7].

INT-MANUS

The 'Intelligent Networked Manufacturing System' (INT-MANUS) project is financed by the European Commission under the 6th framework. Goal of INT-MANUS is the development of a new technology, the so called Smart-Connected-Control Platform (SCC platform) for manufacturing enterprises [8].

5 CONCLUSIONS AND FUTURE WORK

The downscaling of computer and sensor technologies supports the integration of knowledge at all scales of a holistic production system, aiming at increasing the transformability of the factory as a whole. The presented Smart Factory approach represents a real-time, context-sensitive manufacturing environment that can handle turbulences in production using decentralized information and communication structures for an optimum management of production processes. Based on the Nexus Platform it integrates heterogeneous information systems in a manufacturing enterprise both horizontal (e.g. between systems in the shop floor) and vertical (e.g. between shop floor and management systems like ERP or MES) in order to reduce information deficits. In the framework of the research area “Smart Factory” at the Universität Stuttgart, IFF, a Smart Factory environment has been developed and will be enhanced, aiming at implementing the vision of the next-generation real-time and context-aware production systems.

6 REFERENCES

[1] Westkämper, E.; Hummel, V.; Rönnecke, T., 2005, The Stuttgart Enterprise Model Integrated Engineering of Strategic & Operational Functions. In: IEMS Proceedings, March 14-16, Cocoa Beach, Florida USA, p. 379–384.

[2] Weiser M., 1991, The Computer for the Twenty-First Century, In: Scientific American, 09/91, p. 94–110.

[3] Westkämper, E.; Jendoubi, L.; Eissele, M.; Ertl, T., 2006, Smart Factory - Bridging the gap between digital planning and reality. In: Manufacturing Systems 35, No. 4, p. 307-314.

[4] Sonderforschungsbereich 627 Nexus-Umgebungsmodelle für mobile kontextbezogene Anwendungen, Available at: http://www.nexus.uni-stuttgart.de/, 24.01.2008.

[5] Hazas, M., Ward, A., 2002, A Novel Broadband Ultrasonic Location System, In: Ubicomp 2002 Proceedings, September 29 – October 1, Göteborg, Sweden, Springer Verlag, p. 264-280.

[6] Bauer, M.; Jendoubi, L.; Rothermel, K.; Westkämper, E., Grundlagen ubiquitärer Systeme und deren Anwendung in der "Smart Factory". In: Industrie Management 19 (2003), Nr. 6, p. 17-20.

[7] Smart FactoryKL, Available at: http://www.dfki.uni-kl.de/ smartfactory/, 24.01.2008.

[8] INT-MANUS, Available at: http://www.int-manus.org/index.php?id=212, 24.01.2008.

Sustainable Manufacturing - Challenges and Possibilities for Research and Industry from a Swedish perspective

Magnus Wiktorsson[1], Monica Bellgran[2] and Mats Jackson[3]

[1] VINNOVA - Swedish Governmental Agency for Innovation Systems, Sweden

[2] Haldex AB, Sweden

[3] Mälardalen University, Sweden

Abstract

The need for environmental sustainability and a globally increasing manufacturing, drive a gigantic need for technology and strategies that will reduce CO2 emissions globally. The paper reflects on the situation for Swedish research and industry, with a focus on technology for environmentally sustainable manufacturing. Sweden has competitive equipment and system suppliers to the manufacturing industry with potential of contributing to the development of a sustainable manufacturing industry globally. Environmentally conscious actions need no longer be seen as challenges and contrary to financial considerations. On the contrary, it can be the basis for successful companies. Enabling features includes manufacturing systems and supply chain structures for sustainability e.g. local manufacturing providers, high energy efficiency in manufacturing processes and techniques for replacing fossil based energy generation by CO2 neutral generation.

Keywords:

Sustainable Manufacturing; Sweden; Manufacturing System; Manufacturing Technology

1 INTRODUCTION

There is a long seen need for an environmental, economic and social sustainable society – a society meeting the needs of the present without compromising the ability of future generations to meet their own needs [1]. With a focus on environmental issues, the need for sustainability has lead to emerging legislation and industrially accepted emission targets.

The Kyoto Protocol contains legally binding commitments for industrialized countries to reduce their emissions of greenhouse gases by a total of at least 5% [2]. In this total effort, the manufacturing industry has an important role. In Sweden, the industry contributes with 11% to the CO2 emissions. In a recent governmental commission the target is set to reduce the industrial oil consumption with 25-40% by 2020 [3]. The EU is committed to green growth through the Lisbon strategy from 2000, later materialized in the Göteborg strategy. The European Commission initiated a development and wider use of environmental technologies through implementing the Environmental Technologies Action Plan (ETAP), with 28 defined actions to be implemented at European, national, regional or local level [4].

Meanwhile, a global wealth increase is evident. Through the means of globalisation, values are created with a worldwide 60% increase of GDP in current prices over 1996-2006 [5]. By increased welfare the product demand increases, leading to increased manufacturing activities – we have seen a 43% increase of manufacturing activities worldwide 1996-2006 [5]. In figure 1 the increase in economic activity in current prices within manufacturing over 1996-2006 is presented for the largest manufacturing countries. For instance China has seen an unprecedented manufacturing increase over the recent decade.

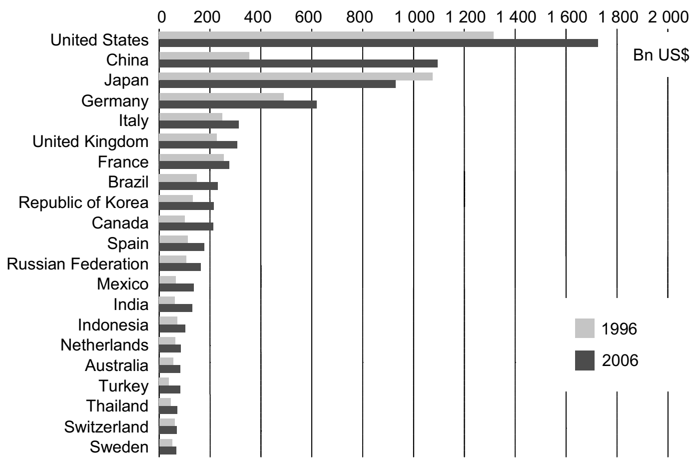


Figure 1. Manufacturing activities 1996 and 2006. Source: UNSTAT [5].

These two trends, the need for sustainability and the globally increasing product demands and manufacturing activities drive a gigantic need for technology and strategies that globally will reduce environmental impact from manufacturing. There is a need of large improvements in terms of resource productivity - "doing more with less". The challenge is to reduce non-renewable material and energy usage in a fast increasing economic activity.

This paper reflects upon technological possibilities to address this global challenge for environmentally sustainable manufacturing, based upon Swedish industrial and research challenges.

2 THE INDUSTRIAL PERSPECTIVE

2.1 Introducing Swedish industrial structure

Sweden has seen a 59% increase of GDP in current prices over 1996-2006 [6]. The value adding from manufacturing activities have in the same period increased by 45% in current prices. Manufacturing companies' contribution to Sweden GDP was 18% in 2006. The structure of Swedish manufacturing industry is illustrated in Figur 2.

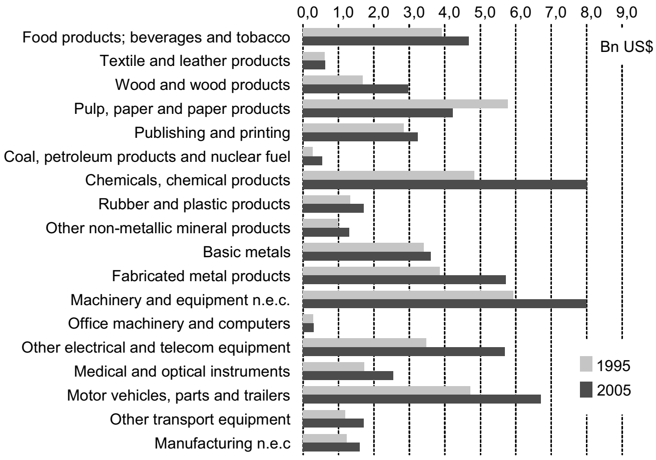


Figure 2. Value added from manufacturing sectors in Sweden 1995 and 2005. Source: Statistics Sweden [6]

Sweden has a large and globally active **machinery and equipment industry** with competitive suppliers to the manufacturing industry. Companies such as SKF, ABB and Alfa Laval operate on global markets, offering industrial products and services that can be considered as clean tech investments, improving operational performance, productivity or efficiency while reducing costs, inputs, energy consumption, waste or pollution. Other significant manufacturing sectors seen in Figure 2 are automotive, pulp and paper, life science and telecom.

2.2 Sustainability challenges and possibilities

Focusing on machinery and equipment industry as one example, this traditional sector can play an important environmental role by progressive actions. Not only by decreasing the environmental impact from their own processes by e.g. energy efficiency actions, remanufacturing and material reuse, but by offering radically new energy preserving technologies and solutions for the fast growing global market. The clean tech investment offerings can support carbon smart manufacturing (to take whatever steps possible to minimise the impacts of emissions) and carbon neutral manufacturing (zero carbon release).

2.3 Haldex AB: One example

Haldex is a provider of proprietary and innovative solutions to the global vehicle industry, with focus on products in vehicles that enhance safety, environment and vehicle dynamics. Haldex, headquartered in Stockholm, Sweden, has a yearly turnover of 1,2 bn US$ and employs 6,100 people. Demand is driven in part by that vehicle production is increasing world wide, and in part because demand for technology to improve safety, the environment and vehicle dynamics is also growing. Haldex is building up an increasingly global presence in order to meet this demand.

Today, the environmental policy at Haldex is based on a life cycle perspective taking the total environmental impact on manufacturing, use and disposal of the products into account. Climate and sustainable requirements are put on products and service purchase, and 90% of the corporate business is certified for ISO 14001 or similar environmental standards. Haldex was recognized as best in its business, and No 3 in Sweden according to Climate Index 2007 made by the organization Carbon Disclosure Project [7].

Already providing a product portfolio with a sustainable focus creates new challenges to move the sustainability approach forward as today's environmental order-winners are most certainly tomorrow's qualifiers.

One future challenge is to design sustainable production systems, including e.g choice of material and equipment, energy efficiency and generation, media consumption etc. Another challenge involves the supply chain and how to reduce climate impact of the whole sourcing process. Sustainability is no longer only a cost centre, but rather an investment for competitiveness.

3 A STRUCTURE FOR ENVIRONMENTAL ACTIONS

It is generally accepted that environmentally conscious actions need no longer be seen only as challenges and contrary to financial considerations. On the contrary, it can be the basis for competitive companies contributing to global environmental improvements. As Hart [8] suggested; 'the basis for gaining competitive advantage in the coming years will be rooted increasingly in a set of emerging capabilities such as waste minimisation, green product design, and technology cooperation in the developing world'.

Environmental actions within manufacturing industry do typically have a product or a process focus. Product focused actions concerns developing products with radically less environmental impact for the user. Process focused actions concerns decreasing the environmental impact of the development and manufacturing of the products. These two aspects are often integrated in a total life-cycle estimation of the environmental impacts from a product.

At the same time, competitiveness can be described as creating great value by low costs, where the values and costs are given at the market place. Improvements in competitiveness can thus be structured in value improvement and cost cutting actions. These two perspectives are closely linked to Hill's order-winning and order-qualifying factors [9]. As concluded by Shahbazpour and Seidel [10], sustainability can be classified as an order winning or an order qualifying objective depending on market, society and technology.

By using these perspectives on environmental actions and competitiveness, four fields of action for leverage on environmental sustainability are described in Figure 3.

	Value improvement	Cost decrease
Product focus	**"PRIUS"** Material and technology selection improving customer value	**"XEROX"** Material and component reuse. Remanufacturing.
Process focus	**"ORGANIC AGRICULTURE"** Selling by sustainable processes	**"SUSTAINABLE LEAN"** Waste reduction

Figure 3.4 types of actions for environmental sustainability.

3.1 Value improvement through product focus ("PRIUS")

A typical proactive sustainability action is to develop product and service solutions with neutral environmental impact while used. The actions are closely linked to material and technology development creating customer value. Examples are alternative fuel solutions for vehicles (thereof the Prius label) or energy saving technologies that implies a higher value/cost ratio for the customer. Actions focused on the product, specifically for industrial products can also lead to fulfilling qualifying regulations and environmental criteria set by the customer.

Enabling features include:

- CO2-neutral materials for replacing fossil based materials in products.
- Technology development for environmental and energy preserving offerings.
- Business development including product service solutions for total life cycle sustainability.

3.2 Value improvement through process focus ("ORGANIC AGRICULTURE")

The economic objective with this type of actions, is to be able to create added value for the customer by sustainability actions within the manufacturing process. An illustrating example is organic agriculture products where the customer is willing to pay a price premia, due to the sustainability actions in the food production process.

Within manufacturing this type of actions correspond to both added value for the customer such as local manufacturing giving fast customer response and customisation, as well as fulfilling qualifying regulations and sustainability criteria set by the customer on the suppliers manufacturing process. Jackson and Zaman [11] presents the Factory-In-a-Box-concept as one illustrative example on manufacturing concepts well suited for local manufacturing and fast customer response while having efficient resource utilization.

Enabling features include:

- Manufacturing capabilities for new types of environmental and energy preserving products.
- Standard interfaces enabling rapid and customized manufacturing system set up.
- Environmental impact assessment and certification tools.
- System modelling and solutions for local manufacturing providers.

3.3 Cost decrease through product focus ("XEROX")

The concept of remanufacturing is often quantified in terms of cost advantages and refering to activities designed to reclaim value from a product at the end of its useful life. Industries that apply remanufacturing typically include automobiles, electronics and tyres.

Lindahl et al [12] argues for a more proactive view on remanufacturing and identifies general environmental pros and cons with remanufacturing, such as the material resource perspective. The authors further review literature concerning the environmental impacts of remanufacturing. Through product design based on component and material reuse, more drastic cost advantages can be gained. Kerr [13] performed a comparison between the remanufacturing of a traditionally-designed XEROX copy machine and a copy machine that was designed to facilitate remanufacture. For the model which has been designed for remanufacturing, the savings of energy equal a factor of 3.1 and those of materials and landfill waste a factor of 1.9.

An even more comprehensive review on literature regarding green supply chain management is presented by Srivastava [14]. It includes the aspect of green manufacturing and remanufacturing.

Enabling features include:

- Design and industrialise products for sustainability.
- Reuse/remanufacture material and components within manufacturing processes, especially fossil based material.

3.4 Cost decrease through process focus ("SUSTAINABLE LEAN")

Cutting cost by a process focus concerns efficient resource utilization. One dimension is energy efficiency, where techniques such as heat recovery and energy co-generation can contribute. In the literature a number of indicators are presented, basically reduced to theoretical thermodynamic efficiencies, and various economic measures of efficiency e.g. energy per output sometimes called energy intensity or technical energy efficiency; energy per $ of GDP (or profit, etc); or energy cost ($) per $ of GDP (or profit, etc) [15]. However, as Love and Cleland [16] concludes, the evaluation of energy efficiency is ultimately a comparative exercise; to make meaningful decisions about energy efficiency the measured efficiency of a process must be compared to a benchmark. Furthermore, the aspects of pricing and legislation have large impact on energy efficiency and its means and measures.

The most appropriate process is the most resource efficient. The trick is however to make it appropriate over time. The concept of lean manufacturing deals with resource efficiency of all kinds, and can also have positive impact on sustainability. Linden and Carlsson-Kanyama [17] note that in the Swedish energy efficiency program education was an important factor in energy saving achievements by industrial groups. An effort to estimate the linkage between energy efficiency and productivity is presented in [18].

Enabling features include:

- Manufacturing systems and supply chain structures for sustainability and minimized waste and media consumption – manufacturing beyond lean.
- High energy efficiency in manufacturing processes.
- Techniques for replacing fossil based energy generation by CO2 neutral generation.

3.5 The importance of a total perspective

Even though actions are structured in the earlier categories, the need for a total sustainability assessment can not be over estimated. Sub optimizations are easily made when new products with apparent positive environmental impact is launched, while the total life cycle impact is negative.

Enabling features include:

- Decision support tools for sustainability and LCC evaluations.
- Certification methods and regulations for total system environmental sustainability.
- Management systems for monitoring sustainability.

4 SUMMARY: RESEARCH & INDUSTRY CHALLENGES

Based on the industrial challenges and the proposed structure for environmental actions, more detailed research challenges must be identified for specific actors and sectors.

- For studies on **manufacturing systems** important fields include: lean operations; systems modelling; multi-criteria analysis; operation/value chain control and economics.
- For studies on **manufacturing technologies** important areas include: robust and carbon neutral processes for new technologies, materials and combinations of these; control and sensor technology; tribology.
- For studies on **product and process development** important areas include: life cycle design for product service solutions; material science; bio based materials.

The need for environmental sustainability creates both restrictions and opportunities for the manufacturing industry. Products and transports/logistics have been first in focus for environmental impact improvements. However, to reach a total effect, the manufacturing process also needs to be looked upon from an environmental perspective. Process industry has worked with these issues longer than the manufacturing industry due to high energy consumption and potentially polluting processes. Another challenge for manufacturing companies are recruitment of talent. A focus on the environment can for a company be the difference in creating a positive and modern image.

The manufacturing company has a choice of merely following the new legislations, customer demands and competitors actions – or deciding to take lead by utilizing good environmental performance with low or zero carbon effect as an order-winner.

5 REFERENCES

[1] UNWCED - World Commission on Environment and Development, 1987, Our Common Future, Oxford: Oxford University Press.

[2] UNFCCC, 1997, Kyoto Protocol UNFCCC: New York. Available online at: http://unfccc.int/resource/docs/convkp/kpeng.html

[3] Commission of oil independence, 2006, Making Sweden an OIL-FREE society. Government Offices of Sweden, Prime Minister's Office, Information material.

[4] Calleja, I., Delgado, L., 2007, European environmental technologies action plan (ETAP), Journal of Cleaner Production, 16, 181-183.

[5] UN Statistics divions, 2008, Data available online on: http://unstats.un.org/

[6] SCB, 2008, Statistics Sweden – government authority for official statistics. Database available online on http://www.scb.se .

[7] Carbon Disclosure Project, 2007, CDP5 Nordic Report. Available at http://www.cdproject.net

[8] Hart, S.L., 1995, A natural resource-based view of the firm. Academy of Management Review. 20(4): p. 986-1014.

[9] Hill, T., 1995, Manufacturing Strategy—Text & Cases. London: Macmillan Press Ltd.

[10] Shahbazpour, M., Seidel, R., 2006, Using Sustainability for Competitive Advantage, Proceedings from 13th CIRP International conference on Life Cycle Engineering of Life Cycle Engineering, Leuven, May 31-June 6, 2006.

[11] Jackson, M., Zaman, A., 2007, Factory-In-a-Box – Mobile Production Capacity on Demand, International Journal of Modern Engineering (IJME), 8(1).

[12] Lindahl, M., Sundin, E., Östlin J., 2007, Environmental Issues within the Remanufacturing Industry. Proceedings from 13th CIRP International conference on Life Cycle Engineering of Life Cycle Engineering, Leuven, May 31-June 6, 2006.

[13] Kerr, W., 1999, Remanufacturing and Eco-Efficiency: A case study of photocopier remanufacturing at Fuji Xerox Australia, in Environmental Management and Policy, International Institute for Industrial Environmental Economics (IIIEE), Lund University: Lund, Sweden. 1999, pp. 56-68.

[14] Srivastava, S. K., 2007, Green supply-chain management: A state-of the-art literature review, International Journal of Management Reviews, 9(1) 53-80.

[15] Patterson, M., 1996, What is Energy Efficiency? Concepts, indicators and methodological issues. Energy Policy 24(5) 377-390.

[16] Love, R., Cleland, D., 2007, Tools for Energy Efficiency in Industrial Processes. Proceedings of the 2nd International Conference on Sustainability Engineering and Science, Auckland, NZ, Feb 21-23, 2007.

[17] Linden, A., Carlsson-Kanyama, A., 2002, Voluntary agreements — a measure for energy-efficiency in industry? Lessons from a Swedish program. Energy Policy 30 897-905.

[18] Boyd, G.A. and Pang, J.X., 2000, Estimating the linkage between energy efficiency and productivity. Energy Policy 28 289-296

Characteristic of a Proactive Assembly System

Kerstin Dencker[1,2], Johan Stahre[3], Åsa Fasth[3], Peter Gröndahl[1], Lena Mårtensson[4] and Thomas Lundholm[5]

[1] Royal Institute of Technology, Department of Production engineering

[2] Swerea IVF, Industrial Research & Development Corp.

[3] Chalmers University of Technology, Division of Production Systems

[4] Royal Institute of Technology, Department of Industrial economics and management

[5] Royal Institute of Technology, Centre for Design and Management of Machining Systems

Abstract: Competitive assembly systems must cope with frequent demand changes, requiring drastically shortened resetting and ramp-up times. Characteristics of assembly systems capable of rapid change are e.g. *Flexibility*; *Robustness, Agility*, and ability to handle frequent changes and disturbances. This paper proposes *proactivity* as a vital factor of semi-automated assembly systems to increase speed of change. Proactive systems utilize the full potential of human operators and technical systems. Such systems have ability to dynamically change system automation levels, resulting in decrease of time consumed for assembly tasks. Proactivity criteria for assembly systems are reviewed based on theory and industrial case studies

Keywords:
Assembly System, Proactive, Automation

1 INTRODUCTION

Present practices for development, design and use of assembly systems may not be adapted to the needs and future challenges of industry. Competitive systems for manufacturing, specifically assembly systems, will have to cope with increasingly frequent changes of product variants as well as increased variation in production volumes. At the same time, cost efficiency and "leanness" require appropriate quality, lower cost, and strong focus on value-adding activities. Preferably, product customization should be made as late as possible in the value-adding chain, i.e. mass customization [1]. Such demands require assembly systems, which is reliable, have high availability and have ability to produce the right product correctly. This means a combination of short resetting time and robustness of the system as whole and its resources. A major challenge is to reduce and minimize the lead-time that direct has influence on order-to-delivery time, while maintaining flexibility and robustness to absorb late requirement market changes. This includes throughput time and cycle times for individual assembly processes. Also, there is a need to minimize resetting time between batches, time to repair, time for disturbance handling, and time to prepare for new variants or products in the assembly system. How can a radical reduction in the time needed to fulfil an assembly sequence be achieved?

The objective of the ProAct project is to identify proactive solutions for time minimization at operational shop floor level in assembly systems. The approach is based on the concept of proactivity; taking action by causing change towards a state and not only reacting to change when it happens [2]. ProAct is a collaborative effort involving six Swedish industries and three Swedish technical universities.

2 BACKGROUND

To minimize the time factors suggested, while remaining competitive, assembly system behaviour has to be proactive and that requires highly transformable system resources.

Dencker et al. [3] suggest that a proactive assembly system has the ability to prepare for:

- changes and disturbances during operation
- planned and long-term changes, and sustainable evolution of an assembly system

Main features required to prepare the assembly system are:

- flexibility
- robustnessspeed of change
- ability to handle frequent changes
- evolvability

In terms of system resources, we suggest that the proactive assembly system constitutes an integrated combination of the competence of *"knowledge workers*" [4] *information*, and *automation*. The proactive assembly concept was presented by Dencker et al. [5], along with arguments for focus on levels of automation, information, and competence as main drivers for the degree of proactivity.

A Proactive Systems should be able to respond quickly to a dynamically changing environment in real-time contexts. Such a system should also have the ability to act proactively and preventive before a situation emerges to a confrontation or crisis.

Proactivity should be employed in two different time perspectives; to be able to act proactively during operation and to maintain the proactive ability over time. The operation perspective involves manual, semi-automated, and automated operations and tasks to be handled proactively. The second perspective focuses on continuous changes to the system, in order to absorb disturbances and changing requirements.

3 THE PROACTIVE ASSEMBLY SYSTEM STRUCTURE

3.1 System model

According to Chapanis [6], a system is an interacting combination at any level of complexity, of people, material, tools, machines, software, facilities, and procedures designed to work together for some common purpose.

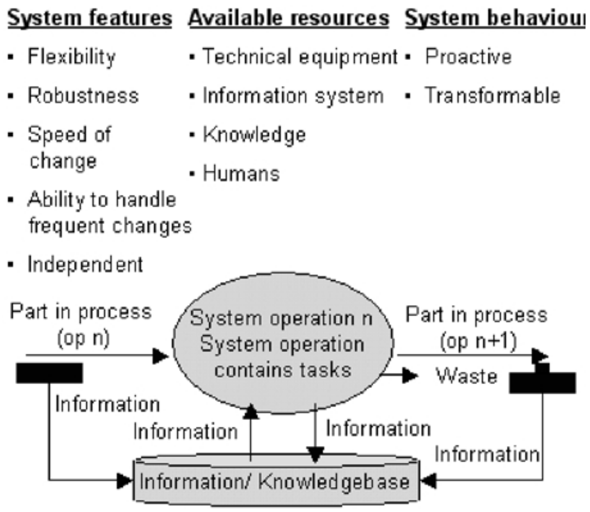


Figure 1 Theoretic model of an assembly system operation

Tharumarajah et al [7] describes that Self-organised systems can be developed through a Bionic Manufacturing System (BMS) protomodel [8]. Self-regularized and dependent systems can be used by applying Holonic Manufacturing System concepts [9]. The difference between evolvability in an operationally and managerially dependent and independent (BMS) system is that a dependent system is able to evolve in one general direction with internal congruence. For the independent system, each subsystem possess has that ability, making the whole system far more agile and responsive to internal and environmental changes. According to [10, 11], an Evolvable assembly system is a system which is "being based on many simple, re-configurable, task-specific elements (system modules), this allows for a continuous evolution of the assembly system. "The proactive assembly system idea is to add skills from a human "agent" as a core assembly system resource, this will make the system more independent and it will receive a larger range by independence.

Development of semi-automated assembly system structures is generally framed by technical, ergonomic, and human work requirements. These contribute with constraints for the system structure. De Toni et al. [12] describes **system process flexibility** as 'the ability to produce a given set of part types', see figure 1. There are two distinct ways of being flexible for a given dimension of change and time period, i.e. range flexibility and mobility flexibility according to Upton [13]. *Range number* is a strict numerical count of the number of possible options that a system or resource can achieve. Mobility, or transformability, is defined as the ease or effort with which the system moves from one state to another. For each change of state within a specific range, transformability is assessed via transition penalties.

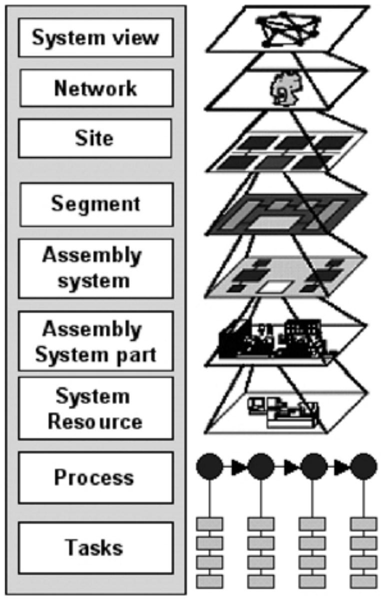


Figure 2 System process definition

The definition of assembly system flexibility focuses on the range flexibility and transformability that is described as the set of options or alternative process performance in a predefined dimension and aims for a more operational and managerial independence, see definition below. The transformability is the ability to move between optional states or system configurations. State-changes should be performed with a minimum of transition penalties, e.g. setup time for reconfiguration or re-routing.

Bjelkemyr [14] defines operational and managerial dependence (a) and independence (b) that can be seen in figure 3 In the left view, the tinted sectors signify that each subsystem is unable to function satisfactory on its own, i.e. it is operationally dependent on the other sectors to create a whole. The striped circle in the middle signifies that they all have one joint management looking after the whole system.

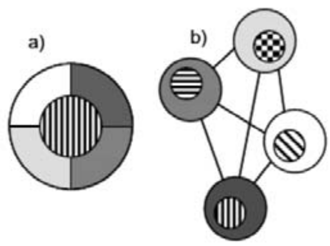


Figure 3 operational and managerial dependence (a) and independence (b)

Independent and constituent systems are useful in their own right and generally operate independent of other systems.
The independent system has a larger range but the transformability must be agile if the system should be considered as efficient. The Reconfigurable Manufacturing Systems paradigm [15] is an example of predefined – dependent- range flexibility and fast transformability. *Robustness* is the ability to handle predictable and

unpredictable variations with minimal loss of functionality causing penalty time. It determines the range of magnitudes of changes within which the feasible respond occur. Low values of transition penalties imply agile transformability in the system and transition penalties are dependent on time parameters.

3.2 System resources

Resources in an assembly system are:

Technical equipment
Automation that is flexible and quickly adjustable to different levels of automation in the assembly system. This applies to mechanical/physical as well as information/cognitive levels of automation. [16-18]

Information system
Efficient and dynamic flow of predictable as well as unpredictable information between assembly, product development, production planning preparation, suppliers, marketing etc. [19-20]

Knowledge
The system knowledge, facts and methods, as the knowledgebase carried by humans or computer.

Human operators
Operator team that have the knowledge of system tasks and competence to perform demanded knowledge based task. Quick, precise, and efficient competence development for assembly operators

The ISO definition SS 62 40 70 [21] defines competence as *Ability and willingness to carry out a task by applying knowledge and skills.* When defining competence the following implications have been to the words used: *Ability – experience, comprehension and judgment to use knowledge and skills in practice, where willingness is the attitude, commitment, courage and responsibility; knowledge means facts and methods –to know and skills is to carry out in practice – to do.*

Resources as they are defined in figure 3, include human resources considered as specific means with a given capability and a given capacity. Those means are considered as being able to be involved in the manufacturing process through assigned tasks. That does not include any modelling of an individual or common behaviour of human resource excepted in their capability to perform a given task in the manufacturing process (e.g.: transformation of raw material or component, provision of logistic services). That means that human resources are only considered, as the other, from the point of view of their functions, their capabilities and their status (e.g. idle, busy). That excludes any modelling or representation of any aspect of individual or common «social» behaviour. By adding the concept of ability to vary level of automation [22] the range of possible system solutions becomes higher but puts new requirements on the competence of the operator team and the information. This Level of Automation concept was defined in Frohm [23]. To realize the productivity potential of automated equipment human involvement is needed [5] to support technical equipment. As stated, the combination of human and machine skills and the possibility to allocate tasks is widening the range of possibilities. Relevant human involvement will be different depending on the situation and the assembly system design. The information system must provide assembly systems with correct and essential information, contributing to the operators' decision support.

3.3 System transformation and System tasks

Ability to vary levels of information, levels automation and competence has earlier been identified as three important factors in a proactive assembly system [3]. The system transformation process defines the tasks that should be performed during set up and the task allocation for desired processes. Such an assembly system needs improved information exchange and this requires an adequate information and decision support system. [19]. Further, the system needs competent operators with ability and mandate to make decisions concerning short term planning, recourse allocation and actions in disturbance handling. Proactive assembly system require a team of operators that are able to exercise their own judgment, "think by themselves", proceed on their own initiatives, and anticipate future problems, in order to avoid exceptional production-, or work situations. The tasks performed by the proactive assembly system are e.g. performing assembly tasks, short-term planning, real-time recourse allocation and actions in disturbance handling. The assembly system task are structured by using Stahre, [24] planning, programming, monitoring, performing, intervening, and learning. For the proactive assembly system to achieve and maintain an acceptable level of performance considering time parameters defined on section 1, the interaction between human and machine demands that the collaboration of systems must correspond to a set of tasks that the operators and/or machine are able to perform at any time. These include the demand to perform the operations included in the process, and to response within the constraints provided by time and system operating requirements. To achieve this correspondence, the design must be based on a thorough analysis of the tasks included in the operations, meaning the tasks that are supposed to be carried out in the process. This includes how the tasks will be carried out in practice and what resources the system has like information, e.g. task description, and competence.

The system transformation process should be as short as possible over a running time during a continuous product variation flow. According to Rasmussen's [25] task definition the skill and expertise of the operator affect the human behaviour, which can be categorized in Rasmussen's three levels of performance: skill, rule-, and knowledge-based (SRK) behaviour [25].

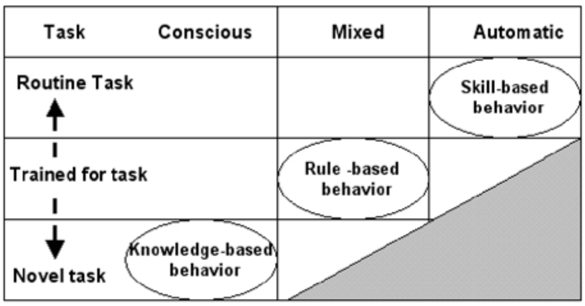


Figure 4 Rasmussen SRK form a task perspective

Each level of the SRK taxonomy defines different ways of representing the constraints in the environment and thus different level of cognitive control of the operator, see figure

4. The skill-based behaviour takes place without conscious attention or control and consists of smooth, automated, and highly integrated patterns. Rule-based behaviour is characterized by pattern matching with stored rules derived from previous successful experiences. At the rule-based level people are aware of their cognitive activities, and hence, can verbalize the used rules. The knowledge-based behaviour is required in unfamiliar situations and demands a conscious, focal attention of the operator. In these situations previous experience is no longer valid and a solution must be improvised by functional reasoning. The task analyze

4 TASK ALLOCATION

The allocation of tasks has traditionally been based on rather simple principles, which tend to consider the system in terms of its parts rather than as a whole. The left-over principle means that tasks that have not been automated, due to either technical or economical reasons, are assigned to the humans in the system. The compensatory principle, also referred to as Fitts' list, uses a list or table of the strong and weak features of humans and machines as a basis for assigning functions and responsibilities to the various system components [26]. More recently, a complimentarily principle for the allocation of functions has been advocated. Instead of focusing on the capabilities and limitations of the components in the system, the focus is on how the humans and machines can complement and support each other. Human and machine functions are not seen as being in competition or as being replaceable, but as being mutually dependent and necessary to achieve the overall purpose-Joint System [27]. The allocation of tasks should thus serve to maintain control of the situation and support the retaining of human skills according to Grote [28].

The possibilities to dynamically allocate tasks between human and technical resources are required. What do we expect from the system and what duties, tasks and actions will the system be responsible for e.g. information handling, communication and decision making in normal operations and occurrences (predicted and unpredicted)? Instead of having the view that the operator is a user of the system the humans and the machines are resources within the system, figure 3. A structured analysis of the skills of a human “agent” as a core assembly system resource should be an integrated part of system development. To support the idea of resource allocation the system tasks has been defined from a SRK view but only by its character not assuming that the task is allocated to the human.

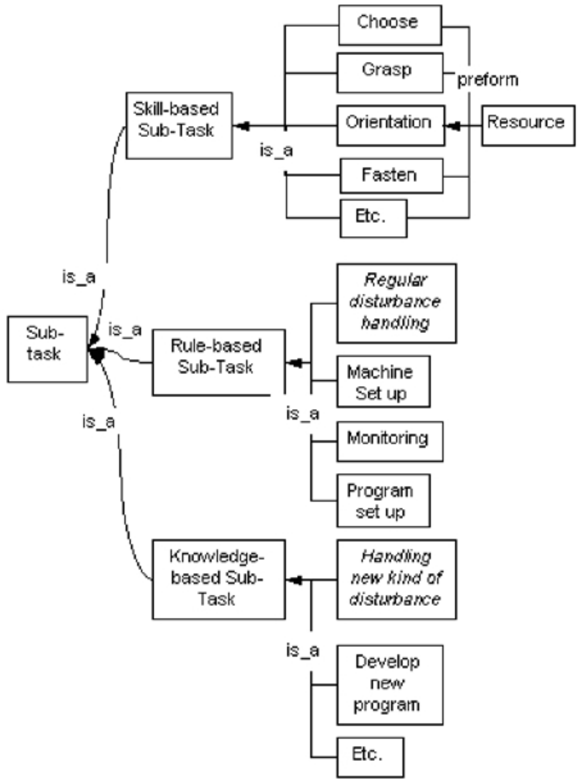


Figure 5 Categorization of assembly tasks according to Rasmussen SRK

5 INDUSTRIAL CASE STUDIES

As previous suggested, To enhance of understanding the relation of LoA and proactivity four case studies was performed during autumn 2007 at four industrial companies with final assembly in Sweden. The case studies focused on the easies of the defined tasks in an assembly system, perform assembly task. The measurement of LoA was conducted with a methodology for level of automation assessment called DYNAMO [22]. In the case studies evaluated the time parameters according to in relation of LoA. The purpose was to detect time difference of performing the task assembly according to different competence. For a proactive assembly system a task like perform assembly defined as a skill-based task but the surveys defined that the individual operator performance time was dependant on experience according to the contribution to lead time. Though the competence for the assembly skill-based tasks goes from low to high in a short time if the instructions of the tasks support novel operators.

Given a fixed competence in the operator group and a fixed LoI an increase of LoA at first decreases the time to perform the task, investigated system are represented by varying LoA solutions illustrated in fig 6. Frohm [22] argues however, that when the system reaches an increased LoA the effects of automation deviates negative from the expected because of e.g. breakdowns. According to Harlin et al. [28] is the condition of realizing the potential of automation to decrease the negative effects caused by system complexity originate from high LoA.

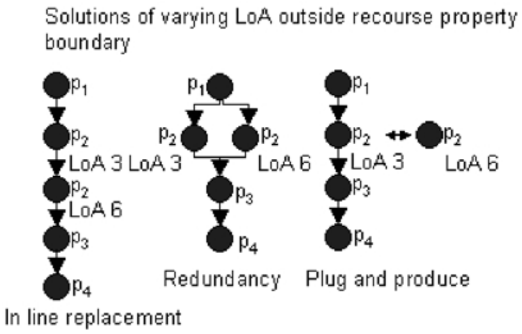


Figure 6 Examples of investigated LoA solutions.

According to the survey the situation with fixed competence in the operator group and fixed LoA but the possibility to vary the amount of information gave reduction in the setup time as well as the task time. The decreased setup time is a result of when the assembly system knew ahead what products that in the queue their made a last minute optimization of the future setups within predefined frames. They where able to change the order of the next ten orders and the sequence was delimited of product similarities. A product that had to many similar but not same components, was not allowed to be assembled after each other because of the high mistake rate. The possibility to rearrange between the orders gave possibilities to make the setup times shorter and more setup was performed as external setup. The higher amount of information to the operator team gave the effect that if a machine did not fulfil the requirement the group allocated the desired task to an operator until the machine got repaired.

6 DISCUSSION

Proactive system behaviour may not be described in terms of classic hierarchical manufacturing system models. A modern assembly system must be independent operational and managerial to support an agile transformation of the resources. Given this independency, tasks can be reallocated proactively. However our studies show that tasks cycle time are affected by LoI, LoA, LoC and their interrelations. The results shows that there is a connection between the parameters LoA, LoI and LoC. Naturally; the lead-time is directly related to the operator’s experience to perform the assembly task. This also influences potential proactivity. Further development of tools for defining LoC and LoI will give the possibility to measure the variation of all three parameters in real time. By enabling task reallocation in real time, decreased operational times like setup and task time will be possible. The LoA concept gives the system a wider range of possible solutions but the range can be delimited by the level of competence of the operator group when it comes to difficult tasks like disturbance handling and reconfiguration.

7 CONCLUSION

The proposed proactive approach for structuring future, semi-automated assembly systems is a strategic solution for integrating human and technical resources in an efficient system. It is not the person, nor the machine that is capable of proactivity, but rather the fully integrated system of human and technical resources. Thus, proactive systems will decrease the lead time in assembly. By reducing time parameters in the chosen areas; LoA, LoC and LoI. It is our belief that interrelated adjustment and transformability received by ability to vary the level of automation, the level of information, and the level of competence will result in a radical decrease of time consumed for assembly tasks.

8 ACKNOWLEDGMENTS

The research is financially supported by the Swedish Agency for Innovation Systems (VINNOVA) as it is a part of the ProAct project (http://proact.iip.kth.se/). The authors would like to thank the companies participating in the ProAct project.

9 REFERENCES

[1] Jackson, M., An analysis of flexible and reconfigurable production systems: an approach to a holistic method for the development of flexibility and reconfigurability. Linköping, Univ., 2000, Lindköping, Sweden ISBN 91-7219-797-8.

[2] Cambridge Advanced Learner's Dictionary, 2008.

[3] Dencker, K., Stahre, J., Gröndahl, P., Mårtensson, L. Lundholm, T., and Johansson C., 2007, An approach to proactive assembly systems – towards competitive assembly systems, IEEE international symposium on assembly and manufacturing (ISAM), University of Michigan, Ann Arbor, US.

[4] Drucker, P. F,1999, Management Challenges for the 21st Century. Harper Business, USA.

[5] Dencker, K., Stahre, J., Gröndahl, P., Mårtensson, L., Lundholm, T., Bruch, J. and Johansson C., 2007, Proactive assembly systems – realizing the potential of human collaboration with automation, In proceedings of the IFAC-CEA, cost effective automation in networked product development and manufacturing, Monterrey, Mexico.

[6] Chapanis, A., 1996, Human factors in systems engineering, New York: John Wiley & Sons, Inc.

[7] Tharumarajah, A., Welles, A.J., Nemes, L., 1996, Comparison of the bionic, fractal and holonic manufacturing system concepts, International Journal Computer Integrated Manufacturing, 9 (3) pp. 217-226.

[8] Okino, N., 1980, Bionic Manufacturing Systems (BMS) Bio-Modelon based systems design. CAM-I, Annual International Conference, New Orleans (US).

[9] Bell, R., Rahimifard, S. and Toh, K.T.K., 1998. Holonic systems, in: Handbook of Life Cycle Engineering:Concepts, Models and Technologies,. Molina, A., Sanchez, J.M., Kusiak, A., (Eds), 115-150, Ed. Kluwer Academic Publishers, London, UK.

[10] Lindberg, B., Onori, M., Semere, D., 2007, Evolvable Production System a position paper, In Swedish Production Symposium, Gothenburg, Sweden.

[11] Semere, D., Barata, J. and Onori, M., 2007. Evolvable Production Systems, Developements and Advances, tbp. IEEE Int. Symposium on Assembly and Manufacturing. Ann-Arbor, US.

[12] De Toni, A., Tonchia, S., 2005, Definitions and linkages between operational and strategic flexibilities, Omega, 33:525-540.

[13] Upton, D.M., 1996, What Really Makes Factories Flexible? In: Hayes, R.H., Pisano, G.P., Upton, D.M. (eds), Competing Through Capabilities, The Free Press, New York, US.

[14] Bjelkemyr, M., Semere, D., Lindberg, B., 2007, An Engineering system perspective of System of Systems Methodology, 1st Annual IEEE Systems Conference, Honolulu, HI, US.

[15] Mehrabi, M.G, Ulsoy, A.G, Koren, Y., 2000, Reconfigurable Manufacturing Systems: Key to Future Manufacturing, Journal of Intelligent Manufacturing, Vol. 11, No. 4, pp 403-419.

[16] Fasth, Å., Stahre, J. and Dencker, K., to be submitted 2008, Analysing changeability and time parameters due to levels of Automation in an assembly system, In FAIM, Skövde, Sweden.

[17] Fasth, Å., Frohm, J. and Stahre, J., 2007, Relations between Performers/parameters and level of automation, In IFAC workshop on manufacturing modelling, management and control Budapest, Hungary.

[18] Fasth, Å., Stahre, J. and. Dencker, K., to be submitted 2008, Measuring and analysing Levels of Automation in an assembly system, In The 41st CIRP conference on manufactoring systems, Tokyo, Japan.

[19] Bruch, J. Karltun, J., Johansson, C., and Dencker, K., A proactive assembly work setting – Information requirements, to be submitted 2008. In The 41st CIRP Conference on Manufacturing Systems Tokyo, Japan.

[20] Bruch, J., Johansson, C., Karltun, J. and Winroth, M., 2007, Considering design demands of a proactive assembly system – a position paper, In Swedish Production Symposium, Sweden.

[21] SIS/SS/624070, 2003, Competence management systems – requirement. ICS 03.120.10.

[22] Frohm, J., 2007, Levels of Automation in Production Systems, Department of Production Development, Production System, Chalmers University of Technology, Gothenburg, Sweden.

[23] Frohm, J., Lindström and Bellgran, M., 2005, A model for parallell levels of automation within manufacturing, In proceedings of the 18th International Conference on Production Research, Italy.

[24] Stahre, J., 1995, Towards Human Supervisory Control in Advanced Manufacturing Systems, Doctoral Thesis, Chalmers University of Technology Sweden.

[25] Rasmussen, J., 1983, Skills, Rules, Knowledge, Signals, Signs and Symbols and other Distinctions Human Performance Models, IEEE Transactions on Man, Systems and Cybernetics, SMC-13/3:pp. 257-266.

[26] Fitts, P. M., Ed., 1951. Human Engineering for an Effective Air Navigation and Traffic-Control System, Columbus, Ohio State University Research Foundation, US.

[27] Hollnagel, E., Woods, D., Levenson, N., et al., 2006, Resilience Engineering. Ashgate, Hamp-shire, England.

[28] Grote, G. 2004, Uncertainty management at the core of system design, Annual Reviews in Control 28, pp. 267-274.?

[29] Harlin, U., Frohm, J., Berglund, M. and Stahre J., 2006, Towards efficient automation implementation. In: Mayer, F. And stahre, J. (Eds) Proceedings of the 9th symposium IFAC on Automated System based on Human Skills and Knowledge, Nancy, France.

Virtual Manufacturing Work Systems

Peter Butala[1], Ivan Vengust[1, 2], Rok Vrabič[1], Lovro Kuščer[1]
[1] Department of Control and Manufacturing Systems, University of Ljubljana, Ljubljana, Slovenia
[2] PS d.o.o., Logatec, Slovenia

Abstract
Contemporary computer and information technologies enable digitalization and virtualization of real systems and their exploration in a virtual space. In the paper, a virtual CNC machine tool is presented. It is build up to support development of CNC controllers and control algorithms on the basis of the Hardware-in-the-Loop principle, where the object of control is substituted with a virtual one. The developed digital model of a machine tool may serve also as a model of a digital factory.
In the case study the virtual model of a desktop CNC engraving machine LAKOS 150 is described. The model is integrated in the control loop with a real CNC controller via the CNC2VML interface, which converts the controller's signal into digital information and vice versa. This enables communications between the CNC controller and the virtual machine tool model in real-time. The model is visualized in a 3D graphical environment. The applied programming techniques are generic and based on the open architecture principle and standard graphic library.

Keywords:
Digital model, CNC machine tool, Mechatronic system design, Hardware-in-the-loop, Virtualization

1 BACKGROUND AND MOTIVATION

Digitalization and virtualization of work systems open new perspectives for development and operations of complex manufacturing work systems. A virtual work system model is an effective tool for demonstration of complex work structures and their control and operational principles. The virtual model embeds explicit knowledge, which can be explored and reused. For example, users can investigate and make experiments with the virtual model by themselves and thus can gain better understanding and can learn much more naturally and effectively.

The contribution reveals development of a virtual CNC machine tool. The objective of the research is to develop a digital model (1) as a building block of a digital factory for simulation of machine operations and (2) for visualization of machine tool behavior for remote control, educational purposes and, last but not least (3) for development of a real CNC controller based on the hardware-in-the-loop principle. Several other applications of a virtual model of a machine tool are possible [1, 2].

The virtual model of a desktop CNC engraving machine LAKOS 150 is described as an example.

1.1 Digital factory

Manufacturing nowadays relies on computer controlled work systems and computer aided technologies, which are more or less integrated in a complex cybernetic structure – a factory. Development and operations of such structures open challenging engineering and managerial issues, which have to be addressed with adequate methods based on knowledge.

The digital factory represents an approach to explicit formulation of manufacturing knowledge and it's coding into software. The objective here is to efficiently support design, development and operations of a real factory. The digital factory is a set of digital models and computer aided tools for design of new manufacturing systems and planning of production for new products [3]. In the future, production of any new product will be examined through simulations before its realization in a real environment. This will significantly contribute to better decisions in the development process, increased quality of solutions, accelerated development and decreased development costs. In order to build up an integrated digital model of a real factory, digital models of individual work systems are needed.

1.2 Hardware-in-the-loop principle

A typical manufacturing work system is a mechatronic system, which is composed of a complex electro-mechanical structure, i.e. an object of control, and a controller. The controller is composed of control hardware elements and highly specialized control software.

In design and development of a new work system, the object of control as well as the controller hardware and software have to be developed concurrently and then integrated in a prototype. In early development phases, neither the object of control nor the controller elements are available. The problem that arises from this fact is that the control software has to be developed on the basis of a conceptual solution and its specification, which is as a rule incomplete and a matter of change. The software also cannot be tested on the target hardware during the development. Hence, inadequate solutions and bugs in the software are common and have to be resolved in the integration phase.

The Hardware-in-the-Loop (HiL) principle enables much more effective concurrent development of a mechatronic system. The principle is based on substitution of a real object of control in a control loop with its software model [4].

Figure 1 shows a typical HiL system structure. The object of control is substituted by its digital object. If one introduces prototype hardware, the control loop can be closed already at the very beginning of its development.

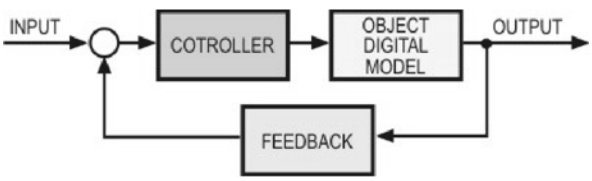


Figure 1: Block diagram of the HiL system.

Software development can start effectively on this configuration. Algorithms can be developed, tested and verified in the loop. HiL may benefit to better and faster system development. The quality of the source code can be significantly improved and final integration with real elements is straighter.

The HiL principle can be applied in development of different mechatronic systems, such as machine tools, robots, devices, storage/retrieval systems, energy systems, etc. In this paper the object of control is a computer numerically controlled (CNC) machining system.

2 DIGITAL MODEL OF A WORK SYSTEM

In order to obtain a virtual model of an entire CNC work system, several subsystem models have to be developed and integrated. According to Peklenik [5] such an elementary work system is composed of a process (e.g. cutting), a process implementation device (machine tool), and a logic controller (CNC).

For the implementation in the HiL concept where the real controller is implemented, the digital machine tool model is needed in the first step. The cutting process model is not indispensable while there is no direct interaction between the process and the controller. In the process there are interactions between a workpiece and a tool on a machine tool. Simulation and visualization of these interactions would benefit to better understanding of the entire CNC system but need a lot of efforts to be accomplished.

The machine tool digital model is composed of several sub-models as shown in Figure 2. The geometric and kinematic model simulates the geometry and kinematic movements of the machine. The dynamic model simulates dynamic responses of the machine to input signals and disturbances. The virtual reality model visualizes the model behaviors.

The machine tool digital model has to be interfaced with the real controller on its input and output sides as shown in Figure 2. The key issue here is to convert the controller signals into digital information on the input side and vice versa on the output side for feeding signals back to the controller. These transformations must be accomplished in real-time with no delay.

The digital model of a machine tool requires appropriate hardware to become a virtual model. A standard or an industrial PC is a suitable choice for the hardware. Different input/output and communication interfaces and drivers can be added for control of digital and analog signals and communication. The operation system has to support real-time (RT) operations. RT extensions of MS Windows or Linux can be implemented. For realization of digital model software different modeling environments, such as Matlab/Simulink and LabView, are available. The software can be also developed with own solutions programmed in the universal programming language C/C++.

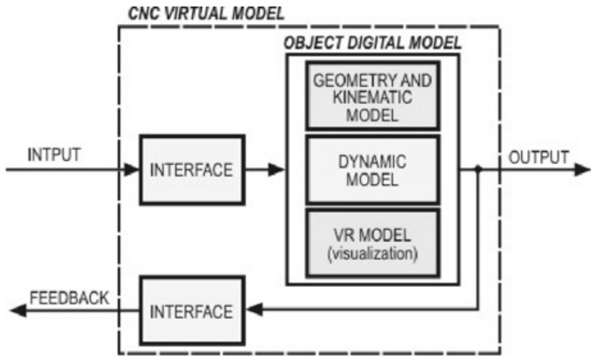


Figure 2: Block diagram of the HiL system.

3 VIRTUALIZATION OF OBJECTS

A virtual image of an object can be produced with adequate programming tools on a computer and presented on a display or screen. Several tools for visualization of virtual worlds are available today. First standardization of virtual object description provided the Open Source VRML standard (1997). Later, it was extended into X3D description. A powerful tool for computer graphics and visualization is the open source program VTK. It is based on an open source C++ object library and includes several algorithms for visualization of 3D objects. It enables stereoscopic visualization but does not enable interaction. A high-end virtual environment is the Cave VR environment, which enables interactive simulation of the virtual world. The Cave system is composed of a spatial projector, which project images on all five view surfaces around a user, and sensors indicating user motions. The images are adopted according to the position of the user. The user can explore the virtual world by walking and/or moving objects by hand. The mentioned systems are highly specialized and expensive.

For own development of digital object models, the graphic library OpenGL [6] can be used. OpenGL is a graphic standard which was introduced in 1992. The library includes building blocks for building and visualization of 3D graphical objects. It is composed of application interfaces (API), which can be implemented by different programming languages, such as C, C++, C#, FORTRAN, Ada and Java. Its main advantages are stability, reliability and ease of use. The library is permanently improved and is well documented [7].

In computer graphics, the basic element is a point, described with a unique 4-dimenzional vector named ‘quaternion’. An object is described with a set of points organized in a matrix. All transformations are performed on matrices. For example, if an object is moved, the model matrix is multiplied with a position vector in order to obtain new positions of points. A group of points is joined in a polygon, which define a surface. A 3D object is composed of several surfaces. In order to obtain the final image of an object, one can define model illumination, surface textures, edge smoothing, etc.

4 VIRTUAL MACHINE-TOOL VML 150

The principles of modeling of virtual objects are applied in the development of a virtual CNC machine tool. The virtualized object is the 3-axes machine tool LAKOS 150, which is a desktop CNC engraving machine developed mainly for educational purposes. The real instance of the machine with the CNC controller is shown in Figure 3. Three positioning drives are carried out with stepping motors and spindles. The CNC controller is implemented on a PC with Linux RT operation system. The controller software is based on the open source software EMC [8].

The structure of the HIL system with the virtual model in the loop is shown in Figure 4. The control loop consists of a real CNC controller – the same or similar as in the case of the real system LAKOS 150 - and the virtual model VML 150. Input is a NC-program in form of the standardized G-code – the same as in the case of the real system. Output here is, of course, virtual – in terms of visualization of the machine tool and its movements displayed on a screen. In this setting, the main question is how to connect the virtual model with the real controller. Electric signals generated by the controller have to be converted into corresponding information, which is feed as a reference into the virtual model. The virtual model also generates feedback information, which in turn has to be converted into corresponding electric signals and fed back into the controller in order to close the loop.

In order to solve the problem of the interconnection the CNC controller with the virtual model, an appropriate interface for the main and feedback loops is developed. Let us look in the solution of the interface.

4.1 CNC2VML interface between a real CNC controller and a machine-tool virtual model

The CNC controller generates two control signals of the 'step/dir' type for each machine axis drive. The 'step' signal defines the number and the frequency of steps to be performed by a stepping motor and the 'dir' signal defines direction of rotation of the stepping motor. Direct implementation of these signals in the MS Windows XP environment without modules for real-time control is not possible. Therefore, the CNC2VML interface for signal processing is developed and implemented on the microcontroller μRD2, which is based on the Atmel microprocessor AT89C51RD2 [9]. The role of the interface is to convert input signals into digital information which is fed to the PC where the machine tool digital model is running over the RS232 serial communication. Feedback information is realized by virtual limit and home-position switches implemented in the VML 150. When triggered, switch state (0/1) information is fed to the microcontroller, which converts it into an electrical signal for the CNC controller.

Figure 3: CNC engraving machine LAKOS 150.

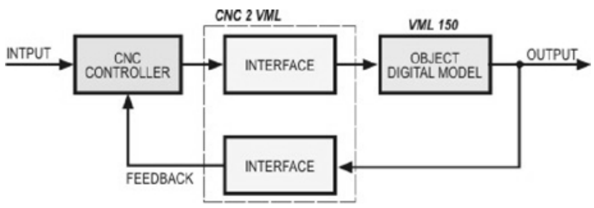


Figure 4: Interconnection of a CNC controller with a virtual machine tool model.

The 'step' signals are of pulse type with frequency up to 10 kHz. The CNC2VML interface has to count each pulse and to register each direction change for three machine drives simultaneously. Any missing signal would result in inadequate functioning of the virtual model. This fact sets high requirements for signal processing in CNC2VML. The CNC2VML software is based on an interrupt routine. The controller signals are connected over a keyboard interface with eight input pins. Six of them are used for 'step/dir' signals for the machine drives and are treated as independent interrupt sources. Each state change of any signal triggers the interrupt routine, which registers the pulse in an interrupt vector. The interrupt source is than recognized and interpreted in by the CNC2VML software.

The CNC2VML communicates over the standard RS232 serial interface with the PC platform where the VML 150 model is running. This is a bi-directional communication which enables sending of interpreted signals data from the interface to the VML and data describing logical states (0/1) of the VML switches in the opposite direction.

4.2 Digital machine-tool model VML 150

The digital model VML 150 is computer software, which represents the virtual version of the LAKOS 150 machine tool. It interprets control signals in a corresponding state and/or movement of the machine drives and visualizes them on-line in a corresponding image displayed on a screen.

The VML 150 software is based on the 'C# OpenGL Framework', which is a graphic engine for the OpenGL library. The .net programming environment, the universal programming tool C# and the standardized graphic library OpenGL are the key enabler for efficient and effective model development due to their availability and performance.

The program manipulates geometrical object and communicates with the CNC2VML interface in order to provide visualization of machine movements according to the input signals and dynamic image refreshing.

An important part of the VML 150 software is a graphical user interface. It is composed of two windows, as shown in Figure 5. The left OpenGLControl window displays the 3D model image as a composition of geometrical objects which perform relative motions according to the reference commands form the CNC controller. The OpenGL commands enable turning and zooming of the model so that the model can be observed from different viewpoints. The right Control Window displays current positions of the machine-tool axes and enables user interactions.

The final result is the integration of the VML 150 virtual model with the real CNC controller over the CNC2VML interface in the hybrid hardware-in-the-loop system, as shown in Figure 6. The system elements are interconnected according to the HiL conceptual scheme (Figure 4).

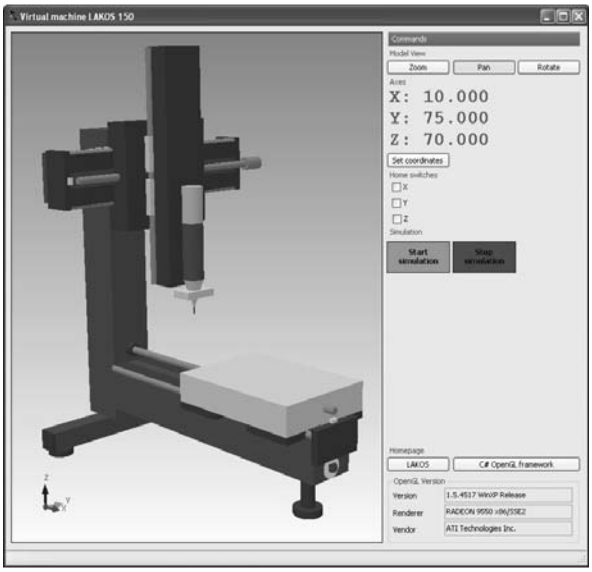


Figure 5: VML 150 graphic user interface.

The system can be operated on the basis of manual commands set directly on the CNC controller or by any NC-program in the G-code serving as a reference for the CNC controller. The VML 150 model exhibits the same kinematic behaviour as it would be realized by a real machine-tool. Thus, the full functionality of the hybrid real/virtual HIL system is achieved.

5 CONCLUSION

Digitalization and virtualization set new frontiers in design, simulation and testing of complex manufacturing work systems. Digital work systems models are building blocks of future digital factories, which will serve as test-beds for investigations of design solutions in different design and operation stages of the next generation manufacturing systems. The digital models also represent knowledge coded in software, which can easily be re-used, adapted, upgraded or recycled.

The presented example describes development of a virtual digital model where the virtual instance of the existing real machine-tool LAKOS 150 is realized. The objective of the model is to support development of CNC controllers with a virtual object of control in the loop. Thus, the development can be accelerated and improved while the entire loop is available already in the early design phases. The virtual model also enables testing of NC-programs. Besides, visualization of the virtual model enables better understanding of machine tool behavior, and surveillance and diagnostics of remote operations, etc.

The applied approach is generic, while it is based on an open architecture and standard or open source programming tools. This makes the solution independent from commercial applications. The proposed concept and solutions enable development of digital models of other manufacturing work systems.

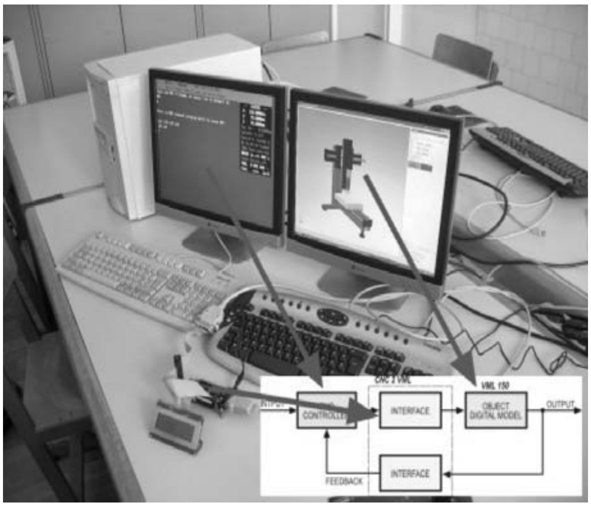


Figure 6: Integrated hardware-in-the-loop system in operation.

6 ACKNOWLEGEMENT

This work was partially supported by the Slovene Ministry of Higher Education, Science and Technology, Grant No. 400-17/2004 and inspired by the EU FP 6 project NOE VRL KCiP, Contract No. FP6-507487-2.

7 REFERENCES

[1] Altintas Y., Brecher C., Weck M., Witt S., 2005, Virtual Machine Tool; Annals of the CIRP Vol. 54/2

[2] Erkorkmaz K., Altintas Y., Yeung C.-H., 2006, Virtual Computer Numerical Control System; Annals of the CIRP Vol. 55/1, 399-402

[3] Bracht U., Masurat T., 2005, The Digital Factory between vision and reality, Computers in Industry, Vol. 56, 325-333

[4] Brendecke, T., & Kucukay, F., 2002, Virtual real-time environment for automatic-transmission control units in the form of hardware-in-the-loop. International Journal of Vehicle Design, 28, 84–102.

[5] Peklenik J., 1998, Fertigungskybernetik – eine neue wissenschaftliche Disziplin für die Produktionstechnik, Sonderdruck der TU Berlin, Berlin, 1-25.

[6] OpenGL Library Documentation: http://www.opengl.org/documentation/

[7] C# OpenGL Framework www.csharpopenglframework.com

[8] EMC: http://www.linuxcnc.org/

[9] AT89C51RD2 Hardware Manual: http://www.atmel.com/dyn/resources/prod_documents/doc4235.pdf

Virtual Enterprise Model for Integrated Planning in Manufacturing Supply Network

Jayeola Femi Opadiji[1], and Toshiya Kaihara[2]
[1] Graduate School of Science and Technology, Kobe University, Japan
[2] Graduate School of Engineering, Kobe University, Japan

Abstract
Value creation in manufacturing supply network often involves complex interactions among autonomous enterprises. Consistency is required in decisions made in all the planning stages if enterprises in such a supply network are to continuously have feasible and flexible plans which respond to volatile trading environments. We propose a virtual enterprise model based that integrates the strategic and tactical planning stages of materials procurement activity in a manufacturing supply network. In this model, each enterprise is represented as an agent participating in economic activities in multiple virtual markets. We employ a market-based approach to find satisficing tactical plan for in a supply network given strategic constraints.

Keywords:
Virtual Enterprise, Supply Network Planning, Resource Distribution

1 INTRODUCTION

One of the major decisions enterprises have to make as touching their supply network is that of resource procurement. Since all the enterprise units that make up any supply network are autonomous, legally independent entities, they make such resource procurement decisions with the aim of achieving their primary objective of market competitiveness [1]. As in Figure 1, planning in a manufacturing supply network is done with respect to three criteria – product types, network structure and time. The product type axis indicates planning with respect to the manufacture of all feasible products in the supply network, the network structure axis refers to the independent manufacturing units involved in value transformation process and the topology of the network, while the time axis represents the planning period under consideration. Our objective in this work is to present a virtual enterprise [2] model that solves tactical level resource procurement problems subject to strategic level planning constraints.

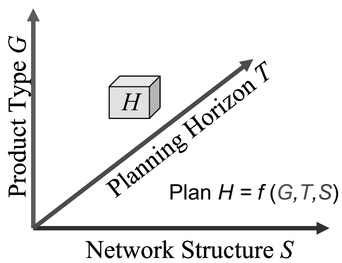


Figure 1: Supply Network Planning Criteria

2 INTEGRATED SUPPLY NETWORK PLANNING

The idea of enterprise units integrating their plans across planning levels stems from the need to prevent infeasibilities at lower planning levels as a result of constraint at the upper levels. This is also necessary to monitor supply network performance at the lower levels with respect to upper level plans. Figure 2 gives an example of inter-level plan integration of an enterprise between the strategic and the tactical levels.

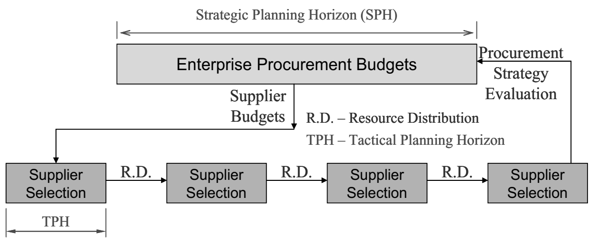


Figure 2: Integration of Strategic and Tactical Plans

From Figure 2, the tactical level problem to be solved is the resource distribution problem across the supply network. As discussed in the previous section, there are three problems to be solved simultaneously: (a) Determination of product types and quantities, (b) Selection of enterprise units involved during that planning time span and (c) The planning span under consideration vis-à-vis supply network demand environment. These three problems spell out the decision variables in an enterprise objective function but the constraints are provided by the enterprise procurement budget in the strategic plan. By enterprise procurement budget, we refer to projected input resource order quantities from every supplier planned for by an enterprise over the strategic time span. Also given as constraints from the strategic level to the tactical level, are the expected production overheads resulting from securing inputs from all the different suppliers. Considering a generalized supplier network of the form shown in Figure 3 with input complementarities and resource contention, we describe the network in the form of a task dependency graph as discussed in [3].

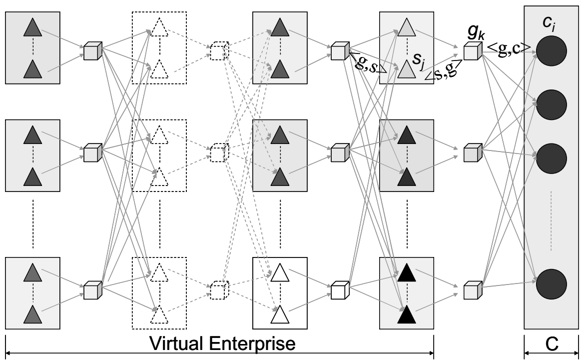


Figure 3: Supply Network with input complementarities

$H:(N,A)$

$N = T \bigcup G = \text{nodes in graph H}; \quad T{:}\{S,C\} = \text{traders}; G = \text{goods}$

$A:\{a_i = \langle g,t \rangle_i \, or \langle t,g \rangle_i \mid t \in T; g \in G\} = \text{set of directed arcs in graph H}$

$\langle g,s \rangle = \text{resource g is an input to enterprise s}$

$\langle g,c \rangle = \text{resource g is consumed by consumer c}$

$\langle s,g \rangle = \text{resource g is produced by enterprise s}$

The tactical level objective is thus defined as:

$$u((N^*, A^*)) = \max_{(N',A') \in (N,A)} (u((N',A')) \mid (N',A') \text{ is feasible}) \tag{1}$$

given that

$$u((N',A')) = \sum_{c \in C} u_c((N',A')) - \sum_{s \in S} \pi_s((N',A')) \tag{2}$$

$u = \text{value accrued to an enterprise}$

$\pi_s = \text{production costs of supplier s}$

The constraints given in the strategic plan are:

$$I = \sum_{u=1}^{l} \sum_{v=1}^{m_u} \sum_{w=1}^{a_v} \sum_{j=1}^{h} p_j f_{uvw}^j \tag{3}$$

$$p_w g_w \ge \sum_{j=1}^{h} p_j f_w^j \ge \sum_{k=1}^{m_{u+1}} p_k g_k^w \quad ; \; \forall \quad w \in M_u; \; 1 \le u \le l-1 \tag{4}$$

$$g_{u11}^w = \sum_{\theta=1}^{TPH} g_{u11}^w(t_\theta) \tag{5}$$

$R_w : \{r_{u1}^w, r_{u2}^w, r_{u3}^w, \ldots, r_{um_u}^w\}$

$r_{u1}^w : \{g_{u11}^w, g_{u12}^w, \ldots, g_{u1n}^w\}$

$I = \text{estimated total investment into the supply network}$

$h = \text{number of primary production resource}$

$l = \text{number of production layers in supply network} \mid L \mid$

$m_u = \text{number of resource markets in production layer } u$

$a_v = \text{population of enterprise agents in market } v \in M_u$

$p_j = \text{price of resource } j$

$f^j = \text{quantity of primary resource } j$

$g_w = \text{quantity of market resource } w$

$R_w = \text{set of maximum budgetary allocations of enterprise } w \text{ for input resources}$

$r_{u1}^w = \text{procurement budget estimate of enterprise } w \text{ for resource } u1$

$g_{u1x}^w = \text{quantity estimate of resource } u1 \text{ expected from supplier } x \text{ over SPH}$

$g_{u1x}^w(t_\theta) = \text{quantity estimate of resource } u1 \text{ expected from supplier } x \text{ within period } t_\theta$

The objective function of equation (1) maximizes the value accrued to enterprise agents from feasible allocation of resources. Equation (3) is the total expected investment into the supply network in the form of capital and labour over the strategic time span, while (4) suggests that the total income received by an enterprise agent from the sale of its product must be at least the amount of budgeted investment which in turn must not be less than the expenditure of the enterprise unit. Equation (5) is the total amount of budgeted input from a supplier over the tactical planning horizon.

With the complexity of a supply network like the one in Figure 3, finding an optimal solution may not be practical with the complementarities involved in network products. We therefore model the network as a single virtual enterprise consisting of multiple virtual markets. This way, we can implement an interaction protocol that will obtain a satisficing solution to the resource distribution problem. We describe the virtual enterprise model in the next section.

3 VIRTUAL ENTERPRISE MODEL

In modelling the supply network as a single virtual enterprise, we a take a cue from current trends in electronic commerce where economic activities are engaged in on a virtual platform such as the World Wide Web [4]. These virtual market places obtain solutions to profit maximization problems for market products by proving an auction where traders update their buy and sell bids and the auction clears at the point where no trader is willing to change their bids. Using this same approach, we assume the entire network excluding the consumer layer is a single virtual enterprise consisting of many virtual markets in each layer. The enterprises in the supply network are deemed to be trading agents which participate in buy and sell activities in the virtual markets in order to locally maximize their own profits. The consumer layer on the other hand, is made up of buying agents who which to maximize the value they derive from market products. They hold the total investment of monetary resource into the market in the form of endowments.

3.1 Trading Mechanism

For markets with no resource complementarity, the existence of Pareto-equilibrium allocation of market resources has been proven [3][5], however, when an enterprise requires complementary input resources, the existence of such an equilibrium cannot be guaranteed. Our approach to solving this problem is using (g+1)st auction [6], where enterprise agents try to secure the inputs required to produce the highest possible units of output products. The fact that enterprise agents may have more than one possible supplier for an input type increases the complexity of the already difficult combinatorial problem. We therefore introduce the idea of average supplier overhead per unit input as known variables for each enterprise agent. These values help them make decisions on which supplier and what quantity to request. The Trading mechanism we have chosen is based on simultaneous ascending price adjustment [3] in which all consumer agents and agents in the virtual enterprise are only allowed to review their bids upward as listed below:

- Step 1: Initialize all trading agents and virtual markets
- Step 2: Consumer agents send bids at current market price (Adjust bid if not winning)
- Step 3: Enterprise Agents inspect number of winning sales bid

- Step 4: Enterprise Agents check if there is enough inputs to meet winning sales bid (if not, adjust procurement bid upward and increment price for sales bid
- Step 5: Auctions compute new market price for all resources and posts bid results privately
- Step 6: If no bid revision for all agents auction clears else go to step 2
- Step 7: Terminate Auction

The termination condition is such that all enterprise agents are able to secure at least all their inputs. This means, there is a possibility of excess supply in the market. Next, we define the enterprise and consumer agents with their bidding tactics.

3.2 Consumer Agent

$$p_c(g) = p'(g) + \alpha_c \quad if \quad p(g) < p'(g) \tag{6}$$

for

$$g = \arg\max_{g \in G}(e_c^g - p_c(g)) \ s.t. \ (e_c^g - p_c(g)) \geq 0 \tag{7}$$

$p_c(g) =$ new bid price of consumer c for resource g
$p'(g) =$ current market price for resource g
$g =$ bid quantity of resource g
$\alpha_c =$ bid adgustment constant for consumer c
$e_c^g =$ valuation of consumer c for market resource g

Equation (6) is the price bidding tactic for the consumer agent. A consumer agent adjusts its bid price by a predetermined value α_c if its last bid price is not enough to make it win all the quantity of that input. It therefore bids above the current market price for that input. Equation (7) represents the quantity of an input a consumer agent will bid for at its current bid value. It bids such that it can get as much units as possible at the current bid price subject to its total valuation for that input.

3.3 Enterprise Agent

$$p^{i+1}(g_w) = \begin{cases} \max((p^i(g_w) + \beta_w), (\sum_{k=1}^{m_{l-1}} p_w^{i+1}(g_k^w) + c_w^k)); & \text{if } p_w^i(g_k^w) < p'(g_k) \\ p'(g_w) & \text{otherwise} \end{cases} \tag{8}$$

$$p_w^{i+1}(g_k^w) = \max(p'(g_k), p_w^i(g_k^w) + \alpha_w) \tag{9}$$

for

$$g_w = \arg\max_{g_w \in G_L \subset G}(\sum_{j=1}^{m_{l+1}} p_j(g_w^j)) \tag{10}$$

$$g_k^w = \arg\max_{g_k \in G_{L-1} \subset G}((p_w(g_k^w) + c_w^k)) \quad \forall \ k = 1,2..m_{l-1} \tag{11}$$

s.t.

$$(\sum_{k=1}^{m_{l-1}} p_w(g_k) + c_o^k) \leq p'(g_w) \tag{12}$$

$$g_k^w \leq \max(g_k^w(TPH)) \tag{13}$$

$$g_w \leq \max(g_w(TPH)) \tag{14}$$

$p^{i+1}(g_w) =$ new bid price for output resource g_w of enterprise agent w
$\beta_w =$ sales bid price adjustment constant of enterprise agent w
$c_w^k =$ overhead cost of procuring resource g_k for enterprise agent w
$p_w^{i+1}(g_k^w) =$ new bid price of enterprise agent w for input resource k
$\alpha_w =$ input bid adgustment constant for enterprise agent w
$g_k^w =$ bid quantity of enterprise agent w for market resource g_k
$\max(g_k^w(TPH)) =$ procurement estimate for market g_k over TPH
$g_w =$ bid quantity of enterprise agent w for output resource g_k
$m_{l-1} =$ total number of markets in input layer

Equation (8) is the price bidding function of the enterprise agent for its product (selling price). It updates this price whenever there is a change in the price of any of its inputs. The price bid for inputs is done in much the same way as in the case of a consumer agent as shown in equation (9). Equation (10) is the output quantity bid function. Equation (11) is the quantity bid function for inputs and is determined by the number of units the enterprise agent is willing to sell at that point in time. The equation shows how an enterprise agent selects the suppliers of an input by considering the allocation that will minimize its average overhead cost, i.e. the most input at the cheapest cost. The constraint of equation (12) is the non-negative profit constraint while equation (13) is the input budget constraint imposed by the strategic plan and equation (14) is the output capacity constraint also imposed by the strategic plan of the enterprise.

4 VIRTUAL ENTERPRISE SIMULATION

4.1 Hypothetical Network

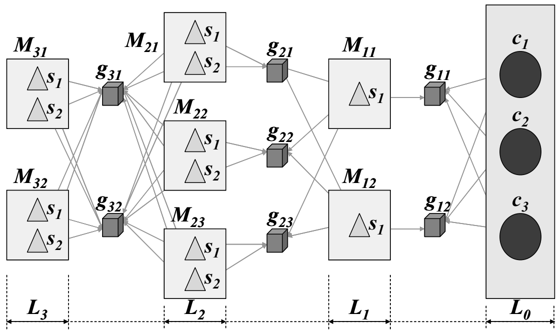


Figure 4: Hypothetical Supply Network

$$|L| = 4$$

$$G: \{G_1, G_2, G_3\}; \ |G_1| = 2, |G_2| = 3, |G_3| = 2$$

$$M: \begin{pmatrix} M_{11} & M_{12} & 0 \\ M_{21} & M_{22} & M_{23} \\ M_{31} & M_{32} & 0 \\ 0 & 0 & 0 \end{pmatrix}; \quad \begin{matrix} |M_{31}| = 2, |M_{32}| = 2, |M_{21}| = 2, |M_{22}| = 2 \\ |M_{23}| = 2, |M_{11}| = 1, |M_{12}| = 1; \ C: \{c_1, c_2, c_3\} \end{matrix}$$

For simplicity, we assume an input ratio of unity for all enterprise agents, i.e. a unit of every input is necessary to produce a unit of output.

4.2 Simulation Results and Discussion

First, we present results on how product supply changes for all enterprise agents in Figure 5.

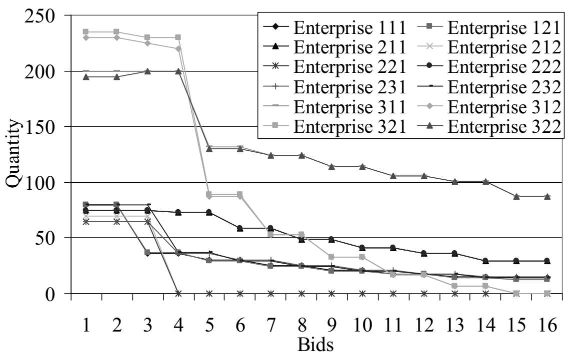


Figure 5: Enterprise Agent Supply Bids

The shape of the supply plots illustrate a trend where enterprise agents are forced to respond to decreases in product demand induced by price increases in the virtual markets. Initially, all the enterprises are willing to sell all quantities equal to their output capacity but this changes as prices force consumers to demand less of the final products of the virtual enterprise.

Next, we show in Figure 6, the price change graph for virtual markets in the final layer (layer L_1) of the virtual enterprise. This illustrates the price change trend in all the other layers of the virtual enterprise.

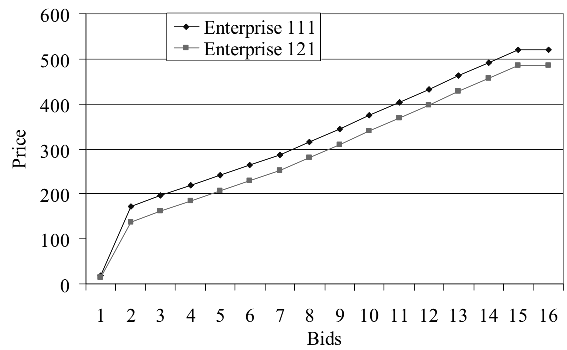


Figure 6: Enterprise Agent Output Price Bids

The shapes of the price plots for each of the enterprise agents are determined by equation (8). The incremental pattern follows the ascending nature of current market price computation after every iteration.

The last set of results presented is the resource distribution graphs in Figure 7(a-b). Figure 7(a) shows the demand for market resources by all the agents in the system, while Figure 7(b) reveals the total supply by all the enterprise agents in the virtual enterprise.

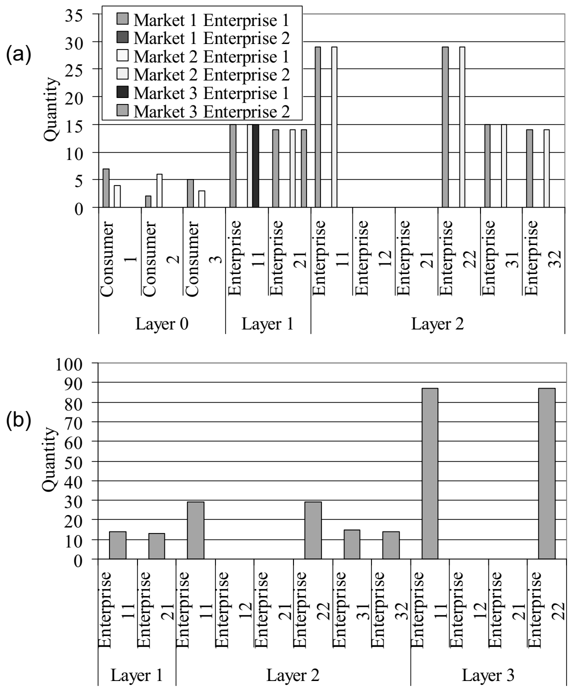


Figure 7: Trading Agents' Resource Distribution

Figure 7 shows the consumption and production profile of each of the agents in the supply network. This demand and supply profile is a reflection of the competitiveness of enterprise agents in a particular tactical planning horizon.

It was observed that the choice of suppliers is affected by the values of the buying price adjustment constant, α and the selling price adjustment constant, β. The lower these values, the higher the chances of an enterprise agent winning bids in its market. Another factor that affects the choice of supplier is the relative overhead cost of suppliers in a market as seen by a buying enterprise agent.

An enterprise agent which looses out in a preceding planning horizon can adjust its bidding strategies to make it more competitive in the next. Given that the supply links have been formed in a planning span, all enterprise agents update their strategic constraints for the next planning span. This means that if an agent has exhausted all its budgeted input from a particular supplier, it cannot send bids to that supplier again in the next planning period because this will violate the strategic plan.

5 CONCLUSION

The market-based approach to integrated planning of manufacturing supply networks makes it possible to simulate *reality* in the sense that the autonomy and distribution of enterprises that make up the network is preserved. Although the virtual enterprise model proposed does not guarantee arriving at an optimal solution, this model provides a platform for satisficing resource distribution and supplier selection at the tactical level.

For an enterprise in the network, its performance at the tactical level spanning the strategic planning horizon will aid it in detecting areas of improvement in future strategic plans as it relates with resource procurement.

One area that will definitely require further research is how agents can learn such that they can automatically update their bid adjustment factors to increase competitiveness and what will be the effects of agent learning on the overall performance of the manufacturing supply network.

6 REFERENCES

[1] Porter, M.E., 1998. Competitive Strategy, Techniques for Analyzing Industries and Competitors, Free Press, New York.

[2] Camarinha-Matos, L.M., 1999. The Virtual Enterprise Concept. Infrastructures for Virtual Enterprises. Dordrecht: Kluwer. pp. 3 – 14.

[3] Walsh, W.E. and Wellman, M.P., 2003. Decentralized Supply Chain Formation: A Market Protocol and Competitive Equilibrium Analysis, Journal of Artificial Intelligence Research 19, pp. 513 – 567.

[4] Lucas, H.C., 2002. Strategies for electronic Commerce and the internet, MIT Press, Cambridge.

[5] Scarf, H., 1969. An Example of an Algorithm for Calculating General Equilibrium Prices. Cowles Foundation Discussion Paper No. 276. Cowles Foundation for Research in Economics at Yale University.

[6] Babaioff, M. and Nisan, N., 2004. Concurrent Auctions Across the Supply Chain. Journal of Artificial Intelligence Research, 21, pp. 595 – 629.

Manufacturing system to support design concept and reuse of manufacturing experience

Andersson Petter[1,2], Isaksson Ola[1,2]

[1] Department for Product Development and Methods Improvement, Volvo Aero, Trollhättan, Sweden

[2] Functional Product Development, Applied Physics and Mechanical Engineering, Luleå University of Technology, Sweden

Abstract

Life cycle responsibility for manufacturing companies increases the service content coupled to the product. One consequence is that transferring knowledge gained from all life cycle phases has an even more decisive impact on the definition of the product concept, here referred to as the functional product.

The paper focuses on transferring experiences from the manufacturing phase and how to account for these in the design phase. Based on an empirical study at two companies, an automotive and one aeronautical company, current practices were identified. Manufacturing experiences are captured and managed in a manufacturing context whereas the use of experience in the design phase is discussed. Finally a generic approach to support the use life cycle experiences in earlier phases of product development is suggested, where the design and manufacturing case serves as an example.

Keywords:

Product Development; Manufacturing Experience; Manufacturing; knowledge sharing; Engineering design

1 INTRODUCTION

Life Cycle Responsibility increases amongst manufacturers today. The industry is challenged to understand different life cycle phase's impact on the product, e.g. manufacturing, operations/usage, disposal etc.

These factors are "product" focused whereas life cycle dependency implies accompanying services, such as maintenance and repair, customer training etc. The term used for products, having a service contend, is called "Functional Product" [1], whereas Product Service Systems is used in a similar manner. The service integration in manufacturing challenges the competences, roles and responsibilities of manufacturing companies [2].

Consequently the emphasis on information and knowledge increases for manufacturing companies as the use and re use of experiences from various life cycle phases' increase [3]. Meanwhile, the impact that up stream processes such as the design phase has on the robustness and efficiency of manufacturing is well known.

A structured reuse of manufacturing experience involves incorporating learning from current or previous products in the design process in order to avoid recurrence of manufacturing issues on new products.

In the present study we have investigated how experiences gained in the manufacturing phase can be identified, adopted and eventually used in a designer's context. This serves as a relevant example of where experiences are used from one life cycle to another.

The challenge to manage experiences and learning within manufacturing obviously cover a broad range of issues, where knowledge takes on many different forms. Even so the competitive power of succeeding in managing experiences in the organization is a strong motivator to continuously improve the experience management process [4]

In particular the feedback from manufacturing to tailor engineering design systems accounting for manufacturing experiences have also been discussed by Brissaud [5]. They point out that the different context of the engineering designer verses the manufacturing context is missed out due to that experience management systems are often defined from a manufacturing context.

One approach is to take the viewpoint of the engineering designer, where the engineer's context is enriched by integrating information from later life cycle phases. Boart et.al. [6] have shown that using the functional product development approach manufacturing process alternatives can be used as design parameters in early stages of product development. They argue that both the capability to quantitatively assess impact of varying Manufacturing Design Parameters, and the availability of these Design methods are needed to succeed as an early phase design method.

In the Design engineers' toolbox, the CAE system plays an important role. The CAE environment has become a center point for the product modeling and much focus is set on the master model concept. Not only is the CAE used for geometric modeling (CAD), but for actually modeling the virtual product. As an example, template modules of parameterized CAD files are used to provide the design engineer with predefined blocks where rules are embedded in the parametric constraints [7].

At the same time as design methods focus on a master definition, the engineering environment gets increasingly heterogeneous with a dispersed set of data sources.

Baily et al, [8] describes an "An Intelligent System for the Optimal Design of Highly Engineered Products", where Knowledge Based Engineering is fused with product Control Structure, Conventional Master Model and Linked Model Environment to collectively render an Intelligent Master Model. This system provides multi-disciplinary design optimization in a web based environment for global collaboration.

In Europe, a collaborative platform for multi-partners and multi-engineering was developed in the European founded 6th framework project VIVACE (Value Improvement through a Virtual Aeronautical Collaborative Enterprise), [9].

2 CURRENT PRACTICES FOR CAPTURING AND USING EXPERIENCE FROM MANUFACTURING

A study was conducted at two companies, one in aerospace and one in the automotive industry with the aim to understand the current practices for capture and reuse of experience, i.e. engineering knowledge, in manufacturing. The study was defined to cover four product development phases; concept, detailed design, manufacturing preparation and serial production in three different organizational disciplines; design engineering, manufacturing engineering and manufacturing operations, according to Figure 1.

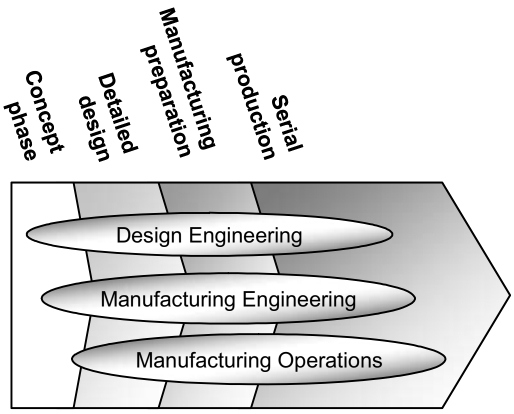


Figure 1. Four PD phases and organizational disciplines.

Questionnaires were used including, one department from each discipline, giving approximately 180 forms to analyze. The questionnaires where performed prior to the interview and both the questions and the preliminary result from the survey was used as a basis for discussions in the interviews.

A report [10] from this study points out that it is common with recurring problems, although the frequency of them is perceived quite differently among the respondents. The perceived involvement where significantly different between the design engineers and manufacturing engineers in early phases, where the manufacturing engineers indicated a much lower level of collaboration.

It was also noted that manufacturing experience from earlier projects is usually made available through the composition of new design teams where competence from the manufacturing disciplines is included. Even so, 90% of the respondents believed there will be less recurrent manufacturing issues if collaboration between manufacturing and design increased. The usage of experience databases where also investigated and it was found that as the amount of information grows, the design engineers prioritize other engineering tasks and are reluctant to follow the procedure to go trough the sources of manufacturing experience.

3 CHALLENGES FOR REUSE OF MANUFACTURING KNOWLEDGE IN ENGINEERING DESIGN

From a design engineering point of view the experiences as perceived, captured and partially logged/documented is typically "atomized", i.e. found in the explicit manufacturing context. These "Elements of manufacturing experience" (EME) are different in character and format, e.g. experience related to manufacturing is; Problem reports, statistical information, list of operations, product structure as well as reports of experience from projects that are stored in Lessons learned databases.

3.1 Heterogeneous environment

The information is stored in different vaults and is usually accessed through special tools. Consequently EME exist in a heterogonous environment, See Figure 2.

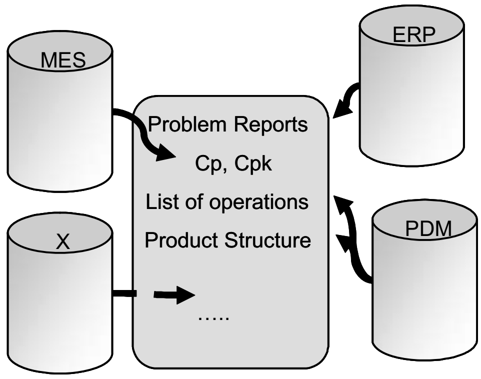


Figure 2. Heterogeneous environment

Manufacturing Execution System

The Manufacturing Execution System is a set of integrated functions which provides an infrastructure and a production management system. One of these functions is to collect statistical outcome from the production e.g. Cp, Cpk, etc. This data is used to follow up manufacturing requirements to ensure a robust manufacturing process.

Enterprise Resource Planning

In the companies Enterprise Resource Planning (ERP) system, various data and processes of an organization is integrated into a unified system. Examples of modules in an ERP system are, Financials, Projects, Human Resources, Customer Relationship Management, Supply Chain Management and Manufacturing, where the latter provides information about Manufacturing Process, Manufacturing Flow, Quality reports, etc. Consequently, the ERP system can provide a large amount of manufacturing experience.

Product Data Management

The PDM environment is usually tightly integrated with the CAD system for the management of product data related to the geometry definition. This system is also providing the link between product definition and manufacturing engineering task, such as lists of operations sequences and NC programs.

3.2 Design context verses manufacturing context

Different character of more or less isolated data elements that is stored from a manufacturing context/view. This problem has been approached in Data Mining [11] where intelligent tools for extracting useful information and knowledge has been developed but the context of usage in a designer's context remain. As mentioned in chapter 2, experience Databases tend to be large and often difficult to grasp.

Although a product development project is working with the same goal, to produce the best product possible, it is a natural tendency in larger organizations to experience a distance between people in different organizations. This gap is not only manifested in human to human communication, but is also apparent in the surrounding system environment, See Figure 3.

As an example, the design engineers work in a CAE environment that provides full access to component structure and the master definition. Although the manufacturing engineers work in the same system, they have a different view and limited access of the product structure, as there role is to grab an existing component definition and create a list of operations to be executed on the shop floor.

From the other side, manufacturing operations has a set of tools, generally referred to as the Manufacturing Execution System (MES), providing an interface between the manufacturing engineers and the operator of a machine.

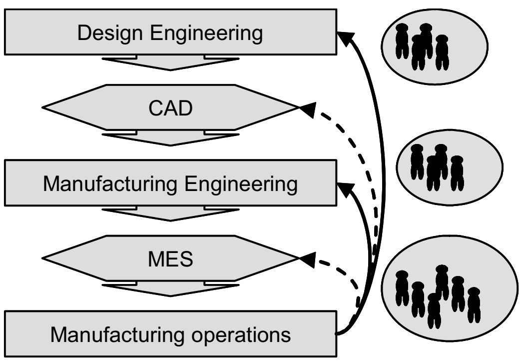


Figure 3. Reuse of manufacturing experience.

The study revealed that although systems for capturing manufacturing experience existed within the manufacturing organization, the knowledge of its existence or how to access the information was not common knowledge among design engineers. In Figure 3, the experience feedback loops from manufacturing operations are also visualized, both the explicit type with a system integration shown with dotted lines as well as the implicit type with a human to human transfer.

1. The shortest feedback loop goes from manufacturing operations back to the production system (MES) and can be a fully automated process where NC programs are adjusted based on sensor signals integrated in the machine. Experiences here are quite close to data patterns, and local in character. The context is far from the designer's context.

2. The feedback of information from manufacturing operations back to manufacturing engineering effects decisions regarding production flow, tools and machines. The manufacturing engineer has a central role in managing experiences in this phase.

3. Knowledge about manufacturing impact of design decisions made by the design engineer has an ever greater impact on the PD life-cycle and therefore a possible greater impact on product cost.

4. Manufacturing feedback to the CAD environment is still limited and usually a process of updating embedded rules. If successful, the embedded rules directly in the design tools can be quite powerful whereas the process of doing so may be sensitive and difficult to keep updated.

4 TOWARDS A DESIGN SYSTEM TAILORED TO MAKE USE OF EXPERIENCE

It is a necessity to understand the view of the receiver in the feedback loop and the engineering environment that surrounds him. How does the "element of experience" on the atomic level relate to his view? In more detailed example, how do we make the design engineer understand the meaning of statistical data presented from an individual milling operation? The result could be highly dependent on previous operations and the status of that machine at that particular time. To what type of geometry topology is data related to? What project?

To answer these questions the design engineer needs to have a clear view of how the EME relates in the context of engineering design. Figure 4 describes the feedback loop in a design to manufacturing context where an element of manufacturing experience, in this case a statistic report of characteristics such as Cp and Cpk are presented in;

a) The context of component structure

b) The associated manufacturing process

c) The process activity, the milling operation.

In the same context, a problem report is presented for a drilling operation, prior to the milling.

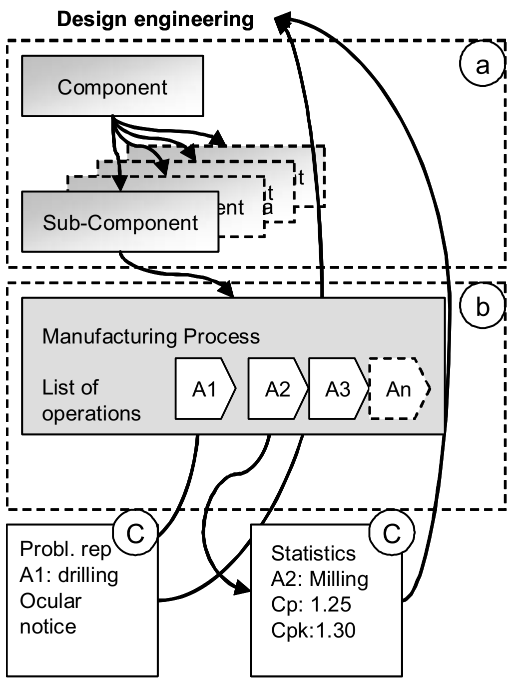


Figure 4. EME in a component and process context.

The consequence of this approach has many dimensions as it relate to several different business systems. It highlights important issues such as setting requirements on transparent interface protocols, neutral formats, etc.

Requirements on a design system that integrates experience use

1. Need to interactively search, find, retrieve and integrate experience related information from several different sources
2. Need to keep the experiences up to date close to real time
3. Need to build on the designer's context and expand functionality rather than building a completely new tool.

5 CONCLUSION AND DISCUSSION

It is noted that the Functional Product approach clarifies the principal need to transfer knowledge and experiences between different domains, illustrated in Figure 5.

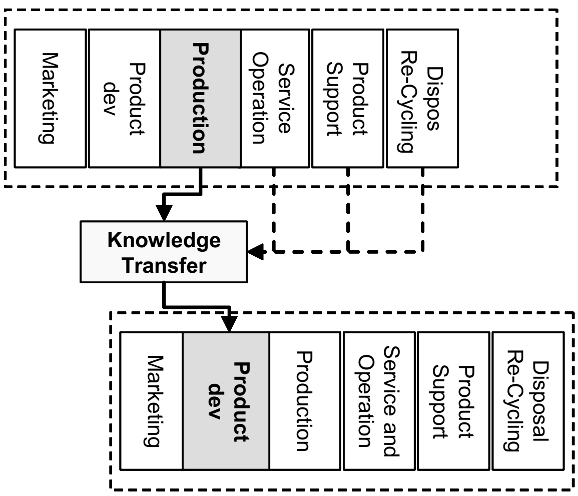


Figure 5. Knowledge transfer to new projects

If the traditional focus has been to define a product based, mainly on a functional requirements perspective - a Functional Product perspective highlights the need to account for knowledge from all life cycle phases. The contextual challenge for design teams increases further, and making experiences available for a designer is a challenge. The situation in this paper has focused on the manufacturing process, but the challenge is universal and the argument is that the contextual diversity increases as life cycle dimensions are introduced in the product concept.

It is also noted that there is a demand for more manufacturing capability information in the concept phase, both in order to predict cost and to avoid recurrence of manufacturing issues.

Finally, to achieve an effective reuse of manufacturing experience for the designer engineer it is important to provide the feedback in the design environment, giving the design engineer access to information in a context he can understand.

6 REFERENCES

[1] Alonso-Rasgado, T., Graham, T., Elfström, T., 2004, Design of functional (total care) products, Journal of Engineering Design, 15/6:515-540.

[2] Tan, A. R., McAloone, T. C., Gall, C., 2007, Product/Service-system development - An explorative case study in a manufacturing company, International conference on engineering design, 334.

[3] MANUFUTURE A vision for 2020, A report of the High-Level Group November 2004, ISBN 92-894-8322-9 (http://europa.eu.int).

[4] Siemieniuch, C. E., Sinclair, M. A., 1999, Organizational aspects of knowledge lifecycle management in manufacturing, International Journal of Human-Computer Studies, 51:517-547.

[5] Brissaud, D., Tichkiewitch, S., 2000, Innovation and manufacturability analysis in an integrated design context, Computers in Industry, 43:111-121.

[6] Boart, P., Isaksson, I., 2006, Enabling variation of manufacturing process parameters in early stages of product development, ASME, IMECE2006-14459.

[7] Hoffman, C. M., Joan-Arinyo, R., 1998, CAD and the product master model, Computer-Aided Design, 30/11:905–918.

[8] Bailey, M. W., 2001, FIPER: An Intelligent System for the Optimal Design of Highly Engineered Products, Performance Metrics for Intelligent Systems, part. II sect. 8.1.

[9] VIVACE project: (www.vivaceprojects.org)

[10] Andersson, P., Wolgast, A., 2008, Current industrial practices for re-use of manufacturing experience in a multidisciplinary design perspective, International design conference - DESIGN 2008 (Submitted for publication)

[11] Wang, K., 2007, Applying data mining to manufacturing: the nature and implications, Journal of Intelligent Manufacturing, 18:487-495.

Implementation and Design System

Globalization-Compatible Product Structure Based on Technological Core Competences

Patrick Grosshennig[1] and Volker Grosse-Heitmeyer[2]

[1]Faculty of Mechanical Engineering, Institute of Production Systems and Logistics, Leibniz University Hannover, Germany

[2]Sartorius AG, Goettingen, Germany

Abstract

The purpose of this paper is to describe a methodology for structuring products in a globalization-compatible way on the basis of technological core competence analysis. This enables a precise identification of core components, in order to connect the know-how intensive and market-relevant components with the technological core competence. The key methodological characteristic is an exact analysis of the product structure with the objective of variant management and competence management. Therefore methods of analyzing product structures in order to classify core competences and to identify the necessity of product changes based on variant and competence aspects are discussed.

Keywords: Core competence, Globalization, Product structuring, Variant management

1 INTRODUCTION

Currently, many changes are taking place in the world, in politics and in business. The environment in which today's companies must work can be regarded as turbulent and is characterized by the internationalization of markets, the individualization of products and the dynamization of customer demands and cycles of technical innovation. A suitable product structure can help a company to obtain an optimal flow of production and materials and can lend it a competitive advantage.

It is important to note that significant cost-related decisions take place already in the development and construction phase, and they therefore have an impact on most other business sectors in the company. Product design, for example, also determines the important steps of the manufacturing process. Although studies by Dilling have shown that roughly 75% of costs are already fixed in a company's development and construction phase [1], these studies do not provide a unified, universal plan of action for structuring a product.

The subject of this paper is to describe a methodology for structuring products in a globalization-compatible way on the basis of technological core-competence analysis.

2 BASIC PRINCIPLES

2.1 Core Competence

The idea of core competence is not a novel concept. As early as 1957 and 1959, Selznick [2] and Penrose [3] introduced the term 'distinctive competence' for a company's unique capabilities that distinguish it from its competitors. These capabilities represent the company's unique strengths and are generated by its own bundling of the resources available to it. In an attempt to define a company's critical factors of success, many studies use various synonyms or related terms, such as distinctive competences [4], core capabilities, core skills, strategic capabilities or strategic assets. As a result, when Prahalad and Hamel introduced the concept of core competence in 1990, it was met with considerable resonance [4],[5],[6],[7]. However, a universally recognized definition of core competence still does not exist.

The following is based on Hamel and Prahalad's description of core competence [8]:

- The competence must contribute to a recognizable customer value.
- It must be uniquely different from competitors' competences.
- It must have the potential for developing new business operations.

According to this definition, a core competence is a combination of capabilities and technologies enabling the company to offer products and services which provide customers with an overproportionally high value. Core competences are thus developed for the purpose of contributing to customer value and for distinguishing the company.

2.2 Competence Portfolio

These two factors are the basis for designing a portfolio that categorizes competences according to their contribution to customer value (market effectiveness) and their relative strength (competence leadership) (figure 1). Market effectiveness refers to the customer value generated by the company's evaluated competence in the market performance/providing the service or product, while competence leadership refers to the company's stage of development in comparison to its competitors. These two criteria are what distinguish the four levels of competence: basic competence, key competence, core competence potential, and core competence. These will be described in more detail below [9].

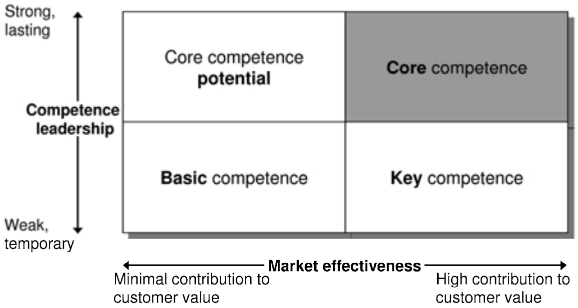


Figure 1: Competence portfolio [9]

A company must have **basic competence** in order to operate its business on certain markets at all. Its competitors also have basic competence, meaning no additional customer value is generated. Basic competences cannot generate any distinct advantages, and they therefore form the bar for entering a certain economic sector. Basic competences are also the first to be outsourced to potential suppliers to save costs.

Key competence contributes to a higher customer value, but competitors are able to reach a similar position relatively quickly. Competitive advantages can be generated by key competences temporarily, for example when a company obtains a one-time technological head-start only for competitors to quickly reach the same level.

A company has **core competence potential** when it has competences which are strong, long-lasting and superior to those of its competitors but have not been transformed into customer value. This potential can be generated in fundamental research, for example.

Core competence is a company's strong, long-term capability which is superior to its competitors' and which generates a distinct customer value. An example of a core competence would be the control of a specific technological process.

2.3 Design guidelines

In addition to the ground rules for designing a 'clear', 'simple' and 'secure' product structure by focusing on the three basic targets of 'realizing technical potential', 'realizing profitability' and 'maintaining safety for people and their environment,' there are other general guidelines. The internationally accepted term for this is 'Design for X.' Design guidelines help to meet specific conditions, support ground rules [10], and must be expanded to accommodate globalization. For this reason, the term 'globalization-compatible' is being introduced here.

3 OUTLINE OF METHODOLOGY

The methodology applied in this paper is an ensemble of several methods by analogy with construction methodology. One of its primary objectives is to identify core assemblies in order to better establish solid connections to technological core competences via know-how-intensive and market-relevant assemblies. The core assemblies consist of the company's product-oriented technological core competences. They are the materialization of these competences on the markets in the form of products. Core assemblies highlight those key components that distinguish the overall system. This can therefore counteract the risks associated with a globalization based solely on labor costs.

This study does not focus on the entire product, but rather on assemblies and component parts. A representative product or a family of product variants can serve as the subject of systematic product structure analysis, the results of which determine how a product can be broken down into its main assemblies and component parts. This serves as the basis for a systematic product structure focusing on product variants and competences. In order to reach the goal of globalization-compatibility, variant and competence management must determine whether each assembly or component part shall be produced in-house or procured from a third party. In some cases, a globalization-compatible product structure may first need to be established through focused reconstruction.

Improvement of the company's already-existing processes is rarely done systematically. With the above arguments in mind, a methodology of creating a globalization-compatible product structure based on technological core competences will be developed in this chapter (figure 2).

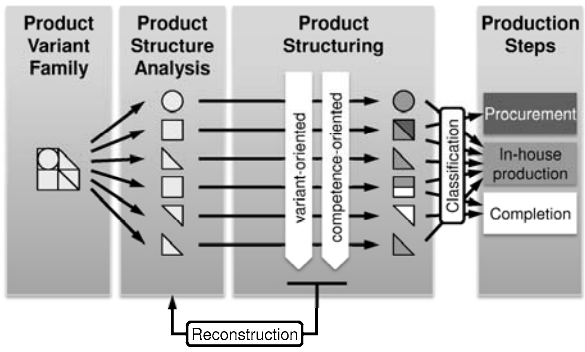


Figure 2: Core methodological concept

3.1 Preliminary strategic considerations

The governing principle of product structure is to have a clear representation of the structure of technical systems. This is contingent on the concrete area of application of the product structure and always focuses on a particular purpose. This method therefore considers the most important options for a globalization-compatible product structure. A variety of fundamental decisions are strategically important and must therefore be reached before product structuring is operationally implemented.

The company's strategy affects the product structure through a variety of mechanisms, for example, through the choice of either a differentiation or cost leadership strategy.

The functionality of a product must be adjusted to meet market demands. Too much variety can be just as damaging as too little; greater variety does not necessarily mean greater customer value. Therefore, it is more important to concentrate on choosing the right kind of variants and not the number. The product must be able to satisfy the demands of its target markets. The necessary number of variants and the resulting product structure must therefore be adjusted to fit these functional requirements.

The technology chosen for realizing a product, however, also has a decisive impact on the product structure. A computer screen manufactured with picture tube technology is made of distinctly different components than a liquid crystal display, although they serve the same basic function. Therefore, a product structure cannot be determined with certainty until the product technology has been chosen. By the same token, production technology must also fulfill certain requirements. If the product structure is fixed on the other hand, this may exclude certain product and production technologies.

A product structure should also not be designed without first taking the company's existing production structures into account. A company with a highly flexible industrial park can hardly produce a bulk product at the same cost as a manufacturer whose production is perhaps inflexible but specialized for this particular product. On the other hand, because of its product structure the bulk product manufacturer cannot all of a sudden produce a large variety of products without driving its costs too high.

These factors must therefore be taken into account and regularly reconsidered when designing a product structure. If a company does not follow these guidelines, it might not be able to save even a very good idea or design from failure.

3.2 Product structure analysis

Depending on the product range's variety, the previously determined representative end products should be taken into account in the product structuring process in order to break down these end products into their product components. These can either be component parts or assemblies or modules. The product structure is uniquely specific to the product and company and can be a multi-level hierarchy of complex products. For this reason, a preliminary product structure analysis makes this process easier.

An assembly-group-oriented product structure based on a structure analysis is the first step toward classifying assemblies. This reduces the complexity of products by enabling further analysis to focus on core assemblies. Each assembly studied is then categorized according to a so-called market know-how matrix (figure 3) in four steps by conducting a series of interviews.

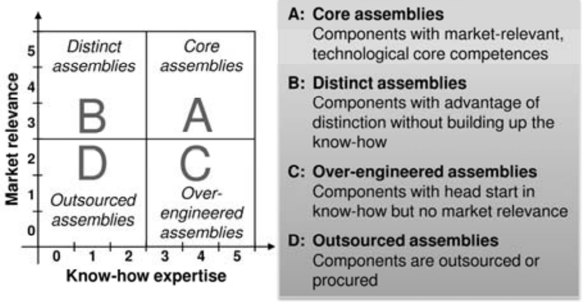


Figure 3: Market know-how matrix [11]

This matrix allows the value of each product element to be more easily visualized and interpreted and helps to identify the core assemblies. The matrix has the two axes of market relevance and know-how expertise which separate the four segments of core assemblies (A), distinct assemblies (B), over-engineered assemblies (C), and outsourced assemblies (D).

Step 1: Product structure

In cases where a product has many variants, classes of variants must first be established in order to classify the assemblies. Each class of product variants groups elements according to their common length, width, height, etc. Individual elements can be distinguished through various forms of these characteristics. All elements of one class are produced in the same manufacturing process.

Later decisions regarding core assemblies may be made for certain classes of product variants or for certain component elements of a class, the so-called representative types. Should a company decide to manufacture a representative type as a core assembly in the future, this decision will affect all assemblies of the same class.

Step 2: Determining market and know-how criteria

The second step is to rate market and know-how criteria according to a scale. Initial market requirements may be generally categorized according to the KANO model [12], which distinguishes between the three categories of basic, performance, and enthusiasm requirements. Basic requirements are not sufficient to satisfy customers. They are minimal requirements that would otherwise result in dissatisfaction if they were not met. A basic requirement of a car would be the safety belts, for example. As to performance requirements, customer satisfaction depends on to what extent these are met. These criteria are what distinguish products from those made by competitors. An example for a performance criterion is a car's engine performance. The third category of requirements concerns customer enthusiasm. These dramatically increase customer satisfaction by offering the effect of surprise. They are highly valued by customers. Returning to our car example, an enthusiasm requirement would be a new extra that is not offered by any other competitor.

The criteria for assessing know-how can be determined by looking at the in-house production depth for an assembly. This is the difference between parts ordered from a catalogue (i.e. norm parts) and simple parts on the one hand, and the in-house production of assemblies and parts requiring expert knowledge and even many years of experience on the other. An example of a detailed scale and listing of market and know-how criteria may be found in figure 4.

Market criteria ranking		Know-how criteria ranking	
0	No purchasing relevance	0	Parts purchased from a catalogue
1	Minimal impact on purchase decision	1	Simple, anyone could do it
2	Possible impact on purchase decision (based on requirements)	2	A certain amount of know-how required for overall context
3	Criterion of distinction compared to competitors' products	3	Expertise knowledge required
4	Important criterion of distinction	4	Difficult to construct as a unit
5	Decisive purchasing criterion	5	Requires many years of experience

Figure 4: Scale of market and know-how criteria [according to 11]

Step 3: Establishing a market know-how matrix

In this step, the variant classes of assemblies are assessed independently by conducting a series of interviews, for example with engineers and distribution personnel, regarding know-how and marketing aspects. The products are assessed according to the determined market and know-how criteria after categorizing their assembly or main components with the help of the hierarchy of variants and intended bill of materials. In this process, it is important that each interviewee is not aware of the results of the other interviews. This allows the different engineering and distribution assessments of subsystems to become clear. After conducting these interviews, the results are entered into the market know-how matrix.

Step 4: Assessment of the market know-how matrix

The last step consists of assessing the assemblies in the market know-how matrix (figure 5), providing the first insight into which assemblies can be outsourced.

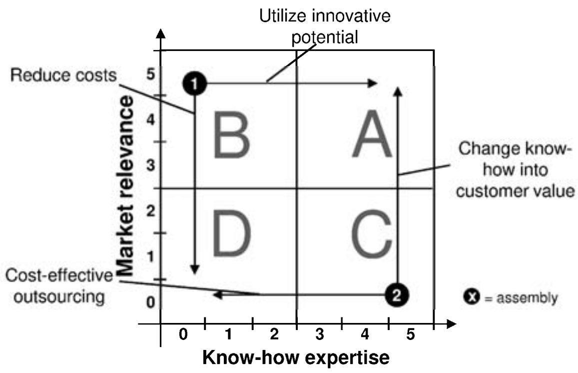


Figure 5: Strategies for assemblies in areas B and C [according to 11]

Those assemblies with no market relevance are predestined for outsourcing. Considering market relevance, action must be taken for those groups with too little or too much technical know-how. For the assemblies (2) in the over-engineered section (C), the company could strive to turn this potential into real customer value and to make this group its core assembly. Distribution might not even be aware of this advantage in know-how. If, however, turning potential into real customer value is impossible to achieve, this assembly should be replaced by standard solutions, or a suitable partner should be found to reduce the costs of this assembly.

Distinct advantages are generated by subsystems in section B's distinct assembly (1) in which long-term know-how has not been built up. This advantage can be quickly lost if competitors copy the system or acquire items from the same supplier. For this reason, it is recommended that the company dramatically cut costs in this sector in order to utilize its time advantage and to build entry barriers. The innovative potential of these subsystems is questionable because so little expertise is required to obtain a distinct advantage and because its market relevance is high.

The assemblies in the areas B and C must be judged as to whether or not they rather belong to area A of core assemblies or area D of outsourced components. The depth of production must be determined. In the methodological section below, core assemblies and/or the other determined assemblies with innovative potential or assemblies with know-how customer appeal will be analyzed.

3.3 Variant-oriented product structure

Variant management functions as the interface between an external variability determined by customers and an internal variability that must be designed by the company. Systematically building up product variants is equal to actively influencing the product structure. For this reason, the product structure is the basis for variant management and the resulting product configuration. This enables many ways to design a globalization-compatible product structure with a focus on variant management, but these will not be discussed in detail in this paper. Further approaches may be found in Grosse-Heitmeyer's dissertation [13].

3.4 Competence-oriented product structuring

Engineering is important for the company's entire value chain, and it goes beyond the company's product-related technologies. Especially in those high-tech sectors where the life-cycle of competences is much longer than the life-cycle of products, the analysis of a company should not only include its products, but a process-oriented analysis should also be done in order to identify its technology. Determining and carrying out measures to further develop technological competences require human and financial resources, which are usually limited. It is therefore sensible to focus on a fixed number of objects of assessment, such as already determined assemblies. If the items to be assessed are predetermined by the method applied, then it is not necessary to limit the number of related assemblies and processes. This could be the case if the company management has decided to review one specific manufacturing technology used for certain related objects and to expand its technological competences in this business sector. Similarly, when assessing the technological competences of potential partners, there will generally be one limited object of assessment. The selection of the object does not play a role for homogeneous product lines and cases of minimal complexity. However, a systematic consecutive procedure is recommended for homogeneous product lines. This procedure will be described in further detail in the next section.

Analysis of technological core competences

Core competences are understood here to be a group of different competences. However, in the following, we will take a closer look at only technological competences. These include bundled technological capabilities employed for a certain purpose. Technological competences and core assemblies therefore interrelate. They can also relate to certain technological capabilities which are not readily apparent in core assemblies but which have been identified in the analysis of core technological processes.

After identifying the relevant process technologies in the core assemblies, it is important to evaluate these technologies from a strategic point of view. For this purpose, the portfolio analysis was chosen because portfolios are an effective way of visualizing complex relations, and, as grounds for discussion, they enable initial strategic conclusions to be drawn and measures proposed.

Assessment of the actual situation

By surveying all of the assessments of the company's current strong competences and future potential for the development of existing core competences, its core competences can be positioned within the competence portfolio (figure 6).

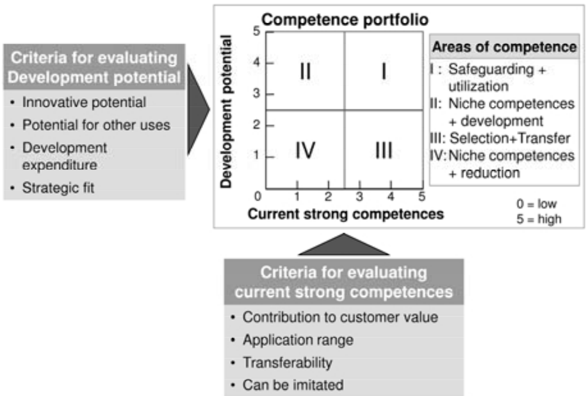


Figure 6: Portfolio of technological competences [according to 14]

This portfolio provides clues as to whether or how to further expand the company's core competences, or whether they should rather be given a smaller role. The portfolio is divided into four quadrants which will be described in detail below. The competence portfolio also allows competence management to determine its priorities so that it and F&E can set their future practices and areas of specialization. A variety of approaches shall be discussed below according to the order of their quadrant.

Quadrant I: Safeguarding and utilization

The technological core competences in quadrant I of the competence portfolio fall under the 'safeguarding and utilization' strategy and are of high strategic relevance for the company's competitiveness in the future. They exhibit strong current competences and a high potential for development. However, it is necessary that these core competences be maintained, supported and further developed. The utilization of these existing core competences therefore has high priority and should be pushed ahead through application-oriented development projects and innovation.

Quadrant II: Niche competences and development

If the development potential of core competences is high but current competences are not very strong, then the company should not give up existing potential. Instead, it should further develop its current core competences and continue to cultivate new core competences in this sector.

Quadrant III: Selection and transfer

The technological competences found in the third quadrant are relatively strong, but their strategic significance for the future is rather minimal. These are often competences that were built up for the current production sectors but which will no longer be in demand for future products. In this situation, the company can strive to establish new areas of application and/or develop new markets and regions. Should this prove unsuccessful and these competences remain insignificant and show only a minimal development potential in other conceivable sectors, then these previous core competences should be discontinued.

Quadrant IV: Niche competences and reduction

Competences in this quadrant are rather insignificant for the company strategically and have a low current level. They are often capabilities which round off the company's product range in a certain sector and therefore have little strategic significance. The company should not invest in further developing these capabilities, and it should rather consider withdrawing from these areas of competence in the future.

The company's product range can be outlined based on the competence portfolio. Especially those core competences in the first quadrant should not be outsourced or discontinued or important know-how will be lost.

3.5 Determining the globalization-compatible product structure

In this paper, variant- and competence-oriented analyses of the product structure serve as the basis for real policy recommendations for how to establish an effective product structure. Each step of the analysis offers its own policy recommendation which is listed in a catalogue of measures. Through this methodological process, a globalization-compatible product structure that meets company targets must be determined. The design and composition of products must be adjusted to fit new global demands. According to the policy recommendations in each step of analysis (sections 3.3 and 3.4), design and composition must be combined in order to find solutions for the entire product. A globalization-compatible product structure must guarantee that certain selected assemblies can be outsourced to a global partner. This applies to both procurement and sales. Core assemblies requiring the company's core competences should be kept on location, while highly standardized and semi-finished consumer goods can be purchased on global markets. However, these components must be technologically reconstructed to accommodate the unique characteristics of the production location. Additionally, the product structure should strive to customize product completion for a country's customer specifications as much as possible.

Product structure is applied to the structure of production according to the motto 'production follows product'. The modulization of production creates autonomous steps of production that meet the demands of different product modules. When a product module is altered, this only affects its related production area. Variants which were developed late deserve special attention here. They could benefit from the implementation of flexible manufacturing units or modules or individual machines for manufacturing residual products [15]. Establishing and maintaining lasting process flexibility for the future also deserves special attention when designing global production steps.

The production stages method of structuring production locations is based on assigning bundles of related manufacturing processes for different items to certain manufacturing locations [Stre00]. An important goal for establishing a globalization-compatible product structure is to create a product structure that adds the dimension of 'core competence' to the production stages approach. Bundling production tasks according to competences leads to similar requirements being made on manufacturing locations. This allows the company to extensively utilize the advantages of each production location by positioning those production steps made for certain assemblies or modules in those locations where the local conditions provide these competences.

The strategy of establishing different production steps also protects the company's own core production competences. Know-how-intensive areas of production can be concentrated in locations where they can be better controlled (for example at the company's headquarters). This means that production steps allow production to be organized according to competence. This organization creates the basis for a global production system with three characteristic steps: **Procurement**; **competence-oriented in-house production**; and **customer-oriented completion**. These steps of production and this method of structuring a global network of production are also the subject of a research project, GVP – Globales Varianten Produktionssystem [16],[17],[18], financed by the German Ministry of Education and Research.

4 SUMMARY

The method presented in this paper enables manufacturers to establish a globalization-compatible product structure in four steps. First, a plan of action should be drawn up. Next, a precise product structure analysis focusing on core assemblies and variant- and competence-based product structures should follow. Finally a globalization-compatible product structure can be established, all in one consecutive approach.

5 REFERENCES

[1] Dilling, H.-J., 1978, Methodisches Rationalisieren von Fertigungsprozessen am Beispiel montagegerechter Produktgestaltung, Diss., TH Darmstadt.

[2] Selznick, P., 1957, Leadership in Administration: A Sociological Perspective, New York.

[3] Penrose, E. T., 1995, The Theory of the Growth of the Firm, Oxford University Press, Oxford.

[4] Hitt, M., Ireland, R. D., 1985, Corporate distinctive competence: strategy, industry and performance, In: SMJ Vol. 6, p. 273-293.

[5] Leonard-Barton, D., 1992, Core Capabilities and Core Rigidities: A Paradox in Managing New Product Development. In: SMJ Vol. 12, 111-125.

[6] Prahalad, C. K., Hamel, G., 1990, The Core Competence of the Corporation, In: Harvard Business Review, 79-91.

[7] Hinterhuber, H., 1997, Strategische Unternehmensführung: I. Strategisches Denken, 7th ed., Berlin, New York.

[8] Hamel, G., Prahalad, C.K., 1995, Wettlauf um die Zukunft, Wirtschaftsverlag Carl Ueberreuter, Vienna.

[9] Hinterhuber, H., et al., 2003, Kundenzufriedenheit durch Kernkompetenzen, Eigene Potenziale erkennen - entwickeln - umsetzen, Gabler Verlag, Wiesbaden.

[10] Pahl, G., 2007, Konstruktionslehre - Grundlagen erfolgreicher Produktentwicklung Methoden und Anwendung, Springer Verlag, Berlin, Heidelberg.

[11] Zehnder, T., 1997, Kompetenzbasierte Technologieplanung - Analyse und Bewertung technologischer Fähigkeiten, Gabler Edition Wissenschaft, Deutscher Universitätsverlag, Wiesbaden.

[12] Kano, N., et al., 1989, Attractive Quality and Must Be Quality, In: Quality Journal, 14, 39-48.

[13] Grosse-Heitmeyer, V., 2006, Globalisierungsgerechte Produktstruktur auf Basis technologischer Kernkompetenzen, Berichte aus der IFA, Garbsen.

[14] Thomsen, E.-H., 2001, Management von Kernkompetenzen. Methode zur Identifikation und Entwicklung von Kernkompetenzen für die erfolgreiche strategische Ausrichtung von Unternehmen, Techn. Univ. München, Diss., Verlag Wissenschaft und Praxis, Munich.

[15] Nyhuis, P., Müller-Seegers, M., 2004, Individuelle Fahrzeuge trotz Serienproduktion, In: Branchenreport 2004 - Automobilzulieferer, 104-105.

[16] Globales Varianten Produktionssystem http://www.gvp-projekt.de.

[17] Großhennig, P., Mühlenbruch, H., Nyhuis, P., 2006, Design of Global Production Networks, Proceedings of the 11th Annual International Conference on Industrial Engineering - Theory, Applications and Practice, Nagoya, JAPAN, October 24-27.

[18] Nyhuis, P., Mühlenbruch, H., Großhennig, P., 2005, Design of Production Stages and Logistics in the Global Variant Production System, Proceedings of the 10th Annual International Conference on Industrial Engineering - Theory, Applications and Practice, Clearwater Beach, Florida, USA, December 4 – 7.

Reducing design complexity of multidisciplinary domain integrated products: a case study

Juan M. Jauregui-Becker, Wessel W. Wits and Fred J.A.M. van Houten

Laboratory of Design, Production and Management, Faculty of Engineering Technology,

University of Twente, The Netherlands.

Abstract

Multidisciplinary product development is well known for the complexity of its design process. It is commonly addressed by domain integration and a modular design approach. The former, often resulting in smaller products and integrated functions, is characterized by a complex non-linear design process. The latter, which may not result in such integrated functions, has a simpler –usually linear– design process, resulting in novel solutions. This paper presents a method for reducing design complexity of Multidisciplinary Domain Integrated Products by decomposing the problem into modular structures. Computational synthesis techniques are used to solve the resulting modules. Printed Circuit Board design is used as case study, as it is well known for its complexity and highly integrated product functionalities.

Keywords:

Domain integration, multidisciplinary design, complexity reduction, computational synthesis

1 INTRODUCTION

Advances in technology are driving modern products towards further miniaturization, better quality, more functionality, and yet cheaper prices. This trend is imposing a great challenge on product development, thus becoming extremely complex. To meet the challenging criteria, several disciplines are often integrated in one product and, as such, require a multidisciplinary product development. However, multidisciplinary product development is a very complex task, due to the amount of engineering disciplines that merge together into one design artifact.

Engineers often use two main approaches to tackle multidisciplinary product development, namely, modular design or domain integration. A modular approach isolates product functions and maps them onto independent modules. Domain integration incorporates different functions into a single module, resulting in Multidisciplinary Domain Integrated Products (MDIP). Designing for MDIP requires the knowledge of multiple engineering fields simultaneously, before arriving at any design solution. Decomposing the design into smaller chunks –a divide and conquer strategy- is often used to reduce complexity [1]. However, the resulting chunks are required to have clear boundaries, governing principles and requirements. According to Goel et al [2], design decomposition has to be based on problem structure rather than on designer's experience. This avoids unwanted side effects, such as multiple interdependencies among sub-problems and disjoints sub-problems, whose articulation constitutes a problem as such. However, given the integrative nature of MDIP, many interactions among decomposed sub-problems usually appear. This increases the complexity of the design significantly. Therefore, intelligent strategies to treat each sub-problem independently and modular while considering the interactions among each other are required.

This paper presents a method to reduce design complexity of MDIP. This method is demonstrated in the field of Printed Circuit Board (PCB) design, which will be introduced in Section 2. The problem is organized and structured as shown in Section 3. The results are then used in Section 4 to decompose the problem. Functional, topologic and physical characteristics are used for this purpose. As many interactions appear, an algorithm is presented to determine the optimal sequence for instantiating each sub-problem. Section 6 discusses how a Computational Synthesis System (CSS) can encapsulate the resulting problem chunks into separate modules, allowing designers to follow a modular design approach.

2 CASE STUDY: PRINTED CIRCUIT BOARD DESIGN

PCB technology is a well established mass-market production method, appreciated for its high degree of integration of mechanical and electronic functions. Polymeric layers, on which a conductive pattern is produced, are laminated together to form a PCB. On top and bottom of the PCB, electronic components (e.g. IC's, resistors, connectors) can be assembled, thus forming a circuit card assembly.

Traditionally, the design process of most electronic products has been dominated by electrical and mechanical requirements. As cooling requirements are (usually) not a primary function of electronic products, its design has been under-addressed. In fact, recent literature study indicates that a limit has been reached for cooling electronics in general [3]. An effective method to conquer this design challenge has been presented by Wits et al. [3]. Here, knowledge of heat transfer and production principles was combined through domain integration into the overall design process of electronic products. Their concept demonstrator, illustrated in Figure 1, incorporates heat pipes produced by the PCB structure itself to effectively transport heat away from the dissipating elements. Heat exchangers can be located independently and further away from their respective heat source (i.e. the electronic components). For PCB technology in general, this concept leads to new manufacturing strategies where thermal management functions and electronic circuitry are fully integrated. It also results in a more compact electronic system, producible at a lower cost. However, inevitably, the design complexity of such a product also increases, thus, requiring new methods and strategies to cope with this family of highly integrated design challenges.

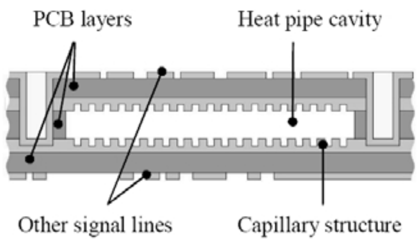


Figure 1: Heat pipes integrated into the PCB [3].

3 PROBLEM STRUCTURE

The formulation scheme of routine design problems presented by Jauregui-Becker et al [4] is used in this paper to structure the MDIP design, as it allows separating different types of knowledge present in design problems. In addition, it structures and models problems such that the computational processes that are required to automatically generate design solutions can easily be identified. The following sections present a short summary of what it embraces.

3.1 Definitions

Element: is a class description of a design artifact component.

Descriptions: characterize an element class by representing its attributes in the form of variables.

Embodiment: is considered as the subset of descriptions of an element upon which instances are created to generate design solutions.

Scenario: is considered as the subset of environment variables, attributed to elements in the natural world and considered in measuring a design artifact's ability to accomplish its function.

Performances: are those descriptions used to express and assess the artifacts behavior, and are calculated using analysis relations.

Analysis relations: use known theories, for instance the laws of physics or economics, to model the interaction of the design artifact with its environment and to predict its behavior.

Topology relations: define the configuration between embodiment and scenario elements.

Objective function: weighs and adds the performances to result in the overall performance of the design.

Confinement constraints: determine the range in which a description can be instantiated.

3.2 Design problem formulation

Embodiment elements and scenario elements, with their attributed descriptions, are used to describe the initial state of a design. Objective function, performance indicators and analysis relations express and assess the goal of the design artifact. Topological relations indicate the set of logical states between elements that have to hold for the design artifact to exist. The RCC-8 [5] formalism is used here to represent the topologic configuration of the design, as it is ideally suited for qualitative spatial representation and reasoning. It comprehends the base relations shown in Figure 2, and allows the formalization of concepts as convergence, connectedness and continuity.

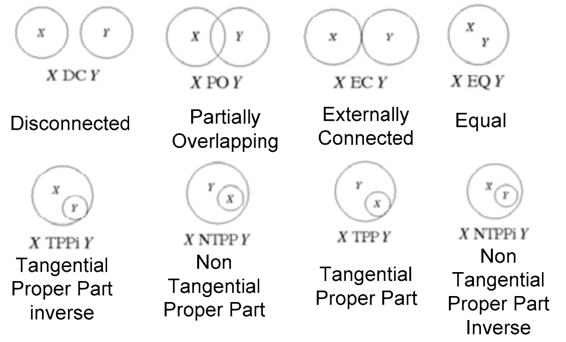


Figure 2: RCC-8 base relations.

In Figure 3, a formulation is presented of the PCB design of our case study, discussed in Section 2. For explanatory reasons, not all descriptions and relations have been included. Figure 3 shows that a PCB consists of laminates, heat pipes and wiring in one embodiment element. The properties of the electronic components –such as location, mass and generated heat– are regarded as scenario. Topology, analysis and objective function indicate how all elements in the formulation relate to each other. These relations also show the complexity of PCB design in general.

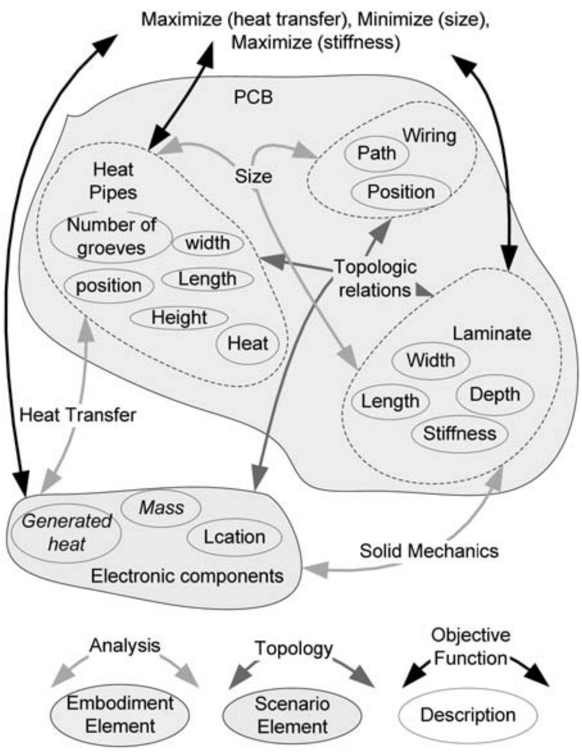


Figure 3: PCB problem formulation.

4 PROBLEM DECOMPOSITION

As a structured formulation of the problem is realized, decomposition is done according to:

1. A model of the elements and its topology is developed, as shown in Figure 4(a) for PCB design. The following convention is used to do so:
    - All elements in the topology must be connected by at least one base relation.
    - Relations have a direction indicating which the reference element of the relation is. For example, in Figure 4(a), the element '*Laminate'* is the reference element to which the element *'Electronic Components*' is related to.
    - The direction of the relation is given by interrogating which element has to exist in order to be able to accommodate the other.
2. The functions of the elements in the embodiment are listed. Functions that are not related to other elements in the topology are not taken into account. For PCB design, the following functions are listed:
    - Laminate: accommodate electronic components, heat pipes and wiring.
    - Wiring: Connect electronic components.
    - Heat pipes: Remove heat from electronic components.
3. Elements presenting several functions are split into new elements, one for each of the encountered functions. For integrated elements with several functions this results in independent elements, for each of the functions of the design artifact. Figure 4(b) illustrates the separation of elements for *'Laminates'*, *'Heat Pipes'* and *'Wiring'*.
4. As the topology has been split into new elements, relations emerging from the separation have to be re-formalized. This is done according to the same guidelines presented previously. The result is a new problem formulation, as shown in Figure 4(b). For instance, here *'Heat Pipes'* and *'Wiring'* are connected to *'Laminate'* by an inside relation (TPP).
5. Now, we categorize the analysis relations by their underlying theories. For this purpose, we make use of perspectives, which are regarded as abstraction levels founded by the underlying theories upon which the goal of the design is assessed. In our case heat transfer, solid mechanics and geometry. Perspectives are identified by assessing the performances. In the case of subject PCB design, three perspectives are defined, as shown in Figure 4(c):
    - Heat transport: Uses heat transfer as underlying theory to determine the amount of heat that the heat pipes are able to remove from the electronic components.
    - Stiffness of the laminates: Uses solid mechanics theory to determine the stiffness of the laminates under the presence of mass of the electronic components and the laminate itself.
    - Size of the PCB: Uses geometry to determine the total volume of the laminates.

    The direction of the relations is determined following the same convention as for the topologic relations.

The result of the applied method is a set of syntactic rules indicating how elements and functional element instantiations is limited by the previous instantiation of other elements. The final decomposition is shown in Figure 4(d). Here, labels are attached for usage in Section 5.

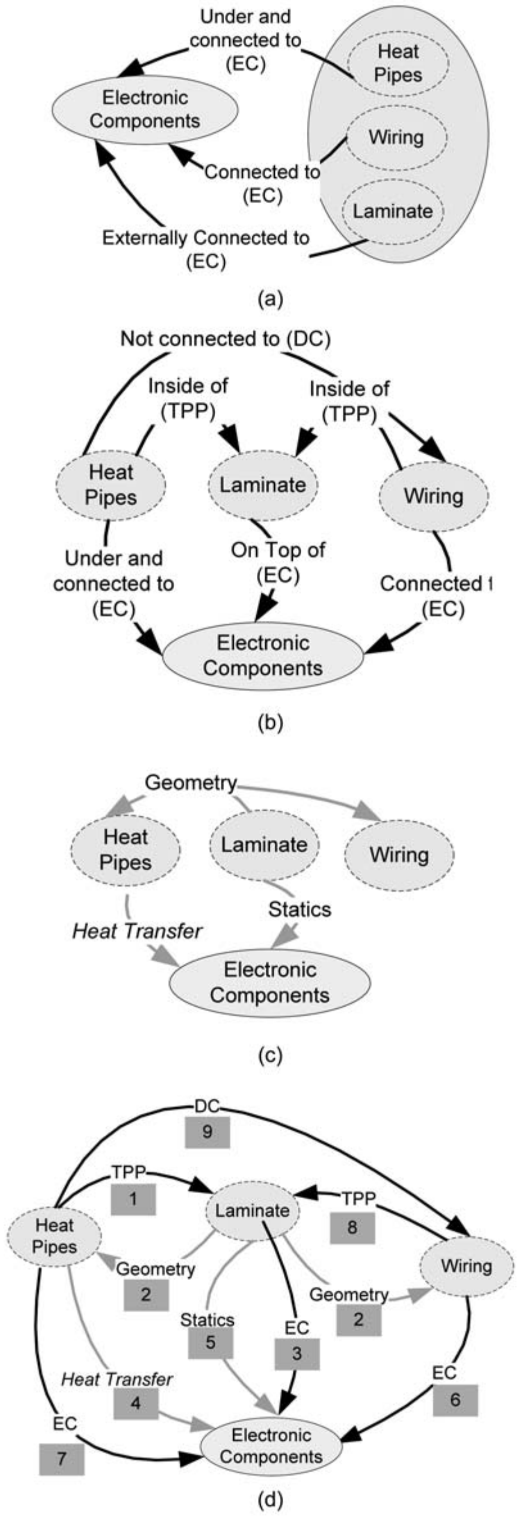


Figure 4: Several steps in PCB decomposition.

5 PROBLEM SOLVING STRATEGY

As shown in Figure 4(d), several additional interactions appear as consequence of the decomposition. Interactions do not necessarily represent an escalade of the design complexity. If managed properly, they determine how to proceed in solving the problem. For this purpose, a method was developed to determine strategies to analyse and optimize the decomposed problem. The basis of this method is found in Operations Research (OR). Here, mathematic algorithms are used to make optimal decisions in complex problems concerning optimization. The aim is to reduce NP-Hard complexities into NP ones. In this section, we will focus on the method itself and not in how it was implemented, as the object of this paper is rather instrumental and not analytical.

5.1 Design Structuring Matrix

The problem we want to solve is that of determining how to optimally proceed in the instantiation of the elements in the design structure, and which relations should be taken into account at each step in the strategy. To do so, we make use of a Design Structure Matrix (DSM). Figure 5 shows a DSM for PCB design strategy optimization. In this figure, E stands for *'Electronic Components'*, H for *'Heat Pipes'*, L for '*Laminate'* and W for '*Wiring'*. Each row and column in the matrix is labeled after one element. To fill the DSM, the following steps are taken:

1. Select one element of the model.
2. List the relations pointing towards that element.
3. Fill the row, attributed to that element, with the listed relations.

For example, the element *'Heat Pipes'* has one relation pointing inwards, connecting this element to the element *'Laminate'*. In the DSM this is represented by the label '2' at row H and column L. As shown in Figure 5, one row and one column are included for measuring the number of constraints to be satisfied in the instantiation of an element, and the number of constraints imposed by one element onto others, respectively. The former is regarded as *effort*, and its value equals the amount of relations present in a column. It represents the degree of complexity attributed to the instantiation of an element. The latter is regard as *potential*; its value equals the amount of relations present in a row. It represents a measure to reduce the complexity of following elements instantiations.

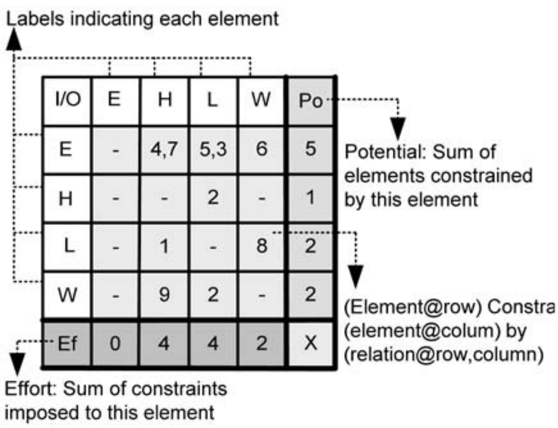


Figure 5: DSM for PCB design strategy.

5.2 Sequential strategy

After filling the DSM matrix, a number of steps follow to complete the design strategy. The design strategy describes the instantiation sequence, and the relations used at each instantiation step. The strategy is developed as follows:

1. Identify the element to be instantiated by:
   a. Finding the elements with the lowest effort.
   b. Selecting the element with the highest potential.
   c. Setting its effort to zero (0), as shown in Figure 6(b).

   The selected element thus becomes the next to be instantiated.
2. List the relations contained by the column of the selected element, and erase them from the matrix. This is shown in Figure 6(b), (c) and (d) by the zeros replacing the erased relations labels. These relations are to be taken into account at the instantiation of the selected element, as indicated in Table 1.
3. Calculate the potential with the remaining relations, as shown in Figure 6(b), (c) and (d) where, as the method progresses, new potentials are calculated with the remaining relations.
4. Calculate efforts considering elements with non-zero efforts.

This process is repeated until all elements and relations have been addressed in the strategy. In the case of subject PCB design, this routine is applied four times. Figures 6(a-d) show how the matrix develops as the method is applied. Table 1 specifies the resulting instantiation order and weighs the reduced complexities by the updated efforts.

I/O	E	H	L	W	Po
E	-	4,7	5,3	6	5
H	-	-	2	-	1
L	-	1	-	8	2
W	-	9	2	-	2
Ef	0	4	4	2	X

(a)

I/O	E	H	L	W	Po
E	-	4,7	5,3	0	4
H	-	-	2	-	1
L	-	1	-	0	1
W	-	9	2	-	2
Ef	0	3	3	0	X

(b)

I/O	E	H	L	W	Po
E	-	0	5,3	0	1
H	-	-	2	-	1
L	0	0	-	0	0
W	-	0	2	-	1
Ef	0	0	2	0	X

(c)

I/O	E	H	L	W	Po
E	-	0	0	0	0
H	-	-	0	-	0
L	0	0	-	0	0
W	-	0	0	-	0
Ef	0	0	0	0	X

(d)

Figure 6: Steps for strategy development of PCB.

Table 1: Design strategy for PCB design.

Sequence	Element	Relation	Initial effort	Updated Effort
1	Wiring	6,8	2	2
2	Heat pipes	4,7,1,9	4	3
3	Laminate	5,3,2	4	2

6 SYNTHESIS TOOL DEVELOPMENT

The result of the decomposition is a linear design process, where each step is regarded as one sub-problem. Results of one step are regarded as scenario (input) for the following design steps. Several loops may be required before finding at least one valid solution. By applying CSS to each of these sub- problems, while designers control the flow of data between their interfaces, allows for further reduction of complexity. In this section, some generalities of CSS development are discussed. We limit the scope of CSS to routine design problems with parametric representations. A demonstration will be presented for heat pipe design in PCB.

6.1 General considerations for CSS development

A well accepted generic model of the design process is shown in Figure 7. According to this, embodiments are generated in a synthesis process. Then they are analyzed to calculate its performance and evaluated to assess whether the design is to be adjusted (path 1), rejected (path 2) or accepted (path 3). The four main processes to capture in a CSS are Synthesis, Analysis, Evaluation and Adjustment, of which Synthesis is the core one. Many techniques for its execution are found in literature, as for instance generation and test, backward reasoning, database lookup, search in a problem space, abduction, generative rules, case-based reasoning, grammar production, computational models, parametric search/generation, constraint-solving, and genetic algorithm.

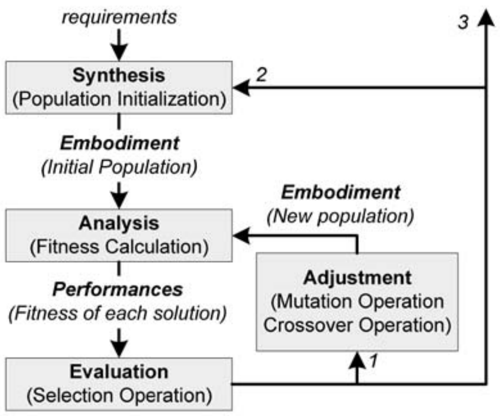


Figure 7: GA vs. generic design process.

Selecting a specific technique depends mainly on the type of model used in the representation. Jauregui-Becker et al [4] presented a classification for descriptions and relation models commonly used for design representations. Using this classification, four main problem distinctions were made: parametric, configuration, layout and shape. Parametric representations are commonly solved by generation and test, constraint-solving and genetic algorithm, whereas grammatical approaches are very common for the generation of configurations and shapes. Layout problems are usually solved by optimization algorithms, as simulated annealing or genetic algorithms.

6.2 Computational synthesis for heat pipe design

The physical model and implementation for a PCB with integrated heat pipe cooling is presented by Wits et al in [6]. This model is used here to derive the parameters and relations considered in the CSS. Figure 8 summarizes embodiment parameters in relation with performance and scenario parameters. For instance, a change in groove dimensions or internal height alters the heat pipe's performance, and thus its capacity to remove heat from the associated electronic components.

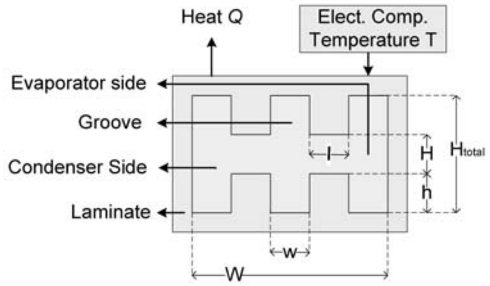


Figure 8: Embodiment and performance for PCB design.

As the model is fully parametric, we have chosen for Genetic Algorithms (GA) as solution generation technique, developed as indicated by Barrell in [7]. GA incorporates the process of Synthesis, Analysis, Evaluation and Adjustment required in a CSS, as shown in Figure 7 where the GA algorithm is presented together with the mentioned processes. This characteristic makes GA implementations straightforward for CSS applications for parametric routine design problems. The GA algorithm was implemented using C# as the programming language. The program consists of five classes: *Population*, *ListGenoma*, *Genom*, *Model and UserInterface*, as shown in Figure 9.

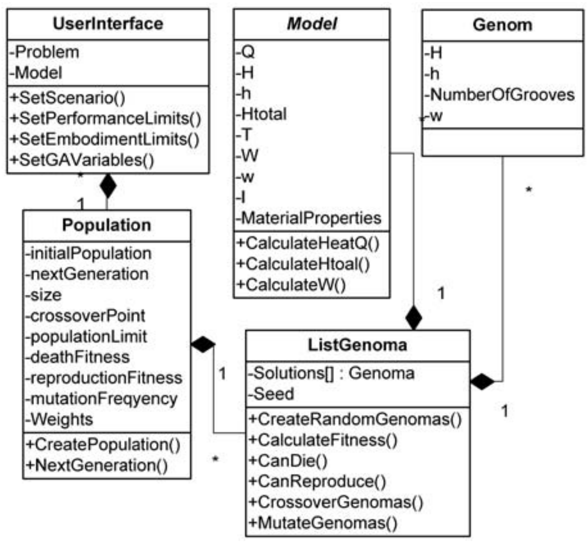


Figure 9: Class diagram of the CSS.

The class *Population* contains variables specifying the size of the initial population, maximum population, death fitness, reproduction fitness, mutation frequency, crossover point and weight of the performance parameters. It stores the generated populations in the class *ListGenoma*. The class *ListGenoma* contains a population of solutions, where each solution is an object of the class *Genom*. *ListGenoma* contain methods to compute the GA operators. These methods are called by the *Population class* using variables that characterize the solution generation and adjustment (e.g. size of population). *Genom* is the class that encapsulates the embodiment parameters of one candidate solution. The *Model* class contains properties of the global physical model related to each Genom instantiation. This class responds to calls from the *ListGenoma* class to calculate the performance of each candidate solution. The class *UserInterface* allows the user to enter input requirements of a specific problem to be solved. The modularity of this architecture allows the re-use of program code by minor changed of the underlying classes.

Simulations were carried out with the developed software tool to determine how fast solutions could be generated. A first simulation is used to scan the solution space. The ranges of the desired requirements for this simulation are shown in Figure 10, where the user input interface of the developed tool is shown.

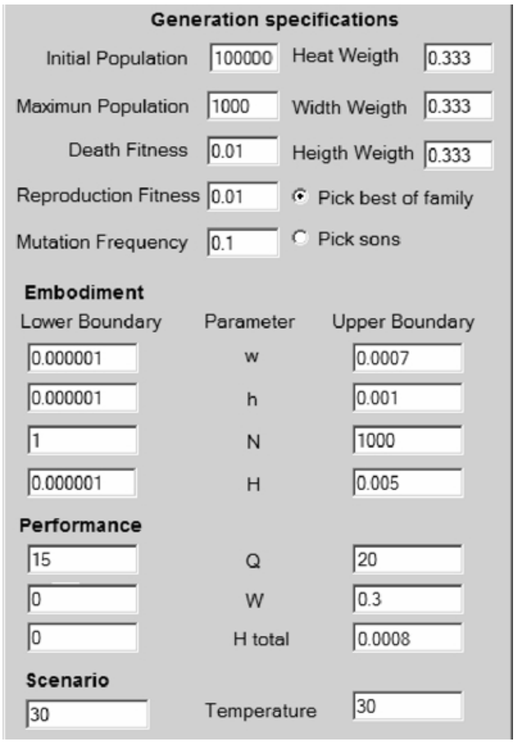


Figure 10: User input interface.

From the range of parameters, solutions are generated randomly and filtered for fitnesses higher than 0.001, in about 2*s*. The results are shown in Figure 11, where the performance parameters *'Heat'* and the *'Area'* of the designed heat pipes are shown. As the figure shows, 'generate and test' gives an overview of how the solutions are distributed in the solution space. The outcome of the initial simulation is used in a second iteration, for instance to zoom in to the selected region in Figure 11. This is done by updating the confinement constraints of the allowed values for the Heat (Q) and the area (WxH). This simulation took about 5*s* to find the local optima illustrated in Figure 11.

7 SUMMARY

This paper presented a method to reduce design complexity of Multidisciplinary Domain Integrated Products. This method allows decomposing of such MDIP design challenges into modular structures. The resulting structures can then be implemented into a Computational Synthesis System. This allows designers to assess the solution space in shorter times, while keeping an insight in the found solutions. The method was demonstrated by a case study to design a complex multilayer Printed Circuit Board with integrated heat pipes. A CSS was developed to optimize the dimensioning of the integrated heat pipes as a function of required cooling capacity and board layout. The results demonstrated the feasibility and applicability of the developed method for MDIP design.

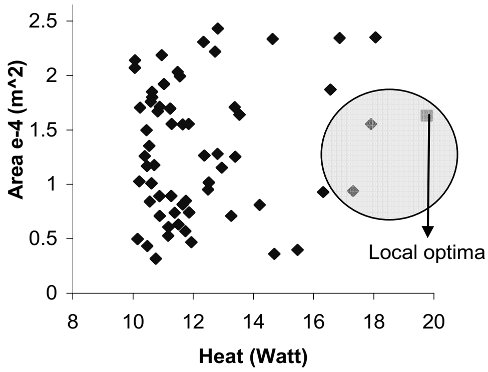


Figure 11: Performances for CSS generated designs.

LITERATURE

[1] Tomiyama, T., 2006, Dealing with Complexity in Design: A Knowledge Point of View, R. Design Methods for Practice, pp. 137-146.

[2] Goel, V., Pirolli, P., 1992, The structure of design problem spaces, Cognitive Science, 16(3), 395-429.

[3] Wits, Wessel W., Vaneker, Tom H.J., Mannak, Jan H., Houten, Fred J.A.M. van, 2007, Domain integration and cost reduction in electronic product design: a case study, Proceedings of 40th CIRP International Manufacturing Systems Seminar.

[4] Jauregui-Becker, JM., Tragter, H., Houten, FJAM. van, 2008, Structuring and modeling routine design problems for computational synthesis development, Proceedings of CIRP Design Seminar 2008. Abstract accepted.

[5] Renz J., Nebel B., 1998, Spatial Reasoning with Topological Information, Lecture Notes in Computer Science, 1404, 351.

[6] Wits, W., Legtenberg, R., Mannak, J. and Zalk, B. van, 2006, Thermal Management through In-Board Heat Pipes Manufactured using Printed Circuit Board Multilayer Technology, Proceedings International Electronic Manufacturing Technology (IEMT), pp. 55-61.

[7] Whitley, D., 1998, A Genetic Algorithm Tutorial Darrell, Statistics and Computing, 4:65-85.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the support of the Dutch Innovation Oriented Research Program 'Integrated Product Creation and Realization (IOP-IPCR)'. This work is also part of the PACMAN project, SenterNovem project TSIT3049. Both activities are founded by the Netherlands Ministry of Economic Affairs. The authors especially want to thank Thales Nederland and in particular Jan Mannak for their effort and energy, contribution to this research project.

PLIB Ontology for Great Group Technology

Qing-sheng Xie[1], Gui-xian zhou[2] and Shao-bo Li[3]

[1] Key Lab of Modern Manufacturing Technology, Guizhou University, Guiyang, China

[2] Dept. of Mechanical Engineering, SouthWest Jiaotong University, Chengdu, China

[3] Key Lab of Modern Manufacturing Technology, Guizhou University, Guiyang, China

Abstract

Great Group Technology (GGT) is the combination of Group Technology and network technology. The core and basis of GGT is an open, across-industrial parts information exchange and integration platform. In this paper, we mainly discuss the realization of this platform. Firstly,the framework and working mechanism of this platform is introduced. Then the two core parts (web-based parts library and part search engine) of this platform are pointed out and analyzed. Through requirement analysis, we think PLIB ontology is fit for the solution of these requirements (unambiguous description, integration automatically, variant design in real time etc). Then a parts library information model is built. Because the function feature and structure feature is not included in PLIB ontology, an Extended PLIB Ontology library is built to solve this problem. The Extended PLIB ontology library is composed of two parts: PLIB ontology based on the PLIB standard and semantic feature ontology. At last, the realization method and farther work are introduced. Now, the prototype system has been tested in our lab.

Keywords:

Group technology; Network technology; PLIB ontology; Web-based parts library

1 INTRODUCTION

Great Group Technology (GGT), which was brought forward by Prof. Xin-jian Gu and Prof. Guo-ning Qi[1], combines the network technology with the traditional Group Technology .Through adopting self-organizing and distributed management mode and using GT in a large scope, enterprise clusters can be optimized holistically; Through grouping design and machining information which come form various enterprises, the set of similar parts can be distinctly widened (Figure.1). By using GGT, manufacturers can provide satisfied products to their customers at a high speed and a low cost. The key for application and popularization of GGT is an open, parts information exchange and integration platform[2]. The aim of the open, across-industrial parts information exchange and integration platform is to (i) help the designer find needed parts and their suppliers via Internet quickly, and help specialized parts manufacturers sale its product via Internet; (ii) realize group analysis and group manufacturing in large scope.(Ordinarily, the part was classified as standard part, common part and special part. In this paper, we mainly aimed at the standard part and common part).

Currently, such parts information exchange and integration platform has gained development overseas. But it is at the initial phase in China.

In this paper, we combine the PLIB ontology in GGT, and mainly discuss the application of PLIB ontology in the open, across-industrial parts information exchange and integration platform (for short: PLIBEI platform).

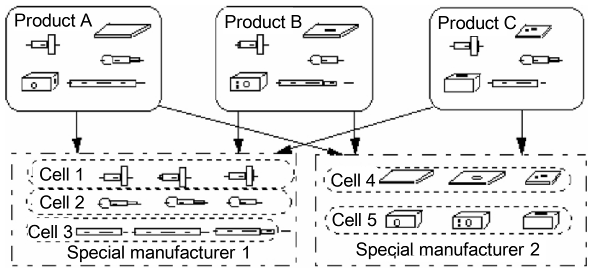


Figure 1: Demonstration of group analysis in large scale.

2 REQUIREMENTS OF THE PLIBEI PLATFORM

Figure 2 sketches the proposed framework and working mechanism of the PLIBEI platform. From Figure2, we can find that the core of PLIBIE platform is mainly composed of two parts: web-based parts library and the part search engine.

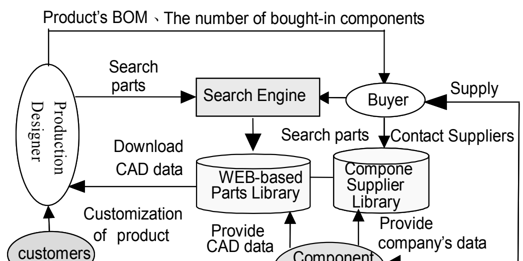


Figure 2: The framework and working and working mechanism of the PLIBEI platform.

Web-based parts library: Parts information (such as CAD data, drawing file, price and date of delivery etc.) which provides by parts suppliers was stored in the web-based parts library. And the parts users acquire these information and data via Internet. So compare to ordinary parts library, web-based parts library has some special requirements: (i) Providing unambiguous parts description method to parts suppliers; (ii) Integrating parts information and data which come from various suppliers automatically; (iii) Variant design in real time: when a user orders a model for a specific part, the web-based parts library can generate particular model in real time instead of storing thousands upon thousands instance models. This allows for significant reduction in cost related to initial CAD creation and future product updates, changes, and deletions.

Part information search engine: Part search engine is the bridge of information exchange and integration. It is differ from the common engine. A part search engine should provide semantic feature search ability besides the keywords search and industry classification search. For example, when a part buyer query a hexagonal machine screw, he should specify what kind of product (hexagonal machine screw), and the characteristic properties (total length, threaded length, diameter, coating etc.) The buyer and the supplier should agree on term of this product.

From previous analysis, we believe with ontologies there's now a model at hand to fit for these requirements. Ontologies serve as a means for establishing a conceptually concise basis for communicating knowledge for much purpose. In our scenario, we restrict our attention to domain ontologies-PLIB ontology. In next sections, we'll introduce what is PLIB ontology and how the PLIB ontology can meet these requirements.

3 PLIB ONTOLOGY

PLIB ontology is modeled according to the PLIB (meta) dictionary(ISO13584-42)[3]. In PLIB ontology, component is highly structured. Components are gathered in parts families that are represented by classes. Parts families are organized as a simple hierarchy which looks like a tree. Properties of parts family are defined at each level, and also applied to lower levels. Each Parts family is described at two level of abstraction: the first level is the semantic description of the abstraction. It includes the part name and meaning, part attribute names and meaning etc. Both the parts family name and the part attribute names are associated with a code (GUI-Globally Unique Identifier), which identifies the abstraction, and with a descriptor, which defines the abstraction. These contents constitute the semantic dictionary of parts families. The second level is the logical description of the abstraction. At this level, the set of parts belonging to one parts family was specified. And the programs used to produce the geometric view or other external files were specified too. The dual description method provides the parts library end user an "understandable part".

4 PLIB ONTOLOGY BASED PARTS LIBRARY

4.1 The information model of parts library

In parts library, a part is not only a set of attribute values, but also a human understandable abstraction .As we relate in section 3, the description of part information adopts the dual description method in parts library, so the information model of web-based parts library should include two layers (as shown in Figure.3): ontology level (semantic level) and instance layer The data in a parts library is stored in the form of database. According to the characteristics of database, the instance layer is divided into schema layer and data layer.

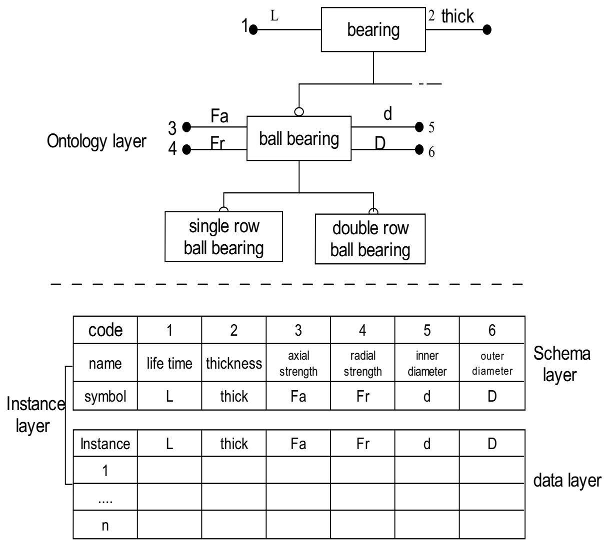


Figure 3: The information model of web-based parts library.

So, the information model of web-based parts library can be described as: $PLIBDB = \{O, I_{sch}, I_{data}\}$,where:

- O is an PLIB ontology. O is defined as

$O =:< C, H_c, P, ApplicP, CharP >$,where:

C is the set of classes used to describe the concept of a given domain, $C = \{c_1, c_2 .... c_n\}$. Each class was defined at least by a code and the explanation for the class.

H_C is a function described the class hierarchy relationship. And here the relationship is subsumption. $H_c(c_i) \in C \quad c_i \in C$,

The value of this function is the class c_i's direct subclasses.

P is the set of properties of a given ontology class.

ApplicP is the set of applicable properties of the given ontology class $ApplicP(c_i) \subset P \quad c_i \in C$

CharP is the set of attributes of a part of the given ontology class $CharP(c_i) \subset ApplicP \quad c_i \in C$

- I_{sch} is a function that describe the given class's Instance schema. $I_{sch}(c) \subset ApplicP(c)$ associates to each ontology class c of C the properties which are effectively used to describe the instances of the class c.
- I_{data} is the collections of the instance of class c that is described by $I_{sch}(c)$.

When the class is an abstract parts family (no-leaf class $H_C(c) \neq \varnothing$), there only I_{sch} exits and

$I_{data} = \varnothing$. $I_{sch}(c) = ApplicP(c) \cap (\cap I_{sch}(H_c(c)))$.

When the class is a simple parts family,

$I_{sch}(c) \subset Applic\ P(c)$, $CharP \subset I_{sch}(c)$.

And $I_{data}(c)$ exits.

4.2 Realization of the unambiguous description of parts

To realize the unambiguous description of parts, the manager of PLIBEI platform can provide an ontology library (see Figure 4).The ontology library is composed of two parts: PLIB ontology library and semantic feature library. Standard PLIB ontology and shared PLIB ontology constitute the PLIB ontology. The standard PLIB ontology is released by the standardization organizations, such as the fasteners ontology (ISO13584-511) and the shared PLIB ontology is provided by the industry association or the PLIBEI platform. Because the typical structure element (such as key, slot, guide etc.) is not included in the PLIB ontology, a semantic feature ontology library was built to solve this problem. The semantic feature ontology library is important to implement the group analysis later. Here we called this ontology library as Extended PLIB ontology library. In our practice, we use OntoEdit[4] to create the ontology library.

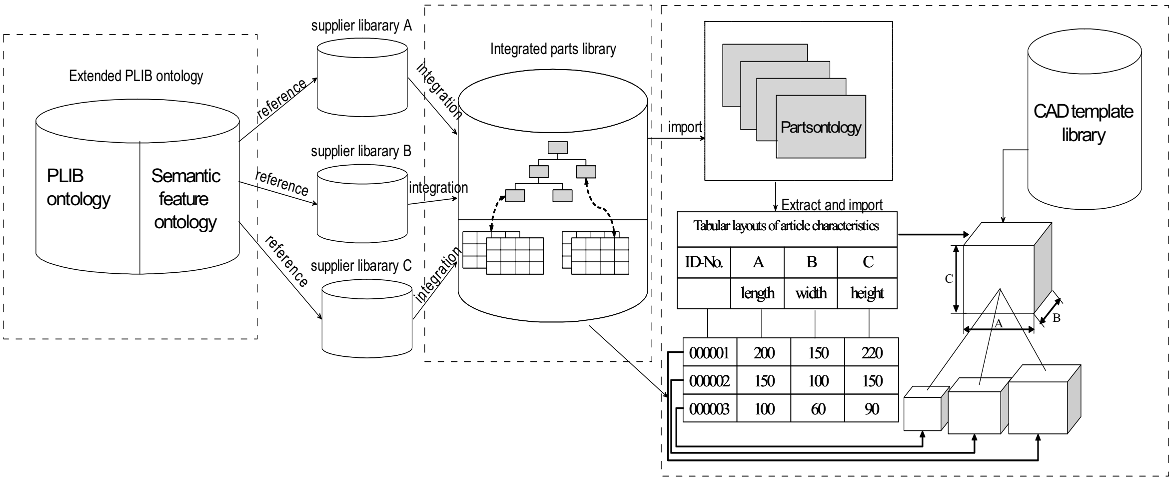


Figure 4: The realization process of web-based parts library.

For standard parts supplier, we propose him reference the standard ontology directly without any modification.

For common parts supplier, we propose him reference the shared ontology as much possible. If the shared ontology can not meet his description requirement, he can extend the ontology. The extended classes and attributes can be added in the ontology library as a shared ontology. To avoid the redundancy of information, the reference frequency should be recorded and Low frequency records will be drove out of the ontology library.

4.3 Realization of the integration

The integration of numerous supplier libraries includes three levels integration (ontology layer, schema layer and data layer)according the information model of parts library. To realize the integration of ontology layer, the mapping between local ontology and global ontology is needed (see Figure.5).The integration algorithm has been discussed in detail in another paper. The mapping can be implemented by using the OntoMap,the plug-in of the OntoEdit.

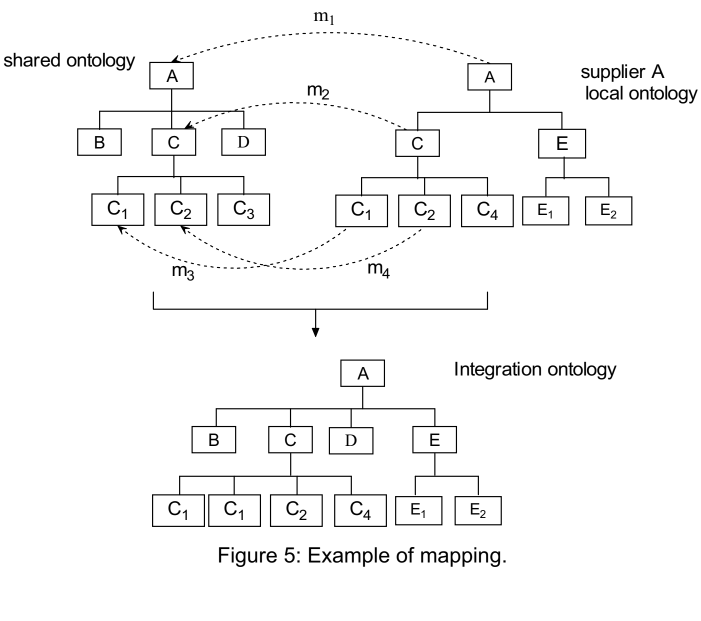


Figure 5: Example of mapping.

4.4 Variant design in real time

The realization process of "variant design in real time" has shown in Figure 5. We could import the part ontology into a table by using OntoEdit. An SQL import has already been realized in OntoEdit. OntoEdit creates an automatically connection to the database by the dbaccessuser-Built-in[4]. For example, we can extract a part class's attribute (mainly the part characteristic) and uses these attribute define the tabular layouts of article characteristics. Using tabular layouts of article characteristics to realize variant design has been researched in our group. Concrete realization method was related in [5] [6].

5 PART SEARCH ENGINE

Traditional part information retrieval mainly adopt keywords search or product classification search .In order to satisfy different user's multi-hierarchy requirements, the semantic search model is built:

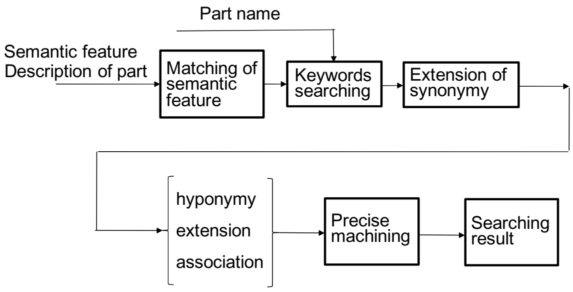


Figure 6. Multi-hierarchy search model

The instance information which has the synonymy, hyperonymy and hyponymy semantic of the user's specified concept can be retrieved by adopting this model. The ontology readwrite API (Jena Ontology API) can fulfill the search of ontology library. The ontology inference engine of Jena was used to implement the reasoning process (see Figure.6).

6 CONCLUSIONS AND FUTURE WORK

The market will be more open and chaotic in 21 century. As a prospective application technology, GGT is helpful for changing the "full-inclusive" or "small-inclusive" phenomenon in Chinese enterprises, building a lifeful component production system and promoting enterprises' international competitive ability. The key for the application and popularization of GGT is to build an open parts information exchange and integration platform (for short: PLIBEI platform).In this paper, the framework and working mechanism of this platform is introduced. The two core parts (web-based parts library and part search engine) of this platform is pointed out and analyzed. Through requirement analysis, we think PLIB ontology is fit for the solution of these requirements (unambiguous description, integration automatically, variant design in real time etc). Then a parts information model is built based on PLIB ontology. Because the function feature and structure feature is not included in PLIB ontology, an Extended PLIB Ontology library is built to solve this problem. The Extended PLIB ontology library is composed of two parts: PLIB ontology based on the PLIB standard and the semantic feature description ontology. At last, the realization method is introduced. Now, the prototype system has been tested in our lab.

In the next phase of the project, we are planning to (1) realize the 3D CAD model browse via internet; (2) use the Extended PLIB ontology library to realize group analysis and group manufacturing in large scope.

7 ACKNOWLEDGMENTS

Project supported by the National High Technology Research and Development Program of China (863 Program)(2006AA04Z130)and National Nature Science Foundation of China (50475185, 50575047).

8 REFERENCES

[1] GU Xin-jian and QI Guo-ning, (1996) (1997) Global integrated manufacturing and Great Group Engineering(1)(2) .In Factory construction and design, No.6 pp.28–32, No.1 pp.30–32.

[2] Liu, D., Wang, ZH., Xie, QS. and Gu, XJ. (2003) Internet-based great group technology. In Third international conference on electronic commerce engineering-digital enterprises and non-traditional industrialization. pp.423–425.

[3] Pierra, Guy. (2003) Context-explication in conceptual ontologies: The PLIB approach. In Proceedings of the 10th ISPE International Conference on Concurrent Engineering Research and Application, Enhanced Interoperable Systems. pp. 243–253.

[4] Y.Sure, Erdmann, J.Angele, S.Staab, R.Studer,and D.Wenke. (2003) OntoEdit: Collaborative ontology development for the semantic web.In Horrocks and Hendler [HH02],pp.221–235.

[5] WU Yang-Dong, YU Jun-He, SU Shao-Hui and LU Yu-Jun. (2004) Web-based variant design system. China mechanical engineering. Vol.15,pp1720–1724.

[6] GU Qiao-Xiang, SU Shao-Hui, YU Jun-He and LU Yu-Jun. (2004) Research on product variant design based on the tabular layouts of article characteristics. China mechanical engineering. Vol.15,pp1713–1716.

Geometrical Properties of Paper Spring

Cheong Chew Min[1] and Hiromasa Suzuki[2]

[1,2] Department of Precision Engineering, Graduate School of Engineering, The University of Tokyo, Japan

Abstract

Recent advances in the research of mathematics and science in origami is rapidly transforming the ancient art of paper folding into a challenging field of study. This paper describes a deployable spring series which is an extension of the original paper spring created by Jeff Beynon. The model was studied and researched through calculations, models and visualizations. We will illustrate and discuss the basic concept and seek possible industrial and space applications of this deployable structure. The original model was also modified for better mechanical applications. The applicability was evaluated at a conceptual stage.

Keywords:

Origami; Deployable Folding; Spring Structure; Energy Absorption; Support System

1 INTRODUCTION

Origami has long been regarded as a form of art mainly as an activity for children. However, over the past few decades, origami has gained the attention of mathematicians and scientists to study the complexity of foldings and to seek possible applications of origami in various fields, transforming the ancient art of paper folding into a challenging field of study. Miura-ori was one of such prominent applications of origami in the folding of a solar sail in a space flight unit [1] as well as a map folding method. Origami has also found its life saving role in airbag folding [2] and its biomedical application as stent grafts [3].

The purpose of this research is to investigate origami concepts applicable to deployable structure and to seek possible implementations in industrial and technological fields. In this paper, we aim to introduce the paper spring structure as depicted in Figure 1 which was created by Jeff Beynon and first described by Tomoko Fuse in [4].

The paper spring structure consists of the repetition of polygonal spring module longitudinally. Figure 2 shows the crease pattern of a 6-sided spring with 4 modules. The mountain fold is indicated by dash-dotted line while the valley fold is denoted by dashed line.

Possible implementations of this paper spring structure are proposed in Section 2, while Section 3 discusses the mathematics underlying the paper spring structure and its geometrical properties. Section 4 shows the experimental observations. Discussions on the mechanical applications can be found at Section 5.

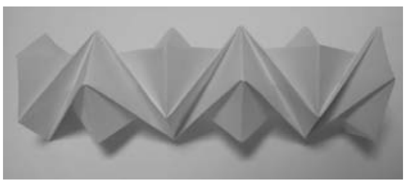


Figure 1: Paper spring by Jeff Beynon

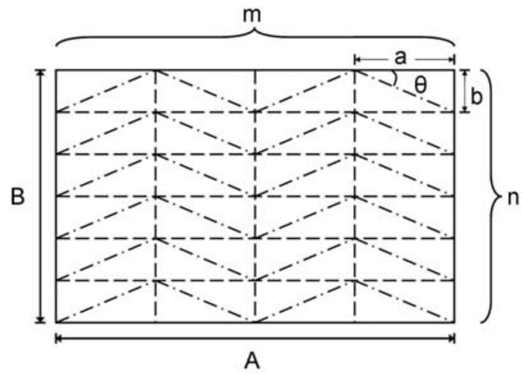


Figure 2: Crease pattern of a 6-sided paper spring

2 IDENTIFICATION OF RELEVANT APPLICATIONS

Comparisons are made between the similarities in conventional applications of folded structures and the paper spring structure.

2.1 Shock Absorbing System

Recent industrial application of folded structures includes the design and construction of high energy absorbing pads for high speed airdrop. The paper spring structure could be useful in industrial packaging for a shock absorbing system as an alternative for the honeycomb structure and Chevron pattern structure as discussed by Elsayed and Basily in [5].

Folded structures offer an ecology friendly packaging solution which is flexible, lightweight and acts as a cushion layer for impact absorption compared to conventional foam cushioning.

Folded spring structure has another advantage over honeycomb structure. The energy absorbing mechanism in a honeycomb structure is due to the progressive buckling which causes the whole structure to deform under pressure [6]. On the other hand, spring structure is deployable by nature, which exhibits the energy-absorbing capability and could be reused as impact merely causes the structure to be compressed.

2.2 Spring structure as deployable truss

Trusses or masts are being used in space applications as supporting members for substructures or instruments. The idea of deployability is of great consequences in space development as the structures have to be transported from ground to space in packaged configurations. Similar concept can be found in the variable geometry truss that has been proposed by Miura, Furuya and Suzuki in [7]. The model can be deployed through a sequential mode transformation whereby the whole truss can be classified into three zones during deployment, namely the deployed zone, the transient zone and the folded zone as depicted in Figure 3.

On the contrary, spring structure can easily be deployed by applying horizontal compression forces from two opposite directions as illustrated in Figure 4. The horizontal forces can be applied to any inner module and the whole structure would deploy simultaneously in both upward and downward directions.

3 PAPER SPRING STRUCTURE

The paper spring structure is folded from a rectangular piece of paper, without cutting and gluing. The crease pattern consists of regular arrangement of uniform right triangles. Figure 2 shows the crease pattern of a 6-sided spring structure consisting *m* vertical columns of 2*n* triangles, where

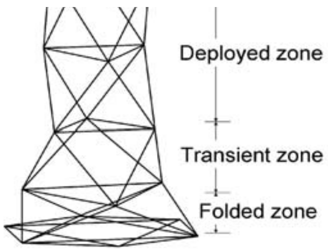


Figure 3: Sequential mode of deployment [7]

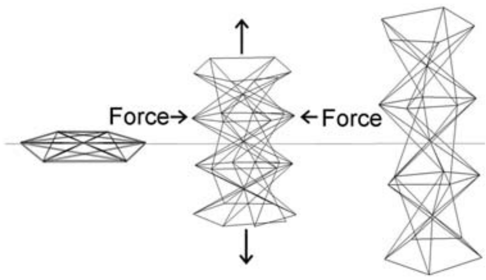


Figure 4: Simultaneous deployment of spring structure

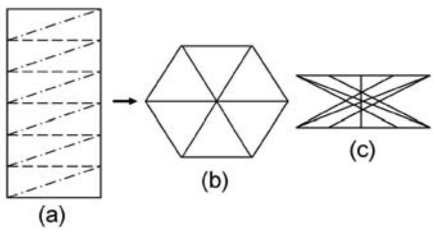


Figure 5: A module of the paper spring

n and *m* denote the vertical and horizontal divisions of the pattern respectively. The spring structure is valid for arbitrary *m* while *n*>2.

Let θ be an angle of the triangle with its opposite and adjacent denoted by *a* and *b* respectively (see Figure 2). This θ is an important parameter which is hereby known as the fold angle θ of a paper spring. It should satisfy

$$\theta = \pi/n \,. \tag{1}$$

Given *n*, *m*, θ, and the width *B*, other parameters of the pattern including *b*, *a*, and the length *A* can be defined as follows.

$$b = B/n \tag{2}$$

$$a = b\cot\theta \tag{3}$$

$$A = ma \tag{4}$$

Since *B* is defined by the size of the paper and *m* can be an arbitrary number, *n* and θ are the essential parameters.

A paper spring is made up of repeating layers. One layer is made up of one column of 2*n* triangles. We call this layer a module. The crease pattern of a module is shown in Figure 5(a). Figure 5(b) and Figure 5(c) shows the top view and the side view of a folded module respectively.

And though it is not obvious from physical model of paper spring structure, we found that even with thin paper it has self intersections, that is, the structure must be deformed to be deployed.

The paper spring structure has a compact center for every module. This is due to the fact that the midpoint (or nearby point) of hypotenuses of all the triangle units (or fan) of a module coincide at a point, forming the center of the module. Due to the impenetrable nature of material, self-intersection of paper is impossible.

Deformation resulted in some complex behaviour in the paper spring which we have not fully analyzed. Hence, in the following subsections, we describe our observations based on the physical experiments. The existing behaviour is not suitable to be used in mechanical system. We will discuss this issue in Section 5.

4 EXPERIMENTAL OBSERVATIONS

4.1 Deployable Limit

Physical experimental on the 6-sided paper spring structure folded from 60.4g/m^2, 90.6 g/m^2, 124.4 g/m^2, 153.4 g/m^2, 204.6 g/m^2 papers show that the structure is deployable to its maximum limit with one complete overlapping fan. The top view of a fully deployed 6-sided spring is illustrated in Figure 6(b). Beyond this limit, severe deformation would occur, causing the paper to tear. Figure 6(a) shows the top view of a 6-sided spring structure before deployment.

For spring models with different number of sides, their respective deployable limits and heights vary when fully deployed. More specifically, while *m* remains constant, the number of possible overlapping fans increases with the increase in *n*. Hence, the overall height of the spring would also increase.

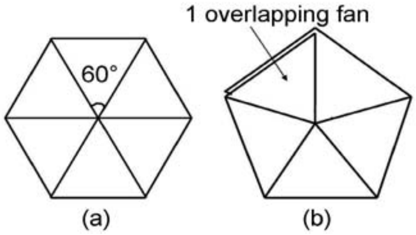


Figure 6: Different views of a 6-sided spring

4.2 Deviations of Fans

Due to the deformations in the center, there are deviations (see Figure 7) for every module, resulting from the thickness accumulated by the overlapping layers of paper. Meanwhile, the width of the spring shrinks in a spiral manner as the spring is being deployed by the horizontal forces acting on it.

4.3 Curvilinear Deployment in Paper Spring Structure

Physical models of the paper spring structure shows significant inclination as the model is being deployed to its maximum deployable limit. The spring structure is deployable until the limiting fans (as shaded in Figure 8(a)) pull each other to the limit. In the case of using paper as a folding medium, the paper will tear beyond this limit. This limiting factor also causes the whole structure to bend as shown in Figure 9(a).

5 DISCUSSIONS FOR MECHANICAL APPLICATIONS

5.1 Comparison with Para-cylinders

Similar studies of deployable structure known as para-cylinders have been done by researchers such as Kobayashi and Nojima [8-10]. The differences between spring structure and para-cylinders in 6-sided cases are being summarized as Table 1. While all the four listed structures are deployable, spring structure has the advantage of being easily deployed. The three para-cylinders can be deployed and compressed by a horizontal force acting over the plane while the spring structure could be deployed by horizontal forces acting from two opposite points.

5.2 Avoiding Curvilinear Deployment

The curvilinear of the spring structure must hinder mechanical applications compared to para-cylinders and hence should be avoided. The inclination of the spring structure can be avoided by modifying the initial model. For instance, in the case of a 6-sided spring, the removal of either one of the limiting fans as shaded in Figure 8(a) would result in a model without inclination when fully deployed as shown in Figure 9(b). The modified crease pattern is shown in Figure 8(b).

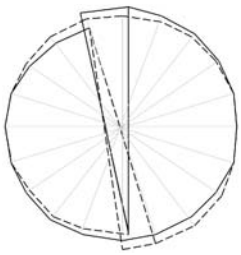


Figure 7: Deviation of fans in a 20-sided spring module

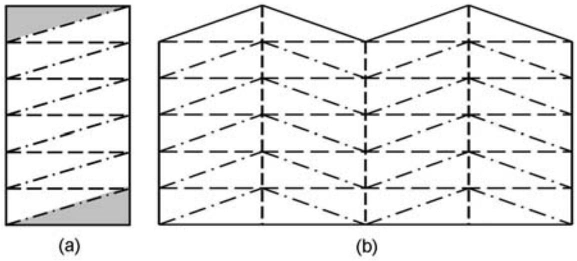


Figure 8: Modified crease pattern

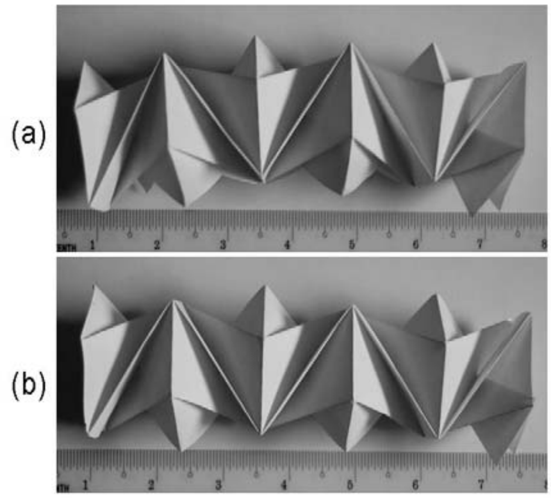


Figure 9: Model (a) before and (b) after modification

6 CONCLUSIONS

The proposed initial and modified models presented here have shown that the innovative concept of single piece origami spring model could be useful in industrial and space applications. The spring structure also exhibits advantages in geometrical properties compared to similar deployable structures as it is instantly deployable. While the initial model deploys in a curvilinear manner, the inclination may also be removed by the modification suggested in Section 5.2. This modification is crutial for mechanical applications where the inclination issue is critical. Further analysis on the spring structure is left for future work.

7 REFERENCES

[1] Miura, K., Natori, M., 1985. 2-D Array Experiment on Board a Space Flyer Unit. Space Solar Power Review, Vol. 5, No. 4, 345-356.

[2] Kuribayashi, K., Tsuchiya, K., You, Z., Tomus, D., Umemoto, M., Ito, T., Sasaki, M., 2006. Self-Deployable Origami Stent Grafts as a Biomedical Application of Ni-rich TiNi Shape Memory Ally Foil. Materials Science and Engineering A, Vol. 419, No. 1/2, 131-137.

[3] Cromvik, C., 2007. Numerical Folding of Airbags Based on Optimization and Origami. MA Thesis, Chalmers University of Technology and Göteborg University, Sweden.

[4] Fuse, T., 1992. Lets Fold Spirals, Chikumashobo.

[5] Elsayed, E. A., Basily, B. B., 2004. A Continuous Folding Process for Sheet Materials. International Journal of Materials and Product Technology, Vol. 21, No. 1/2/3, 217-238.

[6] Ogasawara, N., Shiratori, M., Miyahara, S., Yu, Q., 1997. energy Absorption by Honeycomb Structure (1st Report, Progressive Buckling under Quasi-Static Condition). JSME International Journal, Series A, Vol. 63, No. 608, 118-123.

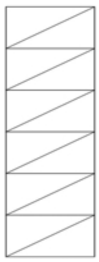
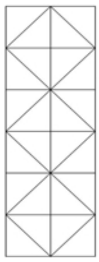
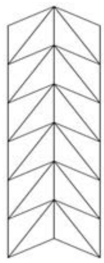
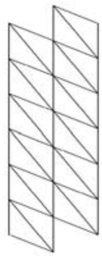
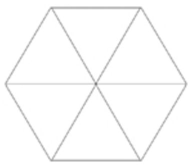
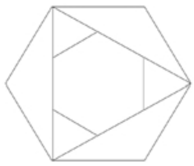
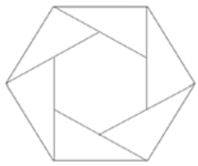
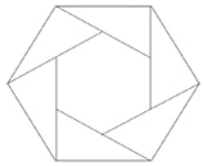
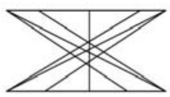
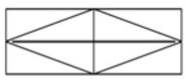
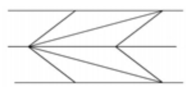
[7] Miura, K., Furuya, H., Suzuki, K., 1985. Variable Geometry Truss and Its Application to Deployable Truss and Space Crane Arm. Acta Astronautica, Vol. 12, No. 7/8, 599-607.

[8] Fujita, A., Daimaruya, M., Kobayashi, H., 2002. Folding and Unfolding Characteristics of Folding Para-cylinder. The Japan Society of Mechanical Engineers, Vol. 42, 172-173.

[9] Tao, X., Wu, Z., Hagiwara, I., 2004. Modification of Cylindrical Origami Structure by Subdivision and Origami Techniques. The Japan Society of Mechanical Engineers, Vol. 6, 73-74.

[10] Kobayashi, H., Daimaruya, M., Nojima, T., 2003. Unfolding Characteristics and Change of Crease Length of Folding Para-cylinder. The Japan Society of Mechanical Engineers, Vol. 2003, 549-550.

Table 1: Comparison of deployable structures

6-sided	Spring	Para-cylinder	Para-cylinder	Para-cylinder
Basic Pattern				
Aerial View				
Side View				
Unit Triangle	Right angle triangle	Right angle triangle	Isosceles triangle	Isosceles triangle
Center Condition	Compact	Hollow	Hollow	Hollow
Deployed inclination	Yes	No	No	No
When pressed horizontally	Deploy	Deform	Deform	Deform
Deploy/ compress rotation	Neutral rotation	No	No	Single direction
Radius shrinking when deployed	Yes	No	No	No
Control over model deployment	2 points	Plane	Plane	Plane

Application of Axiomatic Design to Develop a Lean Logistics Design Methodology

Cleber Favaro[1]

[1] Department of Manufacturing, School of Mechanical Engineering, State University of Campinas, Brazil

Abstract

Supply chain design is a complex task that involves a network of companies and disciplines. Different practices have been applied without a systemic view generating modest results. This research proposes a lean logistics design decomposition supported by axiomatic design theory. The decomposition allows the separation of system objectives and means of achievement. Moreover it connects the high level objectives to the low-level activities and also shows their interrelationship. It guides the designer to achieve simultaneously the business fundamentals of quality, on time delivery, lead time, cost and investment effectiveness. Two case studies are presented to illustrate the framework application.

Keywords:

Axiomatic Design; Lean Thinking; Logistics

1 INTRODUCTION

Logistics system design has been a challenging task to every corporation since the beginning of the twentieth century. From the vertical integration of Ford Model T to the horizontal and integrated supply chain implemented by Toyota, the questions remain the same: how to achieve superior performance and establish a competitive advantage with customers demanding shorter product life cycles, superior levels of quality and more frequent deliveries.

Parker [1] states that superior supply chain performance only can be achieved through superior design and highlights that acting strictly in management restricts the impact in the overall system. Fine [2] goes further defining the supply chain design as the ultimate competitive advantage.

According to Laseter and Oliver [3], in spite of the importance of the discipline, the concept of supply chain has not been correctly explored by companies and therefore the real benefits are modest.

Heckmann et al. [4] conducted a global study and concluded that companies are not satisfied with their logistics performance. Another important finding is that companies that restructured their supply chain operations achieved results from 35 to 55% better compared to companies that only conducted adjust in isolated processes.

Croom et al. [5] and Tan [6] identify a lack of systemic knowledge on supply chain design. Both understand that the literature is restricted to a collection of issues about manufacturing, purchasing, inventory management or operations.

The difficulty to achieve success in the supply chain may be explained due to the absence of complete methodology to design the system.

The effective logistics design should start with the strategic goals of the system and deploy them to operational levels for implementation.

The deployment requires collaboration to create, operate and maintain processes that crosses the borders of departments and plants, increasing the complexity of designing.

It is possible to describe the problem as follows: there is a lack of methodology to guide designer through the stages and levels of lean logistics design.

The objectives of this work can be defined as:

- Present a present a methodology for designing a lean logistics system from discrete manufacturing companies applying the axiomatic design theory;
- Apply the model to evaluate and compare two similar automotive plants.

The paper is organized as follows. Section 2 reviews the literature about lean logistics. Section 3 introduces the axiomatic design theory and its fundamentals as proposed by Suh [7], [8].

The development of the Lean Logistics Design Decomposition (LLDD) is presented at the section 4. In the following section the LLDD supports the assessment of two plants of different automotive suppliers.

The results and conclusions are summarized at section 6, where future works are also proposed.

2 LEAN LOGISTICS

Lean Logistics can be defined by the extension of Toyota Way to the supply chain through a set of techniques from the strategic to operational level in order to maximize flow while reducing waste. [9], [10].

In spite of Toyota's business beliefs and methods are summarized in principles that were born with the company's founder, Sakichi Toyoda in 1935, only in late 1980s business community turned eyes con it [11].

In the years following the release of the book "The Machine that Changed the World" [12] many publications developed lean supply models [13], [14], [15], [16]. Those works were based on empirical research and described lean practices at a strategic level, focusing mainly on sourcing process.

The common practices identified by the authors include:

- Reduced number of suppliers;
- Partnership and long term contracts;
- Suppliers associations and knowledge sharing;
- Mutual problem-solving;
- Quality assurance and zero defect;
- Just-in-time delivery system;
- Target costing and kaizen efforts.

Lamming [13] and Hines [14] suggest roadmaps for creating lean supply according a set of concepts, not proposing a sequence of implementation and moreover always keeping the debate at high design levels.

Extended value stream mapping (EVSM) was presented as a qualitative tool for logistics design supported by lean principles as stated by Jones & Womack [17].

The method starts tracing the existing flows of materials and information from a firm-centric view and then to design a future streamlined state.

EVSM had shown to be a useful tool for preliminary design while allows to define interfaces between subsystems.

Baudin [9] provides a guide for implementing lean tools in both, information and material flows, going into the details necessary to effective implementation of sub-systems in Lean Logistics. Nevertheless, the author does not purpose a methodology for overall system design what may lead to local optimizations.

The Bridge Logistics Model [10] developed a framework merging lean techniques with the six sigma approach. It concentrates in known tools and techniques in order to define how to improve logistics process design.

This model suggests flow, capability and discipline as the high level critical success factors that are deployed in lower level elements. This decomposition is done without a clear method and therefore the elements are a mix of "what to do" and "how to do" and as the previous models does not establish a sequence for implementation.

The current status of the lean logistics literature reviewed may be characterized by:

- Not covering all the levels of design, not going into detail in strategic or in the implementation level;
- Not properly connect the different levels of design;
- Not defining a clear sequence for implementation;
- Not distinguishing requirements from means of achievement;
- Not making use a formal methodology for designing.

The Axiomatic Design theory will be presented in the following section provides a method that fulfills those issues regarding Lean Logistics Design.

3 AXIOMATIC DESIGN

The axiomatic design theory was developed at Massachusetts Institute of Technology during the mid-1970s. It proposes that there are a set of principles that provides a scientific basis for designing engineering systems, by supporting designers with logical and rational process and tools [7].

The concepts of design domains, the design axioms and the decomposition process are the key fundamentals to axiomatic design framework [8].

The design process consists of interplay among four different domains: customer, functional, physical domain and process domains.

In the customer domain the attributes that the customer is looking in a product are identified. The following step is translating those customer attributes (CA) in terms of functional requirements (FR) and constrains (C), what happens in the functional domain. In the physical domain those FRs and Cs are mapped in design parameters (DP). At last, the process to produce the desired product is defined in the process domain through process variables (PV).

The Figure 1 represents the described the process flow through the four domains.

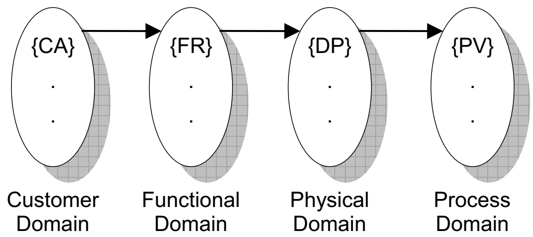


Figure 1: The Four Design Domains

A functional requirement (FR) is defined as the minimum set of independent requirements that completely characterize the functional needs of the product [8]. By definition, each FR is independent of every FR at the time they are established.

The design parameters (DP) are key physical variables that characterize the design that satisfies the specific FR.

Suh [7] posted two axioms that govern design process, based on the examination of common elements of successful projects.

The first axiom, named Independence Axiom, states that the independence of the FR must always be maintained. The second axiom, the Information Axiom, supports the selection of the best design among different designs that satisfies the first axiom. It defines that the best design is the one with the smallest information content [7].

The relationship among the vector of FRs and the vector of DPs is given by the design matrix [***A***] as shown at the Equation 1.

$$\{\boldsymbol{FR}\} = [\boldsymbol{A}] \cdot \{\boldsymbol{DP}\} \quad (1)$$

The elements of the matrix [***A***] can be represented using X or 0 to shows if a DP interferes in a FR or not. The Equation 2 exemplifies the concept.

$$\begin{Bmatrix} FR_1 \\ FR_2 \\ FR_3 \end{Bmatrix} = \begin{Bmatrix} X & 0 & 0 \\ X & X & 0 \\ X & X & X \end{Bmatrix} \cdot \begin{Bmatrix} DP_1 \\ DP_2 \\ DP_3 \end{Bmatrix} \quad (2)$$

By analyzing the design matrix it is possible to define if the first axiom is fulfilled or not by the proposed design [8]. Three types of design are characterized by the design matrix [7].

- **Uncoupled Design**: if the design matrix is diagonal, each DP affects only it correlated FR, satisfying the Independence Axiom.
- **Coupled Design**: when the design matrix is triangular (as shown at Equation 2), some DPs affect more than one FR. In this case the Independence Axiom is only fulfilled if a proper sequence of implementation of the DPs is followed.
- **Decoupled Design**: if the design matrix is not diagonal or triangular, the design is defined as coupled. It is impossible to adjust one DP without affecting many FRs and therefore it is conceptually considered a poor design.

Defining the first level FRs and the correspondent DPs in most of the cases does not give details to implement or create the product. Therefore it is necessary decompose the top FRs in lower levels FRs and defining its respective DPs. During this decomposition the designer moves from the functional domain to the physical domain and then backward until achieve the correct level of details. This process of creating a hierarchy is called zig-zag decomposition [8].

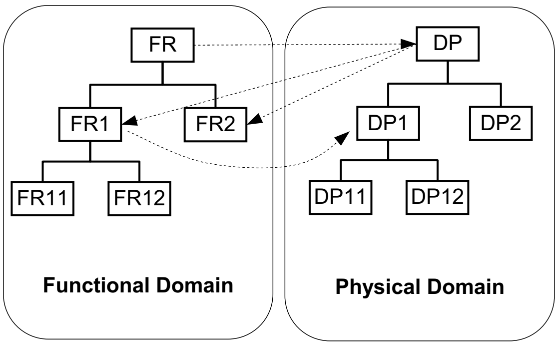


Figure 2: Design Decomposition Process

The advantages of applying Axiomatic Design Theory to develop large system design as LLDD includes:

- Clear separation of the objectives and the means of achieve them;
- Allow connecting high level to low level goals;
- Define the relationships among the elements of the project;
- Establish a common language to communicate the lean logistics design;
- Allow integrating the current techniques and tools in a total framework.

The next section will show the application of Axiomatic Design to create a LLDD.

4 LEAN LOGISTICS DESIGN DECOMPOSITION (LLDD)

4.1 Introduction

The concept of applying axiomatic design to support the development of a lean enterprise was previously addressed by Cochran et al [18] and Bocanegra [19] for Manufacturing and Product Development respectively.

Schnetzler et al. [20] proposed a Supply Chain Design Decomposition, based on the SCOR Reference Model [21] and without influence of lean principles.

The LLDD presented in this section objectifies to put each lean supply tool or practice together in order to create a consistent design methodology.

4.2 Top Level Decomposition

The top level strategic functional requirement (FR1) chosen is to maximize the economic value added (EVA) of the center-plant. According Young and O'Byrne [22], EVA is based on the notion of economic profit which states that wealth is created only when a company exceeds all operating costs and also the capital cost. The LLDD is the correspondent DP1.

The second level of FRs is originated from the decomposition of FR1. Eight variables involved in the EVA calculation allows the derivation of the FRs listed below.

- FR1.1: Maximize sales revenue
- FR1.2: Minimize cost of goods sold due to logistics
- FR1.3: Minimize selling, general and administrative expenses (supporting processes)
- FR1.4: Minimize inventory
- FR1.5: Minimize trade receivables
- FR1.6: Maximize trade payables
- FR1.7: Minimize intangible assets
- FR1.8: Minimize fixed assets

The following DPs are selected to fulfill the second level FRs.

- DP1.1: Maximize customer's satisfaction due to logistics performance
- DP1.2: Minimize non-value added tasks in the order fulfillment process
- DP1.3: Convert support necessary non-value added tasks in unnecessary to eliminate them
- DP1.4: Maximize inventory accuracy
- DP1.5: Maximize the age of receivables
- DP1.6: Minimize the age of payables
- DP1.7: Minimize use of softwares to manage the logistics processes
- DP1.8: Select assets based on long term product strategy

The design matrix show at Equation 3 represents the second level of LLDD.

$$\begin{Bmatrix} FR_1 \\ FR_2 \\ FR_3 \\ FR_4 \\ FR_5 \\ FR_6 \\ FR_7 \\ FR_8 \end{Bmatrix} = \begin{Bmatrix} X & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ X & X & 0 & 0 & 0 & 0 & 0 & 0 \\ X & X & X & 0 & 0 & 0 & 0 & 0 \\ X & 0 & 0 & X & 0 & 0 & 0 & 0 \\ X & 0 & 0 & X & X & 0 & 0 & 0 \\ 0 & 0 & 0 & X & 0 & X & 0 & 0 \\ X & X & X & X & X & X & X & 0 \\ X & X & X & X & X & X & X & X \end{Bmatrix} \cdot \begin{Bmatrix} DP_1 \\ DP_2 \\ DP_3 \\ DP_4 \\ DP_5 \\ DP_6 \\ DP_7 \\ DP_8 \end{Bmatrix} \quad (3)$$

The zig-zag process is done progressively until the framework contains information enough for design implementation.

The total LLDD framework is composed by a set of about 150 pairs of FRs-DPs which are detailed in Favaro [23]. It is organized according nine logistics fundamentals whose are explained in the following sub-section.

4.3 Lower Level Decomposition

The FR 1.1 is deployed in the following fundaments: supply chain structure, quality, on time delivery and lead time.

The definition of the structure of supply chain is the first step in the designing process. It embraces the make or buy process, the strategic supplier segmentation, supplier monitoring and his knowledge development.

The second fundamental of quality include all aspects to produce quality at the supplier and assure that handling, the external and internal transportation, as well the storage will not affect the requested specification of the finished good.

Reducing the mean lead time and delivering on time minimizing the variation are the complementary fundamentals of the FR 1.1.

In a similar way presented at the MSDD [18], the variation in the logistics process may be minimized by identifying the disruptions (FR 1.1.3.1) and acting in order to minimize stoppages in the flow (FR 1.1.3.2).

Lead Time compression is achieved by delay reduction in the information and material flows from the supplier to the customer. The low level functional requirements embrace reducing delay in supplier production, incoming, internal and external transportation, storage and shipment to customer.

The FR1.2 and FR 1.3 represent cost and expenses, respectively. The costs are related to the order fulfillment process including freight, material and labor costs. Furthermore, the expenses are connected with the supporting administrative functions.

The FR 1.4 advocates for reducing inventory. Most of the DPs related with lead time minimization impacts directly in this requirement. The selected DP 1.4 complements the need for inventory reduction by increasing its accuracy.

The payment terms to the customers and from the supplier are treated by the FR 1.5 and FR 1.6. The concept is to reduce the capital employed due to commercial agreements.

At last, the FR 1.7 and FR 1.8 represent the need for minimizing the intangible and fixed assets, requisites that are fulfilled by the general design of the system and a strategy of investments based on the product life-cycle.

The Figure 3 shows the sequence of implementation for the logistics fundamentals, as concluded after the development of LLDD using the axiomatic design theory

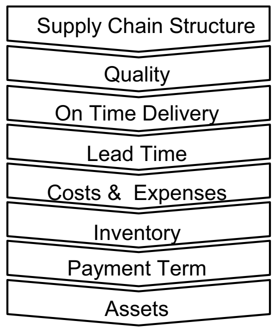


Figure 3: Sequence for implementing LLDD fundamentals

The LLDD can be applied to design complete new systems or as a roadmap to assess and improve existing systems.

The section 5 presents two case studies where LLDD is applied to assess the current condition of the companies.

The assessment tool was created with the intermediate level pairs of FR-DP. For each pair some statements are developed and its fulfillment evaluated according a five points Likert scale (strongly disagree, disagree, neither agree nor disagree, agree, strongly agree).

5 CASE STUDY

5.1 Introduction

The two plants studied in this section are from global auto-parts suppliers of heat, ventilation and air-conditioning (HVAC) systems for passenger cars located in Brazil. The companies will be treated as Plant A and Plant B from now on.

The manufacturing process of both plant is similar embracing pre-assembly and assembly cells.

The LLDD was used to assess the value streams of each plant by following the process from supplier to delivery.

5.2 Plant A

Company A started a systematic conversion from mass to lean production in the beginning of the 1990s, when the topic production system became part of its global strategy.

Plant A was started up in 1992 as a greenfield designed under the lean principles.

The supply chain structure of Plant A is organized according other topic of the global strategy named supplier integration.

The program is deployed to the studied plant and objectifies to offer opportunities to the suppliers with best in class performance in quality, cost, development and delivery. The program suggests a long-term partnership whose advantages to the supplier include:

- Systematic consulted on new projects;
- Priority at equivalent performance on in the assignment of new projects and re-sourcing;
- Knowledge sharing of efficient methods.

Regarding the fundamental of quality, the efforts are concentrated on problem solving activities at suppliers and a carefully designed logistics system to maintain the target specification of the product with a high level of standardized containers and transportation equipments.

The fundamentals of on time delivery and lead time are the strongest in Plant A logistics system. It operates a combination of horizontal storage and internal milk-run system and the internal and external pull leveled system by fixed batch kanban in short replenishment cycles, increasing the inventory turn rate.

The main efforts to reduce cost are derived from the strategy of supplier integration that sets a target for material of 15% per year.

No special practice was identified in the payment term. From the customer side, the conditions are defined by the power of bargain of the automakers. On the other side of the value stream, no special attention was given to the trade payables.

At last, the level of intangible assents was reduced due to the simple controls to manage the system and the support of the visual management.

At last, the reduction of the fixed assets is conquest with high standardization of the logistics items as containers, wagons, storage system and wagons.

5.3 Plant B

Plant B was established in the 1950s to supply engine cooling systems to the early Brazilian automotive industry. In the 1980s the Company B made a joint-venture and was completely incorporated the plant in 1994.

The HVAC production unit started in 1997 in a new building at the same site.

The supply basis is selected by the purchasing department according the lowest price that establish contractual and arm's length relationship, with a lack of support to suppliers' development.

Constant shortfalls in quality and delivery along the supply chain are treated without a systematic method to identify and solve root cause.

Inspection processes at the suppliers are implemented in order to avoid transfer defects to Plant B and final test at Plant B protects the customer. Incoming inspection and selection are applied at Plant B to counteract problems.

The material flow is characterized by high level of work in process at the supplier and at the inbound of Plant B and at the assembly line. The storage system uses vertical solution and forklifts to supply the line.

The line supply is done in lots without a standardized procedure, causing frequent production shutdowns.

The production schedule is a push system performed by a MRP (material replenish system) using the forecast of the customer and releasing production orders to the shop-floor.

The supplier's delivery frequency is irregular and the cost are carried by the suppliers.

No special practices are identified to payment term or assets.

5.4 Summary of Results

The Table 1 compares the results of each plant according the logistics fundamentals.

Table 1: Assessment Score

LLDD Fundamental	Score	
	Plant A	Plant B
Supply Chain Structure	4,0	2,5
Quality	3,5	2,5
On time delivery	3,0	2,0
Lead time	4,2	2,2
Cost & Expenses	3,3	3,2
Inventory	3,0	3,0
Payment Term	3,0	3,0
Assets	3,5	2,0
Final Score	**3,6**	**2,5**

The Plant A presents more aspects related with the lean model and therefore shows better score than Plant B, who is a typical mass production site. The most relevant differences are in the fundamental of supply chain structure and lead time, what can be explained by the deployment of a clear corporate strategy to address those topics.

6 CONCLUSIONS

This paper has presented a framework for lean logistics system design and evaluation, proposing a set of functional requirements and respective design parameters. The model is applicable to repetitive manufacturing in a variety of environments. It fulfills the need for a complete method to designing logistics systems under the principles of lean thinking.

The LLDD allows the designers to understand the connection among the high level system objectives and the means of achievement. By understand the interrelations among the different elements of the system design it is possible to assure the fulfillment of the high level objectives.

Two plants are evaluated using the LLDD and it is possible to conclude that the Plant A, that better achieve the objectives of the framework effectively performs better than the Plant B. The interrelation of the logistics methods and practices in the Plant A shows better fulfillment of functional requirements while the Plant B illustrates the problems that occur when tools are applied without a holistic view.

7 ACKNOWLEDGMENTS

I would like to thank Dr. Matthias Schnetzler (ETH Zürich, Switzerland) and M.Sc. Mattheus Pinotti (Université Laval, Canada) for all the support during the development of the framework summarized here.

8 REFERENCES

[1] Parker, G.G., 1998, Contracting for Employee and Supplier Capability Development, Doctoral Thesis,

Sloan School of Management, Massachusetts Institute of Technology, Cambridge, MA.

[2] Fine, C.H., 1998, Clockspeed: Winning Industry Control in the Age of Temporary Advantage, Perseus Publishing, Cambridge, MA.

[3] Laseter, T., Oliver, K., 2003, When Will Supply Chain Management Grow Up?, Strategy + Business, issue 32.

[4] Heckmann, P., Shorten, D., Engel, H., 2003, Capturing the Value of Supply Chain Management, Strategy + Business enews, June.

[5] Croom, S., Romano, P., Giannakis, M., 2000, Supply Chain Management: An Analytical Framework for Critical Literature Review, European Journal of Purchasing and Supply Management, vol. 6, pp. 67-83.

[6] Tan, C.K., 2001, A Framework of Supply Chain Management Literature Review, European Journal of Purchasing and Supply Management, vol. 7, pp. 39-48.

[7] Suh, N.P., 1990, Principles of Design, Oxford University Press, New York, NY.

[8] Suh, N.P., 2001, Axiomatic Design: Advances and Application, Oxford University Press, New York, NY.

[9] Baudin, M., 2004, Lean Logistics: The Nuts and Bolds of Delivering Material and Goods, Productivity Press, New York, NY.

[10] Goldsby, T, Martichenko, R, 2005. Lean Six Sigma Logistics: Strategic Development to Operational Success, J.Ross Publishing, Boca Raton, FL.

[11] Stewart, T.A., Raman, A.P, 2007, Lessons from Toyota's Long Drive: a Conversation with Katsuaki Watabene, Harvard Business Review, July-August.

[12] Womack, J.P., Jones, D.T., Roos, D., 1990, The Machine that Changed the World, Rawson Associates, New York, NY.

[13] Lamming, R., 1993, Beyond Partnership: Strategies for Innovation and Lean Supply. Prentice Hall International, Hertfordshire, UK.

[14] Hines, P., 1994, Creating World Class Suppliers: Unlocking Mutual Competitive Advantage, Pitman Publishing, London, UK.

[15] Nishiguchi, T., 1994, Strategic Industrial Sourcing: The Japanese Advantage, Oxford University Press, New York, NY.

[16] Dyer, J.H., 2000, Collaborative Advantage: Winning Through Extended Enterprise Supplier Networks, Oxford University Press, New York, NY.

[17] Jones, D.T., Womack, J.P., 2002, Seeing the Whole: Mapping the Extended Value Stream, Lean Enterprise Institute, Brookline, MA.

[18] Cochran D.S., Arinez, J.F., Duda, J.W., Linck, J., 2001, Journal of Manufacturing Systems, Vol. 20, n.6, pp. 371-389.

[19] Bocanegra, C., 2001, Design and Implementation of the Product Development Design Decomposition, Massachusetts Institute of Technology, Cambridge, MA.

[20] Schnetzler, M.J., Sennheiser, A., Schönsleben, P, 2007, International Journal of Production Economics, vol. 105, n.1., pp. 21-42.

[21] Supply Chain Council, 2005, Supply Chain Reference Model SCOR Version 7, Supply Chain Council Pittsburgh, PA.

[22] Young, S.D., O'Byrne, S.F., 2001, EVA and Value Based Management: A Practical Guide to Implementation, McGraw-Hill, New York, NY.

[23] Favaro, C., 2008, Aplicação do Projeto Axiomático para Desenvolvimento da Metodologia de Projeto do Sistema Logístico Enxuto, State University of Campinas, Campinas, Brazil. (forthcoming)

Measuring and analysing Levels of Automation in an assembly system

Åsa Fasth[1] Johan Stahre[1] and Kerstin Dencker[2, 3]
[1]Department of Product and Production Development, Division of Production system,
Chalmers University of Technology, Sweden
[2] Royal Institute of Technology, Department of Production engineering
[3] Swerea IVF, Industrial Research & Development Corp.

Abstract
The level of automation employed in semi-automated assembly systems is crucial, both to system performance and cost. This paper presents a methodology to enable selection of the right Level of Automation. The method thoroughly maps existing product and information flows as well as the automation level in separate parts of the system. It then analyses and identifies future automation possibilities, i.e. the automation potential seen from an industrial perspective. Further development of the method is based on validations and industrial case studies.

Keywords:
Levels of Automation, Assembly system, DYNAMO

1 INTRODUCTION

Products of today are becoming increasingly customized. Smaller batches and decreasing time limits for set-ups of new products are some of the resulting demands on the assembly systems, due to the increasing number of variants in the assembly flow.
Consequently, assembly systems have to get the right things, to the right place, at the right time, in the right quantity to achieve perfect work flow while minimizing waste and being flexible and able to change [1]. To achieve this, the companies can adopt automated solutions, when doing this there is a need to determine the correct amount of automation. It is also necessary to identify the optimal parts of the value-flow to be automated. In automation decisions it is necessary to consider human resources, as well as mechanical technology and information flow. By definition, automation is a technology by which a process or procedure is accomplished without human assistance [2]. Unfortunately, automation does not always fulfil expectations; the need for human intervention in cases of disturbances and system failures is still high. Smart automation is defined by [3] as the human aspect of 'autonomation' whereby automation is achieved with a human touch. However, there is a tendency among industry to consider automation investments as a" black or white" decision. This may be suboptimal, since there is not always a need to distinctly choose between humans or machines. The interaction and task division between the human and the machine should instead be viewed as a changeable factor which can be called the level of automation [4]. Thus, identifying and implementing the right level of automation in a controlled way could be a way to maintain the effectiveness of a system[5].
In this paper we will discuss a methodology that could be used to measure today's automation level and that analysis the level of automation that is possible in the future. This could help companies to choose the "right" level of automation due to their production, requirements and demands. This gives great benefits when it comes to time and cost savings in the planning and implementation phase.

2 THE DYNAMO METHODOLOGY

A method was developed in the DYNAMO project (2004-2007) [6] in association with five companies.
The aim with this project was to develop a methodology for measuring and get an accurate picture of today's information flow and automation level in production systems.
Furthermore to develop a reference scale for different Levels of Automation (LoA) that could be used in the manufacturing area [7], this is shown in table 1.

Table I Levels of Automation

Levels	Mechanical	Information
1	Totally manual	Totally manual
2	Static Hand tool	Decision giving
3	Flexible hand tool	Teaching
4	Automatic hand tool	Questioning
5	Static work station	Supervising
6	Flexible workstation	Interventional
7	Totally automatic	Totally automatic

The concept Levels of Automation was defined as;

"The allocation of physical and cognitive tasks between humans and technology, described as a continuum ranging from totally manual to totally automatic"

By physical support, [7], mean the level of automation for mechanical activities, *mechanical LoA* while the level of cognitive activities is called *information LoA.*
Due to [8] the conclusion is that most tasks in manufacturing often involves a mix of both mechanised and computerised tasks and the companies has to consider both areas when automating their system.

2.1 Validation of the methodology

The DYNAMO methodology to measure the Levels of automation consists of eight steps [9], seen in *figure 2.* These steps were validated at an industrial company which has not been participating in the development of the methodology [9]. The validation group contained of four people; two who developed the method and two that began looking at the methodology in 2006 as part of the ProAct project [10]. The group validated each step, ***except step 8***. This step was not validated because the company where the validation took place did not have any purpose regarding LoA strategy, [9].

3 Method for analysing the Levels of Automation

The methodology for analysing levels of automation is done as one step in the **ProAct [11]** research project. This is done to be able to generate possible improvements for the present assembly system. The skeleton of the methodology [7] is intact.
The further development is based on the validation that was done in the DYNAMO project and six case studies that have been done mainly within the **ProAct** project, throughout the period May-Nov, 2007 [12, 13]. The case studies were performed in SME companies.
The modification of the method contains of four different subgroups or phases, as visualised in figure 1.

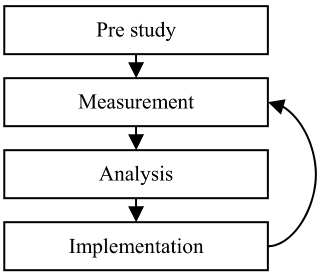


Figure 1 Phases in the measurement methodology

These phases each contains of three steps, shown in figure 2.

Results from the modification are;

- Video or tape documentation has been used in all case studies to easier analyse the assembly system.
- "Lean awareness" in the measured companies is seen as necessary (4 of 6 companies had this) to be able to perform and understand the usage of step 3-10 (modification steps).
- A simplification of the equation has been developed in order to decrease the time in step 4-9 (modification steps), and to decrease the subjective assessment in the work shop.
- No information were given out before the people who were doing the measurement were in place at the companies, i.e. off-site measurement could not been done in any of the case studies.
- Doing a Value Stream Mapping (VSM) to gather information about the time parameters, the information- and material flow in the system to get a deeper analysis of today's system.
- Measure LoA based on value adding and not value adding tasks in the system.
- Logic has been developed to simplify future simulation of levels of automation in an assembly system.
- Consideration also has to be taken about the operators' competence and education about future changes.

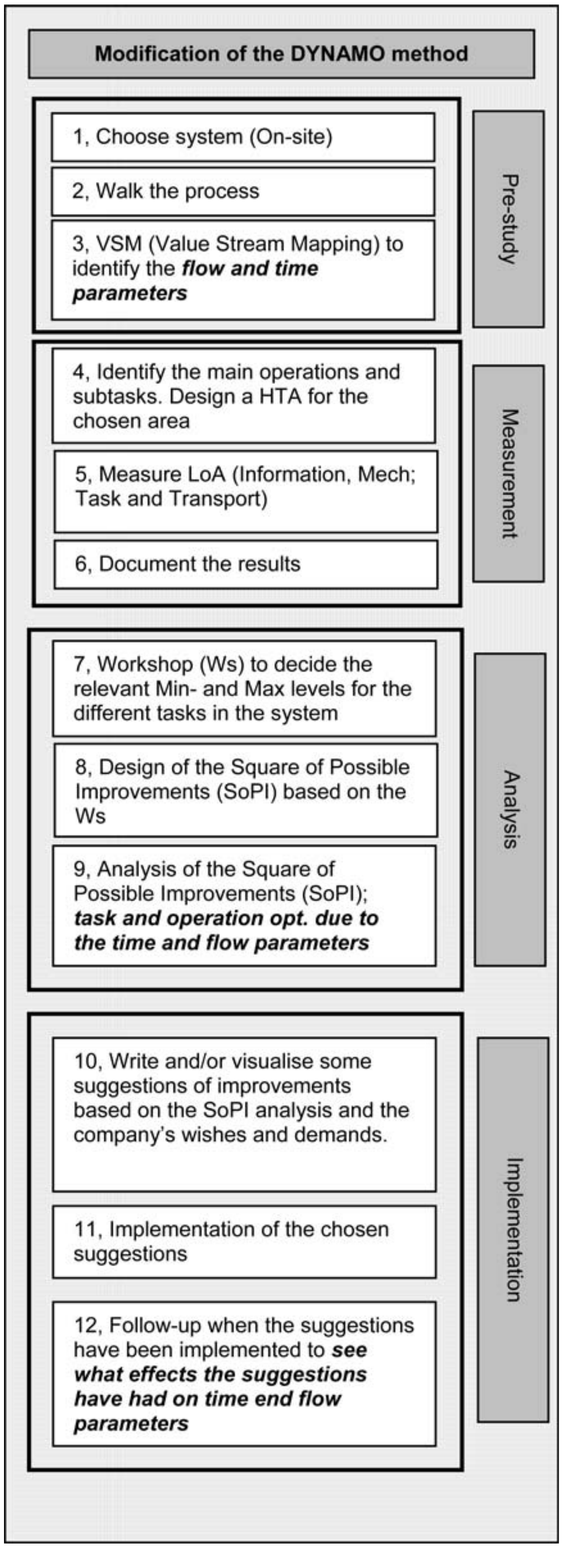


Figure 2 Method for measuring and analysing Levels of Automation

The most important development is the implementation of measurable parameters, time parameters, within the methodology. The logic ha also been very important to develop considering future simulation and visualising of the result in the analysis phase, see subsequent sections.
The analysis and implementation phase is completely new and the sections below will describe these phases.

3.1 THE MEASUREMENT PHASE

The reference scale, seen in table 1 has been developed into a matrix. This is done to get a logical ground and to be able to add dimensions or parameters to the methodology. This matrix is used to visualise the different levels of automation. It is also used in the analysis phase to show the results of the measurements and the suggestions of possible improvements.
The matrix is called LoA_{total} and contains of the vectors LoA_{mech} and LoA_{info} .
The logic for the matrix is seen in equation 1;

$$1 \leq LoA_{total} \leq 49$$
$$LoA_{total} \rightarrow (LoA_{mech}) \wedge (LoA_{info})$$
$$\text{WHERE } LoA_{mech}\ (y) = 1 \leq y \leq 7 \text{ and } LoA_{info}\ (x) = 1 \leq x \leq 7$$

Equation 1 the matrix of LoA_{total}

This means that there are 49 possible solutions that could exist or be developed in the assembly system. It also means that a measured task has to contain both a mechanical and an information part otherwise the structure of the hierarchical task analysis (HTA) is too deep.

3.2 THE ANALYSIS PHASE; Step 8 and 9

These steps are done after the work shop to analyse today's assembly system and to map the possible improvements in the LoA_{total} matrix. This is done with help of the relevant min and max value. These values form a span where the company could move within when it comes to a development of the companies' assembly system; this span forms a square called *Square of possible improvements (SoPI).* The SoPI sets boundaries for the company's future improvements in automation solutions seen from their demands.
This is done to make it easier to analyse the effects when changing/ varying the LoA and also to see if it is possible to make task and operation optimisations within the measured system.
Two different SoPI: s could be designed; task optimisation and operation optimisation. The first step is to design SoPIs´ for all the tasks in the operations.
The logic for the SoPI and $SoPI_{task}$ is described in equation 2:

$$SoPI \rightarrow (LoA_{mech}\ (min; max)) \wedge (LoA_{info}\ (min; max))$$

$$SoPI = LoA_{mech}\ (min; max) * LoA_{info}\ (min; max)$$
$$\text{WHERE}$$
$$LoA_{mech}\ (x) = 1 \leq min \leq max \leq 7 \wedge LoA_{info}\ (y) = 1 \leq min < max \leq 7$$

$$SoPI_{task} \leq LoA_{total}$$
$$SoPI_{task} \subseteq LoA_{total}$$

Equation 2 The Square of Possible Improvements (SoPI) and task optimisatio*n*

To be able to perform an operation optimisation there is one condition, all the $SoPI_{task}$ has to be represented in the $SoPI_{operation}$ in order to make an optimisation, if not, one solution is to do an optimisation with some of the tasks and do a task optimisation on the others. It could be described as;

$$\text{IFF } SoPI_{operation} \subseteq \sum_{task=1}^{n} SoPI_{task}$$
$$\text{THEN operation optimisation is possible}$$

Equation 3 Operation optimisation

The next step in the analysis phase is to evaluate what the effects are when choosing different solutions in the $SoPI_{task}$ or $SoPI_{operation}$ depending on the goal with the measurement.

4 Discussion

From the six case studies, we can declaim that the SoPI is being limited by two obvious things;

- The persons who are leading the work shop - If the leaders of the work shop do not have the skill or knowledge to ask open questions to widen the companies view about the future automation possibilities. It is also important that these persons have a deep knowledge about the methodology and some experience from the industry.

- The companies thoughts and knowledge about future automation - If the companies refrain from “thinking outside the box”

This could result in a smaller SoPi which limit both task and operation optimisations. This has resulted in further development of the assessment in the work shop. This is done to decrease the subjective assessment.
Furthermore to be able to simulate and visualise possible improvements based on the companies' demands and needs rather than their thoughts.
Future research efforts wiil be to investigate how the different solutions in the matrix are related to each other in terms of different time parameters [14]. Furthermore an investigation will be carried out to see if and how it is possible to change level of automation in real time and over time in an assembly system. Deeper knowledge will be used to improve modelling and simulation tools for different levels of automation; this is a part of a research project called SIMTER. Future research aims at simulating and visualising assembly systems with varying LoA in the system's stations and tasks. Improvement of flexibility and the flow-and time parameters will be measured; this will be done to be able to develop a proactive assembly system in the project called ProAct.

5 SUMMARY

This paper has presented a methodology for measurement of the level of automation in assembly systems. While based on the DYNAMO methodology, a development of the measurement- and analysis step has been shown to provide a good visualisation of the present assembly system.
The resulting automation matrix provides 49 different solutions that can be compared and analysed. The Square of Possible Improvements (SoPI) shows the span within the matrix where company personal believe their systems could be improved. The improvement potential is seen from different perspectives described by parameters, resources and demands.

The development of extended method logic and the addition of the time dimension to the existing LoA reference scales will provide opportunities to easier simulate different solutions for assembly systems.
Furthermore, it will provide measurable values that could be analysed in today's assembly system in view future systems. This will give the companies a solid base for decision making in planning and implementation phases of developing their future assembly system.

6 ACKNOWLEDGEMENTS

The authors would like to express their deep gratitude to researchers and industries that participated in the three projects DYNAMO, SIMTER, and ProAct. This research was financially supported by the Swedish Foundation for Strategic Research (SSF) and also by the Swedish Governmental Agency for Innovation Systems (VINNOVA).

7 REFERENCES

[1] J. K. Liker, *The Toyota Way: 14 Management Principles from the World's Greatest Manufacturer*. USA: McGraw-Hill, 2004.

[2] M. P. Groover, *Automation, production systems, and computer-integrated manufacturing*, 2. ed. ed. Upper Saddle River, N.J.: Prentice Hall, 2001.

[3] T. Ohno, *Toyota Production System*: Productivity Press, 1988.

[4] R. Parasuraman, T. B. Sheridan, and C. D. Wickens, "A model for types and levels of human interaction with automation," *IEEE transactions on system, man, and cybernetics - Part A: Systems and humans,* vol. 30, pp. 286-296, 2000.

[5] M. Bellgran and K. Säfsten, *Produktionsutveckling - Utveckling och drift av produktionssystem*. Lund, Sweden: Studentlitteratur, 2005.

[6] J. Frohm, V. Lindström, and M. Bellgran, "A model for parallel levels of automation within manufacturing," in *the 18th International Conference on Production Research*, Fisciano, Italy, 2005.

[7] J. Frohm, "Levels of Automation in production systems," in *Department of production system*. vol. Doctorial Gothenburg: Chalmers University of technology, 2008.

[8] J. Frohm, V. Lindström, M. Winroth, and J. Stahre, "Levels of Automation in Manufacturing," *Submitted for publication in Ergonomica*.

[9] V. Granell, J. Frohm, J. Bruch, and K. Dencker, "Validation of the DYNAMO methodology for measuring and assessing Levels of Automation," in *Swedish Production Symposium* Gothenburg, 2007.

[10] K. Dencker, J. Stahre, P. Gröndahl, L. Mårtensson, T. Lundholm, J. Bruch, and C. Johansson, "PROACTIVE ASSEMBLY SYSTEMS- Relaizing the potential of human collaboration with automation," in *IFAC-CEA, cost effective automation in networked product development and manufacturing* Monterey, Mexico, 2007.

[11] K. Dencker, J. Stahre, P. Gröndahl, L. Mårtensson, T. Lundholm, and C. Johansson, "An Approach to Proactive Assembly Systems - Towards competitive assembly systems," in *IEEE international symposium on assembly and manufacturing (ISAM), University of Michigan, Ann Arbor, USA.ISAM*, 2007.

[12] K. Jonsson, "Flow analysis and simulation of new and old production plant at Eka Chemicals AB," in *Dep. of product and production development*. vol. Master Gothenburg: Chalmers University of Technology, 2007, p. 43.

[13] P. Ottosson and G. Grundström, "Flow analysis and simulation of the cooling module production at NIBE AB," in *Product and production development*. vol. Master Gothenburg: Chalmers University of Technology 2008 to be submitted.

[14] Å. Fasth, J. Stahre, and K. Dencker, "Analysing changeability and time parameters due to levels of Automation in an assembly system," in *FAIM* Skövde, 2008 to be submitted.

Sustainable Design of Machine Tools through Load-Dependent Interventions and Adapted Services

Jürgen Fleischer[1], Matthias Schopp[1], Alexander Broos[1], Jan Wieser[1]
[1] Institute of Production Science (wbk), Universität Karlsruhe (TH), Germany

Abstract
Manufacturers of machine tools face an increasing demand for total cost of ownership (TCO) contracts including an availability guarantee for their products. One key figure in machine availability is Overall Equipment Efficiency (OEE), which is affected by the wear of machine components from both regular use and singular environmental conditions. This paper presents an approach to derive strategies for load-dependent interventions based on the diagnosis and prognosis of the component state, leading to a high level performance with a reliable forecast of machine and component failure in specific working environments.

Keywords:
Machine Tool, Sustainable Design, Lifecycle Engineering

1 INTRODUCTION

1.1 Starting point

One of the main challenges machine and component manufacturers are often required to cope with is lack of information from different areas. After expiration of the warranty period, field experience from operating performance and failure patterns is usually not passed on from the operators of machines and components to their manufacturers [1].

While environmental conditions are difficult to estimate, the regular wear from machining operations can be estimated as soon as the load spectrum during machining operations is known. With this information, it is both possible to derive load-dependent interventions to keep loading within desirable limits, or to derive adapted services to guarantee a high operational availability. Furthermore, feedback to the machine design process facilitates the design of machine tools with high reliability, which starts with dimensioning and selecting appropriate machine components, such as guides and bearings [2].

1.2 Objective

The investigations described in this article are performed in order to allow for the development of an integrated methodology to achieve availability increases for machine tools through continuous load and equipment condition monitoring and to optimize machine operation in terms of its life cycle. Monitoring is supposed to provide the basis for the generation of countermeasures for timely maintenance and diagnosis, and it is to allow for active intervention in the process in order to maintain or increase availability. The main focus is on capitalizing on the machine technology employed in the best possible way to achieve an availability increase. This approach allows for the estimation of potentials and risks required to determine an adapted service scheme for availability at constant costs. The implementation of this approach to machine availability increases requires a deeper understanding of the actual loads bearing on the machine tool as a whole and on its individual components, which will enable the manufacturer to match machine technology design and configuration to the requirements set out by the user in terms of load spectra. This – if combined with overload restriction – allows for machine operation to be always kept just below the allowed limit load.

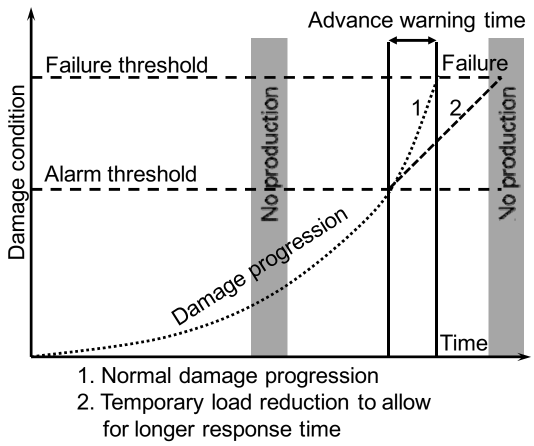


Figure 1: Intervention scenario temporary load reduction (“emergency operation program”).

Figure 1 shows one of the potential scenarios for the implementation of the intended strategy. On the basis of a known damage progression curve and an alarm threshold, the machine dynamics is temporarily reduced via a load controller to enable the operator to replace the component at a suitable time.

2 APPROACH

2.1 Machine tool modularization approach

At the beginning of the project, a comprehensive analysis of machine tools and their operating behaviour under real production conditions is to be carried out. For the different configurations of all machines to be compared with each other in terms of the availability of their individual components, the machine tools were modularized. The machines are broken down into the component assemblies and their components down to the level of individual functions in three steps.

2.2 Data collection

On the basis of the above-described modularization methodology, the machinery of several users is looked into, and the data found is then set alongside the experience of the manufacturers involved in the project and the supporting committee of industry representatives. For representativeness to be ensured, the service data of machine manufacturers and maintenance and failure records of operators are evaluated in addition to the collection and evaluation of own data in form of a field data collection. If the selected modularization method is followed consistently, data for different machine tool configurations can also be collected and used for analysis.

2.3 Data evaluation

In a first step, the failure data is ranked at the levels of component assembly and components. The collected data is entered into a standard software program. Additional analyses on the difference between operator and manufacturer data are carried out to reflect the different perspectives represented within the project group. Besides, the machines are itemized according to their age and configuration structure, and noticeable characteristics during operation are recorded.

The analysis covering more than 250 machines shows the following distribution of TOP4 failure areas (100 %):

1. Axes (38 %).
2. Spindle/Tool changer (26 %).
3. Electrical/Electronic system (23 %).
4. Fluidic system (13 %).

The distribution of all problem causes among ten assembly groups illustrates that these areas hold the greatest potential in terms of an increase of technical availability, which puts the focus of all subsequent investigations on the above-listed component assemblies and the individual failure causes and mechanisms. In general, the perceptions of machine tool manufacturers and machine tool operators match. The method described below is mainly based on axis components. Among others, these include the components ball screws, bearings and guides, which account for approximately 1/3 of all failures in this case and which will be discussed in more detail below.

3 MEASUREMENT-BASED ANALYSIS OF THE OPERATING CONDITIONS OF SELECTED COMPONENTS

3.1 Diagnosis

A thorough understanding of the operational behaviour of the machinery used is indispensable for an early detection of a potential failure [2]. To that end, extensive test runs are carried out on test rigs particularly set up for that purpose and on the component while incorporated in the selected machine (MAG Powertrain Star500). In the tests, the test machine represents a standard machining center of medium size (working envelope 620/630/900mm).

The tests are performed to collect data on the following variables:

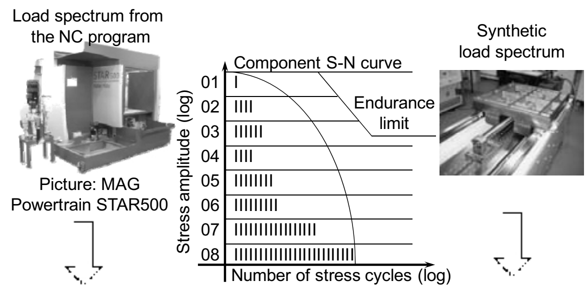
1. Defined differentiators between good and poor conditions;
2. Reaction time during operation (i.e. the time between the definite prediction of a failure and its actual occurrence.)

All selected components of the selected machine are measured on separate test rigs and in the built-in condition. In case of the spindle, not only the balanced but also unbalanced conditions are deliberately measured.

The same method is used for axes components: Good and poor conditions are compared both on the test rig and when the component is actually built into the machine.

3.2 Measurement-based analysis of loads and their impact on damage progression

Not only are undamaged and defective components to be distinguished between, but the effect different loads have on the parts also needs to be looked into. In this respect, the impact of parameters related to the machine dynamics are of particular interest with regard to their effect on damage progression.



Analysis of load-dependent operating conditions of the power train

Figure 2: Load analysis by the example of ball screws.

Figure 2 shows the basic parallel method including measurements both in the machine and on the test rig by the example of the axes assembly. For each test set-up, both intact and defective components are used. Damage progression is recorded on the basis of the dynamics of the predefined loads, and extensions of the reaction time caused by dynamic process intervention are evaluated.

Awareness of the loads resulting from the respective work piece spectrum and all boundary conditions is important for both settings. As part of these investigations, two exemplary work pieces are defined and analyzed by the operators for the purpose of defining dynamic variables for all test series which reflect real conditions.

4 THE PROGNOSIS OF OPERATING LOADS THROUGH THE USE OF SIMULATION METHODS

The operating loads bearing on the machine components (ball screw, bearings, guides) during machining are largely unknown and are only of minor importance for the design calculation. The load spectra used instead are empirical ones

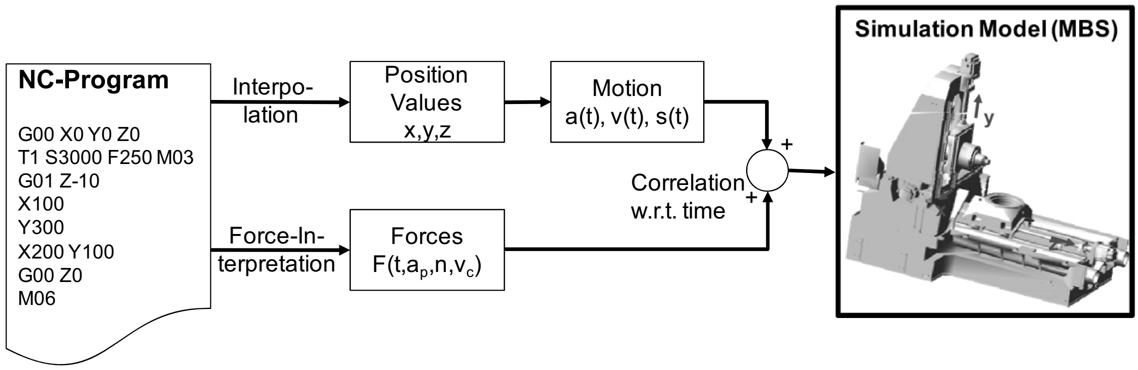


Figure 3: Schematic illustration of model generation.

corresponding to an estimate of the maximum loads to be expected. Knowledge of the actual loads, however, is required to assess the load history and, thus, the expected remaining service life. These loads can be directly derived from the NC program running on the machine. The NC program is generated on the basis of cutting data and machining strategy, whose impact on components can only be analyzed by means of a complex process if determined through measurement. Measurements of a machine during industrial operation are very difficult to perform, as additional sensors would be required.

For that reason it makes sense to resort to the use of simulation methods to collect operating data. Multi body simulation is particularly suitable to connect a virtual machine to the NC program (Figure 3). To that end, a machine model is set up via multi body simulation focusing on the detailed display of the power trains. Besides, the model of a machine control is generated which ensures the controlling of machine movements in co-simulation with the multi body model. An NC interpolator generating the reference values for the control is linked to it. In addition, the NC program is analyzed for the cutting conditions linked to the respective travel command, and a process load model is then generated on the basis of these results.

A model set up as described above which allows for the analysis of effective component loads. Figure 4 shows an exemplary simplified work piece for which the simulation was carried out. The defined machining operation was the trimming of a triangle with 300 and 400mm edge length, whereas the machining forces were extremely simplified for initial investigations and defined as the normal force against the cutting direction. The static basic load is 6 kN. It is superposed by a dynamic load with a 2 kN amplitude at a frequency of 300 Hz, allowing for the simulation of component loads.

The resulting loads can then be used as a basis to determine the nominal service life. According to the design guidelines for anti-friction bearings, anti-friction guides or ball screws, the mean operating load can be calculated from the simulation results according to the formula

$$\bar{F} = \sqrt[3]{\sum_{1}^{i} F_i^3 \cdot \tfrac{t_i}{100}} \qquad (1),$$

the average being determined via the time percentage [3]. The result can be used to determine the service life L10, which is calculated dependent on the dynamic load rating c_α

$$L_{10} = \left(\frac{c_a}{\bar{F}}\right)^3 \cdot 10^6 \qquad (2).$$

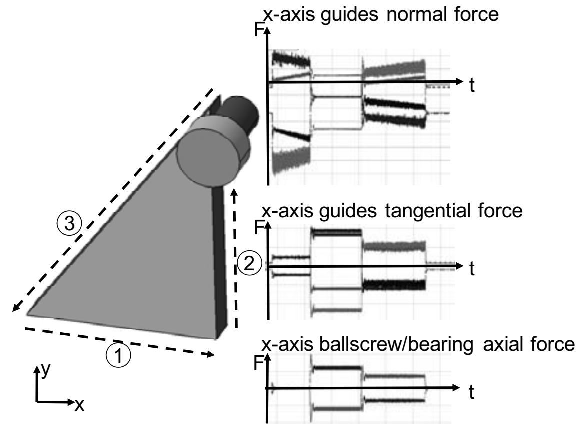


Figure 4: Component loads for an exemplary machining operation.

Depending on the type of component, the service life is expressed as the number of rotations or the travel distance.

With this method, which is actually used for component selection in industry, the calculated service life and the service life recalculated on the basis of machining loads can be compared. Under the assumption that only the sample part is machined during the entire operating time, the result for the sample machine could be as follows:

Table 1: Exemplary discrepancy between the service life of power train components according to design vs. actually used.

Component (X axis)	Service life	
	By design	Used
Guide	100%	75 %
Ball screw	100%	85 %

The nominal service life would be expected to rise/fall by 33/25 % but this may vary distinctly due to environmental conditions.

The next step would consist in analyzing the effect of intervention in the machining process on component service life in the form of either adaptation of the NC control program or changes to the machine parameters (control setting, jerk limitation).

The use of simulation offers several benefits. Firstly, the design calculation for certain machining conditions can be optimized; secondly, countermeasures for the reduction of operating loads can be evaluated a priori in order to conceive strategies for the best possible operation of the machine in terms of its life cycle. Also, it can be an option to derive strategies for a load-dependent disposal of work pieces to machine tools. The use of simulation, thus, allows for an optimization of the design calculation for determined machining conditions [2].

5 DEVELOPMENT OF ADAPTED SERVICES

Both awareness of the machine failure pattern and the possibility to configure a machine for predefined loads during the design phase in the best possible way in terms of its life cycle help machine manufacturers to provide adapted technical services [5, 6]. During future investigations, support services will be devised on the basis of the technical implementation described above and will then be evaluated by machine tool operators. The data on wear behavior gained from this project and the prognoses of the remaining service life of individual components which has now become possible can be used to set up new offers or to improve or complete existing maintenance services. One of the investigations will look into the amount of spare parts required as safety stock and the quantity of consignment stock to be kept at the machine operator's. If the failure prognosis reaches a sufficient level of accuracy, it may suffice to have the spare parts delivered just in time for a planned maintenance operation.

A profitability analysis must be carried out for the emergency operation program shown in Figure 1 which requires a longer machining time and for the load-dependent disposition of work pieces within a production system. The implementation of these strategies represents an effective response to disturbances in interlinked systems, particularly in the light of declining safety buffers in production and decreasing output quantities.

All scenarios illustrated in this article will have to prove economical in profitability analyses evaluating availability increase and costs, a requirement which also applies to extended warranties potentially given by machine manufacturers [1, 6].

If there is a direct intervention in the production process, the conditions for the operation and the type of intervention will need to be determined. The manufacturer guarantees the user availabilities for certain load classes provided that the machine dynamics can be adjusted according to load.

6 SUMMARY

The total cost of production machinery throughout its entire lifecycle is increasingly taken into consideration, which provides machine manufacturers with a key differentiator but also requires existing machine concepts to be optimized. The investigations presented in this article illustrate how on the basis of a failure analysis at the components level specific methods can be developed to detect disturbances of the machinery at an early time, for example through the use of measurement systems. In addition, simulation can be used in order to determine and optimize the critical components of individual component assemblies already during the design phase.

7 ACKNOWLEDGMENTS

The OPTILAST project is funded by the German Federal Ministry of Education and Research (BMBF) as part of the program “New machine tools for tomorrow's production” and is supervised by the Project Management Agency Forschungszentrum Karlsruhe (PTKA), field office Dresden.

Project partners: MAG Powertrain (Ex-Cell-O GmbH), BoschRexroth AG, Heidelberger Druckmaschinen AG, HYDAC INTERNATIONAL GmbH, KSB AG, Siemens AG, Institute for Control Engineering of Machine Tools and Manufacturing Units (ISW/University of Stuttgart), Institute of Production Science (wbk)/Universität Karlsruhe (TH).

8 REFERENCES

[1] Wawerla, M., 2008, Risikomanagement von Garantieleistungen – Methodische Identifikation, Beurteilung, Steuerung und Überwachung der Risiken von Garantieleistungen im Maschinen- und Anlagenbau, Bd. 140, Forschungsberichte aus dem wbk Institut für Produktionstechnik, Universität Karlsruhe (TH).

[2] Fleischer, J., Schopp, M., 2007, Sustainable Machine Tool Reliability based on Condition Diagnosis and Prognosis, Takata, S., Umeda, Y. (Editors): Advances in Life Cycle Engineering for Sustainable Manufacturing Businesses. Proceedings of the 14th CIRP Conference on Life Cycle Engineering, Waseda University, Tokyo, Japan, Springer-Verlag London:323-328.

[3] Bosch Rexroth AG (Edt.), 2007, Handbuch Lineartechnik. Bosch-Rexroth Firmenpublikation Nr. R310DE 2017.

[4] Fleischer, J., Munzinger, C., Broos, A., Schopp, M., Wieser, J., 2007, Einsatz und Potential der Mehrkörpersimulation für die Ermittlung von Komponentenbelastungen bei Werkzeugmaschinen, Tagungsband Dresdner Maschinenelemente Kolloquium, Verlag der Wissenschaften GmbH, Dresden:425-432.

[5] Schuh, G., Klotzbach, C., Gaus, F., 2007, Designing services based on ‘intelligent’ press-die-systems, Takata, S., Umeda, Y. (Editors): Advances in Life Cycle Engineering for Sustainable Manufacturing Businesses. Proceedings of the 14th CIRP Conference on Life Cycle Engineering, Waseda University, Tokyo, Japan, Springer-Verlag, London:201-205.

[6] Meier, H. (Edt.), 2004, Dienstleistungsorientierte Geschäftmodelle im Maschinen- und Anlagenbau: Vom Basisangebot bis zum Betreibermodell, Springer-Verlag, Berlin.

Modeling and implementation of Digital Semantic Machining Models for 5-axis machining application

Fumiki Tanaka [1], Masahiko Onosato [1], Takeshi Kishinami [2] , Kiyoshi Akama[1],
Makoto Yamada [3], Tsukasa Kondo [3] and Satoshi Mistui [4]
[1] Graduate School of Information Science and Technology, Hokkaido University, Japan
[2] Kushiro National College of Technology,, Japan
[3] Department of Mechanical Engineering, Hakodate National College of Technology, Japan
[4] Department of Department of Information System Engineering, Asahikawa National College of Technology, Japan

Abstract
To realize the articulate link between design and manufacturing, it is necessary to realize interoperable digital tools which represent the semantics of information independently of any implementations. It is necessary to assure that the quality of the shape information overcomes the limitation of interoperability of the shape information. In manufacturing, the interoperable CNC machining systems can be characterized as being capable of 1) seamless information flow; 2) feature based machining; etc. In this research, modeling and implementation of the Digital Semantic Model for 5-axis machining that fulfils the above requirements is proposed.

Keywords:
Semantic modeling; Manufacturing System; 5-axis machining

1 INTRODUCTION

The movement toward concurrent engineering drives the need for frequent transformation of the product data between different digital tools. To realize the articulate link between design and manufacturing, it is necessary to realize interoperable digital tools which represent the semantics of information independently of any implementations. However, some fundamental problems which cause manufacturing inefficiency have been pointed out and are as follows [1]:

1. Interoperability among heterogeneous software systems and tools is generally not available. The design, manufacturing, and supply chain software systems are not interoperable across software ownership boundaries.
2. The use of sub-optimal process parameters increases the manufacturing cost because of increased tool wear or under-utilized tools and machines.

Today, all engineering data, such as geometry representation data, tolerance data, etc., are not used throughout product development processes, as shown in Figure 1.

An interoperability error makes product model data unusable when applying digital tools. If an interoperability error arises in collaborative engineering applications, tracking the error back to its source is expensive (if not impossible) and requires a significant amount of time. The National Institute of Standards and Technology (NIST) report titled 'Interoperability Cost Analysis of the US Automotive Supply Chain' estimates that the economic cost of bad interoperability in the US automotive industry is one billion dollars per year [2]. Thus, it is necessary to assure that the quality of the shape information overcomes the limitation of interoperability of the shape information [3].

Manufacturing companies face increasingly frequent and unpredictable customer demands, and interoperable CNC machining systems have become key systems to meeting these customer demands. This type of system can be characterized as being capable of 1) seamless information flow; 2) feature based machining; 3) autonomous CNC etc. [4]. However, current CNC systems and NC data do not have these capabilities.

In order to realize the articulate link between design and manufacturing and the traceable manufacturing system, it is necessary to represent the semantics of information independently of any implementations. Moreover, the quality of data that represent information should be ensured during the product life cycle. In this research, the Digital Semantic Model that fulfils the above requirements is proposed. In this paper, modeling and implementation of the Digital Semantic Model for 5-axis machining is reported.

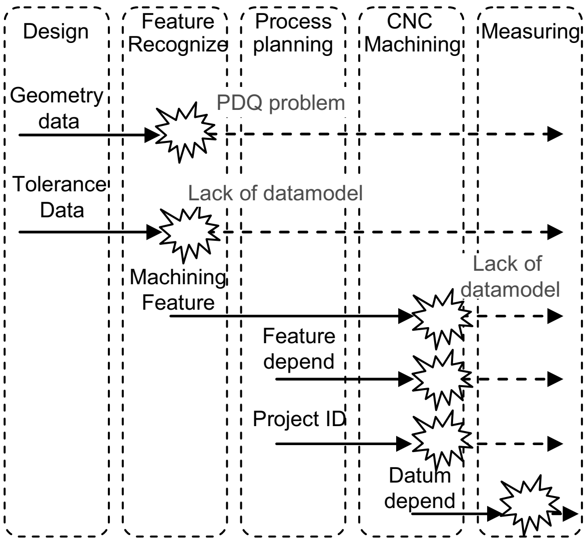


Figure 1: The problem of the communication in product development processes.

2 DIGITAL SEMANTIC PRODUCT MODEL

2.1 Basic concepts of the digital semantic product model

The digital semantic product model consists of the quality assurance of the product model data, the semantic modeling for humans (such as designer and worker), understandable and computer readable data, and the functional modeling for supporting the product life cycle. In order to ensure the product model data, the Product Data Quality Checking method for the ISO Standard product model (ISO 10303-214 [5]) is adopted. In order to support all of the product lifecycle activities, the functional model for each activity is developed based on the related ISO Standard, such as ISO 10303-224 [6], 238 [7], 240 [8], and ISO 14649 [9-11]. In order to represent the semantics of the data, the EXPRESS language [12] for data modeling and XML for representing the data [13], are adopted. The EXPRESS language was proposed and standardized in ISO TC184/SC4. The elements of the digital semantic product model such as PDM data, geometric data, application attribute (machining feature), and application management (project ID) shown in Figure 2, are modeled based on the ISO Standards, In this research, three research themes shown in Figure 2 were chosen to solve typical problems in current product lifecycle.

2.2 Modeling and implementation of digital semantic product models for 5-axis machining application

Figure 3 shows the modeling and implementation of digital semantic models for 5-axis machining. The manufacturing procedures of the product using proposed digital semantic models are described below. At first, the designed product model (AP-214 data) is checked in terms of its data quality based on the product data quality assurance method. Secondly, the machining features for 5-axis machining are recognized and ISO 14649 CNC data are generated. The machining feature for 5-axis machining is also proposed as an element of the digital semantic product model. Finally, the product is manufactured using the machine tools for ISO 14649 CNC data model.

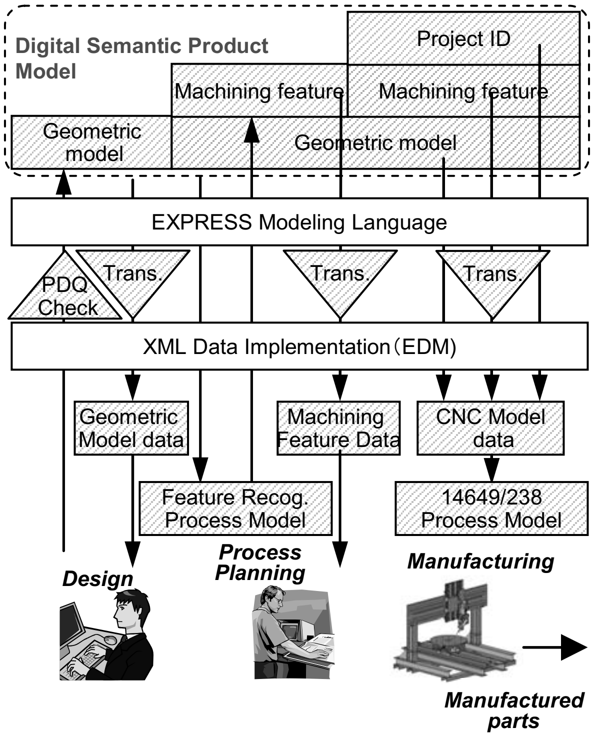


Figure 3: Modeling and implementation of Digital Semantic Models for 5-axis machining application.

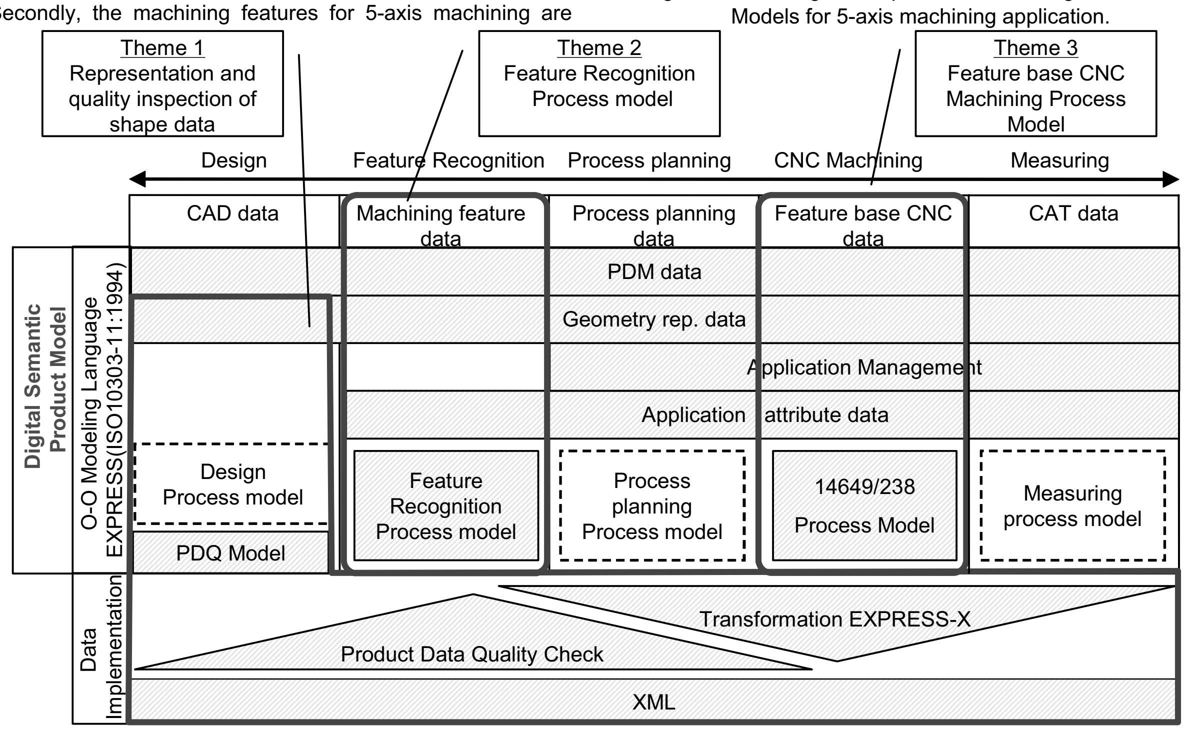


Figure 2: The elements of the digital semantic product model and research theme.

3 PRODUCT DATA QUALITY ASSURANCE METHOD

3.1 Current problems of checking the quality of product data

At present, interoperability errors are one of the critical problems in concurrent engineering. In order to avoid an interoperability error, it is necessary to develop a quality checking mechanism for the shape information of the product model data and to verify the shape data whenever they are transformed. Some commercial software systems have been developed for implementing PDQ guidelines. However, their algorithms are not open, and their accuracy cannot be verified. Even if their algorithms are presented in natural languages, there is no assurance for equivalence between algorithms and implemented software.

Moreover, the PDQ criteria defined for one industry need to be extended easily to other industries. In order to specify the algorithms for the user, the data format of the product data also needs to be open and described in an unambiguous way.

In brief, current problems for PDQ diagnostics are as follows: (Figure 4)

1. In order to avoid interoperability errors in collaborative engineering environments it is necessary to verify the shape data using various PDQ-tools before the shape data are used in different CAx systems. Therefore, many PDQ-tools (or criteria) for specific combinations of CAx systems are needed.
2. PDQ criteria such as SASIG-PDQ guidelines are described in natural language, and a certain amount of ambiguity exists in the measurement methods of the product data quality. Therefore, it is difficult to implement a PQQ-tool.
3. A PDQ-tool uses closed (secret) algorithms in which the consistency between tools cannot be ensured. Even if a PDQ tool presents its algorithms in natural language, there is no assurance for equivalence between algorithms and implemented software, and the PDQ criteria need to be extended easily.

3.2 Proposed method of checking the quality of product data

To solve the problems mentioned above, a software platform for checking the quality of shape data of the STEP product model for CAD/CAM environment is proposed as shown in Figure 5 in this paper. The constituents of the proposed method are as follows:

1. The target shape data of the product model should be a neutral file which is independent of any specific CAx systems. Moreover, the data structure of the product model data should be known by all users in order to specify the PDQ criteria of the shape data. Then, the target product model data uses STEP Application Protocol (AP) 214 [5] for the purposes of the automotive industry, and it is widely supported by many CAD systems. The data structure of AP 214 data is explicitly described in formal language to avoid any ambiguity of the PDQ criteria.
2. PDQ criteria should be described in formal language. Describing the criteria in formal language means representing the diagnostics algorithms in some kind of programming language. In this research, the PDQ criteria are described in the Equivalent Transform (ET) language [14] which is a kind of rule based language. The described quality inspection algorithms are referred to as quality criterion gauges.
3. The description language (or formal language) should have the ability of extending the existing criteria easily, and anyone should be able to use it. As mentioned above, the description language of PDQ criteria is the Equivalent Transform (ET) language in this research. The advantages of using ET include ease of developing the checking programs as the constraint rules. The software platform for checking the quality of shape data of the STEP product model for CAD/CAM environment is also developed.

Using the proposed developing environment of the PDQ inspection algorithm, systematization of PDQ-criteria based on the classification of PDQ criteria from the viewpoint of their geometric algorithms [14] is possible.

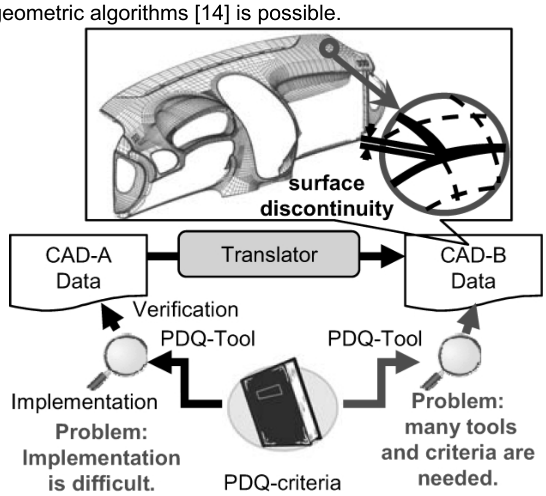


Figure 4: Current problems of checking the quality of product data.

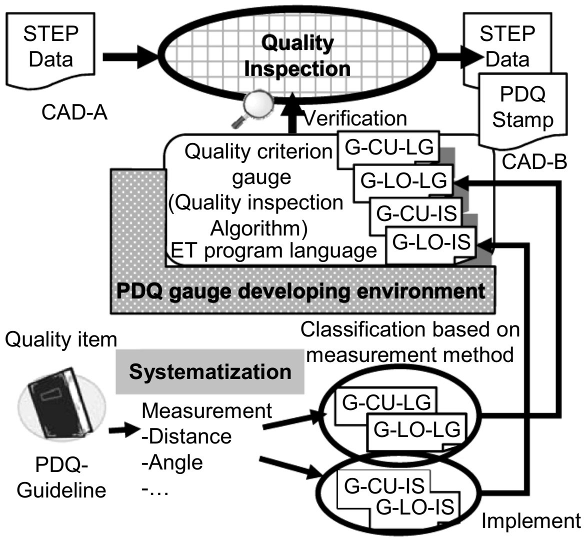


Figure 5: Modeling and implementation of Digital Semantic Models for 5-axis machining application.

4 MACHINING FEATURE FOR 3+2 AXIS MACHINING

4.1 Concepts of 3+2 axis high speed machining

Finishing processes for injection molding molds and die casting dies are done by EDM because these applications require the machining of many thin and deep cavities. To solve this problem, mold and die machining using indexing tilted angles of the tool axis on a 5-axis machining center was proposed [15]. Typical machining operations with a milling tool are classified into 3-axis machining and 5-axis machining as shown in Figure 6. As mentioned above, we propose mold and die machining which is carried out on an indexing tilted tool axis using a 5-axis machining center shown in this figure. In the proposed machining process, the orientation of the cutting tool remains fixed during high speed machining. This kind of machining is referred to as 3+2-axis high speed machining in this paper. Thus far, there has been a significant amount of research on machining in which the orientation of the cutting tool has remained fixed. In this paper, 3-axis machining is performed as "high speed machining" of hard materials.

4.2 Machining feature for 3+2 axis high speed machining

In this machining process, the machinable area for each indexing tilted angle is defined as a machining feature for 3+2 axis machining as shown in Figure 7.

To calculate the machinable area (feature) for each candidate tool indexing angle, the offset surface of the required surface should be obtained, and machining simulation should be done on the obtained offset surfaces. To do so simultaneously, the inverse offset method [16] is adopted in this research. An offset surface can be generated from the enveloped surface of the cutting edge, which is generated by moving the reverse cutting tool on the required shape.

The state flag is added to the inverse offset point to distinguish between the offset surface derived from a cutting edge part and the offset surface derived from a tool holder part. To obtain machinable facets, a reference flag to the required machining facet from the offset point derived by inverse offset operation is also added. The deriving method consists of the following procedures:

1. A grid space area in the model space is set in the computer, and Z-Queue is prepared to memorize the Z coordinate value to each grid point. On that occasion, we set the maximum Z coordinate value as the intersection point of facets and Z-Queue, as shown in Figure 8.
2. An inverse offset operation is applied to the previously obtained Z-Queue for the tool model with a tool shank part and a holder as shown in Figure 9. By this operation, the offset surface which can avoid holder interference, and the required machining surface can be obtained.

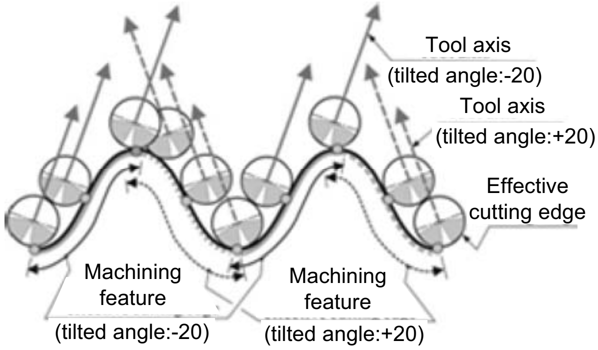


Figure 7: Machining features for 3+2 axis machining

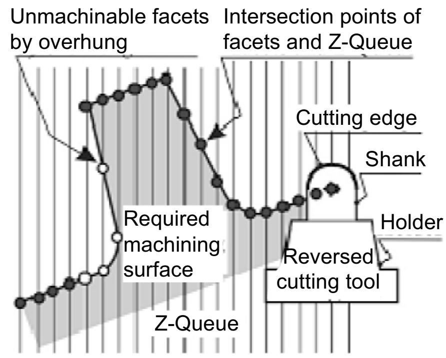


Figure 8: Generation of machining points on Z-Queue, and model of cutting tool

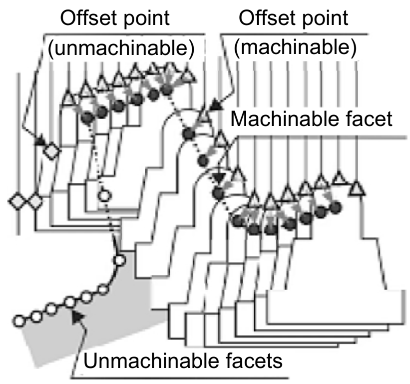


Figure 9: Deriving of machinable facets

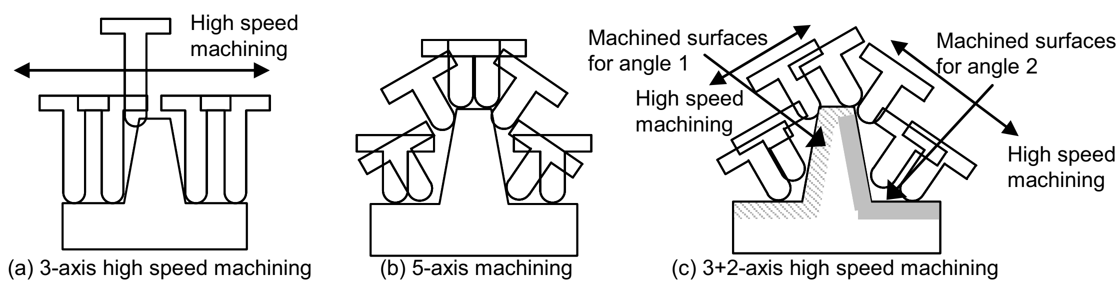


Figure 6: 3-axis high speed machining, 5-axis machining, 3+2-axis high speed machining.

5 MACHINE TOOLS FOR ISO 14649 CNC DATA MODEL

In this paper, a prototype system of 5-axis machine tools for ISO 14649 CNC data model is developed [17].

The fundamental principle of the ISO 14649 NC Data Model is the object-oriented view of programming in terms of manufacturing features [9], instead of direct coding of sequences of axis motions and tool functions defined in ISO 6983. Together with the general process data described in ISO 14649-10 [10], ISO 14649-11:2003 [11] specifies the technology-specific data elements needed as process data for milling.

The overview of the data structure of ISO 14649 is shown in Figure 10. Workingsteps represent the essential building blocks of the machining process. Each workingstep describes a single manufacturing operation for a machining feature using one tool and one strategy. The tool movement within the machining operation is determined by the technology-dependent strategy and additional parameters. In order to determine the tool movement to reach the first cut, optional information about the approach (plunge) strategy can be specified. Moreover, in order to determine the tool movement after finishing the last cut, optional information about the retract strategy can be specified.

The software structure of the 5-axis machine tool is shown in Figure 11. ISO 14649 XML data are read to this machine tool, and machining motion data are generated using machine tool kinematic model. Finally, the product is machined from generated machining motion data.

The machine tool model consists of a kinematic model which describes a kinematic structure of a machine tool and a shape model which describes all the shapes of a machine tool. In order to describe a kinematic structure of a machine tool, particular entities are added to an information model standardized in ISO10303-105 [18]. A data structure of the machine tool model is shown in Figure 12. The 'mechanism' is described as a kinematic structure with a single link fixed to the ground using entity 'kinematic_structure' in ISO10303. The 'kinematic_structure' is represented by means of 'kinematic_joint' which describes a relationship between a pair of links. Each link is represented by a 'kinematic_link'. The 'kinematic_pair' defines the kinematic constraints between two adjacent links coinciding at a joint. ISO10303-105 covers 17 types of 'kinematic_pairs.' Although in this study, types of 'kinematic_pairs' for a machine tool are the only 'revolute_pair' which constrains the motion between two adjacent links to a rotation about a common axis, and the 'prismatic_pair' which constrains the motion between two adjacent links to a translation along a common axis. The 'su_parameters' describe the position between adjacent pairs and consists of 6 parameters [18]. Generally, the kinematic mechanism of a machine tool is considered to be two mechanisms from a base of a machine tool to a tool and from the base of a machine tool to a workpiece. Therefore, two subtypes of entity 'mechanisms' are a 'base_to_tool_ mechanism' and a 'base_to_workpiece_mechanism.' The 'base_to_tool_mechanism' is specified as a mechanism from the base to a tool and the 'base_to_workpiece_mechanism' is specified as a mechanism from the base to a workpiece.

The prototype machine tool is shown in Figure 13. As shown in Figure 13, the machine tool consists of the X, Y, Z axis, the A-axis which is a rotation axis of the main spindle and the C-axis which is a rotation axis of the work table.

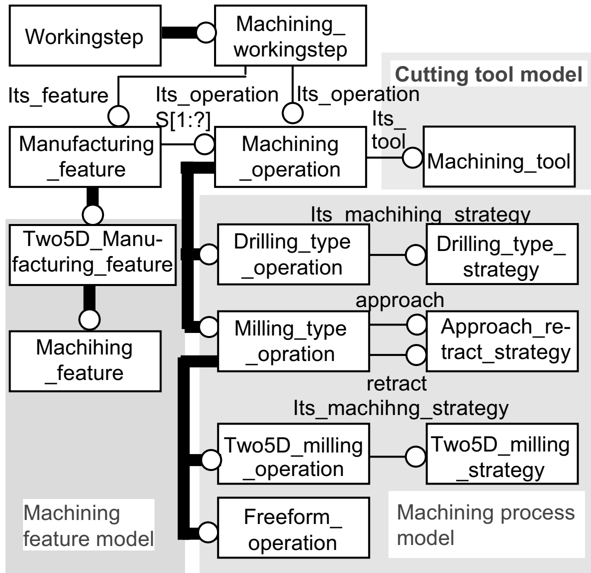


Figure 10: Overview of the data structure of ISO 14649

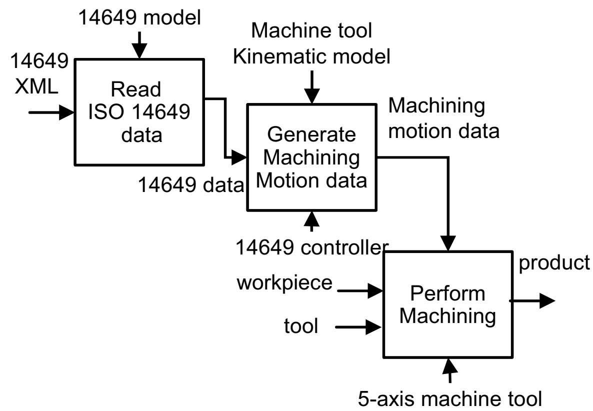


Figure 11: Software flow of machine tools

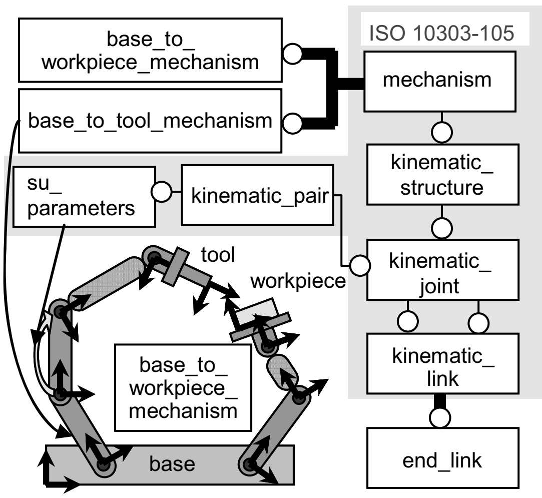


Figure 12: Kinematic model of machine tool.

Figure 13: Machine tool for ISO 14649 CNC data model.

6 CONCLUSIONS

In this paper, modeling and implementation of Digital Semantic Model for 5-axis machining is reported. Conclusions of this paper are as follows:

1. The problems which are faced in current manufacturing systems are pointed out. To solve these problems, the basic concepts of the digital semantic model are proposed. The overview of the Digital Semantic Models for 5-axis machining is described
2. Current problems of checking the quality of product data are pointed out and product data quality assurance method is proposed. The target shape data is the ISO standard 10303-214 and PDQ criteria are described in the Equivalent Transform (ET) language.
3. Machining feature for 3+2 axis high speed machining is proposed and the deriving method from product data is also proposed based on inverse offset method.
4. Overview of the prototype 5-axis machine tool is described. The machine tool model based on ISO10303-105 is also proposed.

REFERENCES

[1] Feng, SC., Stouffer, KA. and Jurrens, KK., 2005, Manufacturing planning and predictive process model integration using software agents, Advanced Engineering Informatics, Vol. 19, No. 2, 135-142.

[2] National Institute of Standards and Technology, 1996, Interoperability Cost Analysis of the US Automotive Supply Chain, Planning Report #99-1, NIST Strategic Planning and Economic Assessment Office, Gaithersburg, Md., http://www.nist.gov/ director/prog-ofc/report99-1.pdf.

[3] Tanaka, F., Kishinami, T., 2006, “STEP-based quality diagnosis of shape data of product models for collaborative e-engineering,” Computers in Industry, vol. 57 no.3, 245-260.

[4] Xu, X. W., Wang, L., Rong, Y., 2006, STEP-NC and Function Blocks for Interoperable Manufacturing, IEEE Trans. Automation science and engineering, vol. 3 no. 3, 297-308.

[5] ISO 10303-214, 2003, Product data representation and exchange Part 214: Application Protocols: core data for automotive mechanical design processes.

[6] ISO 10303-224, 2006, Product data representation and exchange Part 224: Application Protocols: mechanical product definition for process plans using machining features.

[7] ISO 10303-238, 2007, Product data representation and exchange Part 238: Application Protocols: Application Interpreted Model for Computerized Numerical Controllers.

[8] ISO 10303-240, 2005, Product data representation and exchange Part 240: Application Protocols: process plans for machined products.

[9] ISO14649-1: 2003, Data Model for Computerized Numerical Controllers: Part 1 Overview and Fundamental Principles.

[10] ISO14649-10: 2004, Data Model for Computerized Numerical Controllers: Part 10 General Process Data.

[11] ISO14649-11: 2004, Data Model for Computerized Numerical Controllers: Part 11 Process Data for Milling.

[12] ISO 10303-11, 2004, Product data representation and exchange Part 11: Description methods: The EXPRESS language reference manual.

[13] ISO 10303-28, 2007, Product data representation and exchange Part 28: Implementation methods: XML representations of EXPRESS schemas and data, using XML schemas.

[14] Tanaka, F., Iwata, T., Akama, K., Kishinami, T., Onosato, M., 2006, Developing software platform for Checking Quality of shape data of STEP product model for CAD/CAM environment, Proc. of FAIM2006, 823-830.

[15] Tanaka, F., Yamada, M., Kondo, T., Kishinami, T., Kohmura, A., 2004, Software system for sculptured surface machining based on 3+2-axis high speed machining on a 5-axis machining center, Proc. of JUSFA symposium 2004, JL-006.

[16] Kondo, T., Kishinami, T., Saito, K., 1988, Machining System based on Inverse Offset Method, Journal of the Japan Society for Precision Engineering, vol. 54 no. 5, 167-172.

[17] Mitsui, S., Tanaka, F., and Kishinami, T., 2006, Machining Feature-Driven 5-Axis CNC Machine Tools, JSPE Publication Series No.5, Towards Synthesis of Micro-/Nano-systems, 11th International Conference on Precision Engineering, 169-174

[18] ISO 10303-105, 1996, Product data representation and exchange Part 105: Integrated application resource: Kinematics.

Compensation of Thermal Deformations at Machine Tools using Adaptronic CRP-Structures

Eckart Uhlmann[1] and Philipp Marcks[1]

[1] Institut für Werkzeugmaschinen und Fabrikbetrieb, Technische Universität Berlin, Germany

Abstract

Thermal strains represent immediate quality affecting factors for the characterisation of machine tools and production processes. Today, strategies for the compensation of axial and radial displacements are widely used whereas a compensation of angular displacements is not yet realised satisfyingly.

This paper presents an approach to reduce thermally induced angular displacements of main spindle housings using carbon fibre reinforced plastic (CRP) structures.

The compensation of thermally induced angular displacements utilises the negative thermal expansion of CRP-structures in an adaptronic system. This system is a combination of thermal sensors, point sensors, controllers and CRP-actors which include controlled heating of the laminate by heat films and heating filaments.

Keywords:

Compensation; Deformation; Temperature

1 ADAPTRONIC COMPENSATION OF THERMAL DISPLACEMENT OF MACHINE TOOL SPINDLES USING CRP-BANDAGES

Thermal strains of machine-tool-spindles represent immediate quality affecting factors for the characterisation of machine tools and production processes. Today, the compensation of axial and radial displacements can be controlled, whereas a compensation of angular displacement is not yet realised satisfyingly [1].

Carbon fibre reinforced plastics (CRP) provide remarkable potentials for the reduction of thermal displacements due to their low or negative thermal expansion coefficients. Therefore, the contraction of a unidirectional laminate through heating can be used to compensate thermal expansion of the surrounding components [1]. Given suitable choices of the structural geometry regarding thermal symmetry and ideal heating of the unidirectional carbon fibre reinforced laminate, in this paper solutions for an active compensation are described.

In the field of adaptronic compensation of thermally conditioned strains within machine tools or machine tool components, so far no scientific deliberations exist. One reason for this is that predominantly piezo-electric material is used for the adaptronic compensation of dynamic strains. Since the amplitudes of thermal strains are considerably higher than dynamic strains piezo-ceramic materials are not suitable for this application.

However, for scientific purposes research on adaptronic compensation of thermal strains is very important. One approach to solve this problem is to use the negative thermal expansion of a CRP-bandage through controlled heating. The adaptronic system, therefore, is a combination of thermal sensors, point sensors, controllers and actors which provide the heating of the laminate as heat wires and heat films. Depending on temperature- and point-data at the spindle housing, the controlled heating of the CRP-bandage leads to a contraction which in turn compensates the thermal expansion due to thermal dissipation loss.

During this investigation a prototype of a spindle housing was designed using both finite elements analysis (FEA) and analytical calculation. One important aspect within this stage is the design of the CRP-bandage with regard to its ability to compensate thermal strains arising in the spindle housing. The design of the spindle housing is aimed at improvements in the static, dynamic and thermal behaviour. In this stage, different approaches are compared with each other by comparing the results of the FEA and the analytical calculations. The most appropriate design is then built up as a prototype.

Under ideal conditions radial thermal displacement can be reduced by up to 70% using a CRP-bandage compared to the results of a monolithic aluminium body. The distinct reduction of thermal strains at bearing points can considerably reduce the angular displacement at the tool-centre-point. Depending on the bearing clearance this results in an improvement of machining accuracy.

The computed strains refer to a thermally symmetric design with thermal dissipation loss at the bearings according to an existing machine tool spindle. As a result of the lower thermal conductivity of the carbon fibre reinforced plastics, a significantly higher heat power is needed for the heating of the bandage. Additionally, an internal heating of the bandage is necessary to provide a homogenous thermal distribution within the bandage. The internal heating is done by inserting heating filaments into the bandage and therefore reduces the angular displacement at tool-centre-point.

With an adaptronic approach for controlled heating of CRP-bandages, the thermal strains and therefore the radial and angular displacements at the tool-centre-point of machine tool spindles can be reduced. The self controlled system is able to react to changes in displacement and temperature autonomously and represents a new method for improving machine accuracy.

Due to the thermally unsymmetrical design of the spindle housing especially at guideline joints, where a concentration of the mass is located, inhomogenities in thermal heating of the spindle housing arise from lost heat of bearings and the motor-spindle. These inhomogenities within the temperature field prompt a bending of the spindle housing and, therefore, an angular displacement of the spindle. An approach to selectively reduce this angular displacement of the spindle housing is the application of adaptronically controlled CRP-bandages. These bandages are arranged radially around the spindle housing, Figure 1.

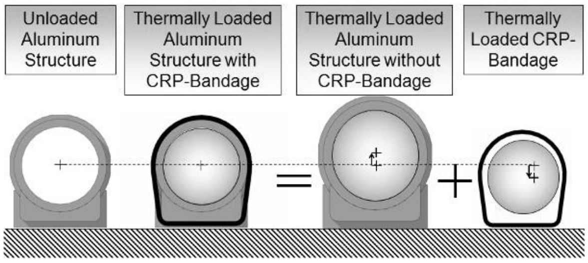


Figure 1: Principle of compensation using CRP-bandages.

By using selective heating of high modulus CRP-bandages, a defined contraction can be generated, which reduces the thermal displacement of the spindle housing. The required heating power can be provided with heat films, heat filaments or peltier elements. Heat films and heat filaments in this case are preferable because of their higher heat output compared to peltier elements. Considering the relatively low thermal conductance of CRP and the radial arrangement of the bandages, a high heat output of the heat films and heat filaments is needed for this specific application.

The objective of the project is the design of compensation structures, consisting of a spindle housing suitable for the application of CRP-bandages. By using an adaptronically controlled heating a directed manipulation of thermally induced angular displacements is done by making use of the contraction of the CRP-bandages. The static, dynamic and thermal behaviour of the newly designed spindle housing is additionally analysed using the Finite Element Analysis (FEA). Additionally, a prototype of the spindle housing is built up and used for experimental analysis.

1.1 Material selection

The selection of suitable materials for the use in structures which compensate thermal strains in machine tools is limited by the mechanical properties of the available material used. To compensate thermal strains of an aluminium spindle housing, thermal displacement, resulting from a positive linear expansion coefficient has to be thwarted. Materials with negative linear expansion coefficients are suitable to meet this task.

Basically, a range of fibre material with negative linear expansion coefficients is available. Here, the anisotropy of these materials has to be kept in mind. The desired material property of negative thermal expansion only exists in the direction of the fibre. For aramid the ratio of the thermal expansion across the fibre in relation to the expansion in the direction of the fibre is 35, whereas the value in direction of the fibre is negative and the value across the fibre is positive. In addition to aramid reinforcement fibres carbon fibres with also negative thermal expansion coefficients exist.

A comparison of the values for Young´s modulus, strength, strain at failure and for the thermal expansion coefficient in the direction of fibre shows that aramid fibre due to its lower strength is less suitable than high modulus carbon fibre. An additional problem arises from the tendency of aramid fibre to absorb water, which leads to a loss of adhesion between fibre and matrix material and a loss of strength. Therefore, the use of high modulus carbon fibre as material for the bandages appears to be the best choice. As matrix material, epoxy is used because of its good material properties in combination with carbon fibres. Here, the excellent dynamic material properties have to be underlined [2].

2 FE-ANALYSIS OF THE NEWLY DESIGNED SPINDLE HOUSING

The dynamic behaviour of the newly designed spindle housing was analysed using the FEA. For better comparability, a reference spindle housing was analysed as well, Figure 2. The reference spindle housing is in use in the HSC-machining centre LPZ 500, MAP-GmbH, Berlin, Germany, which is located at the Produktionstechnisches Zentrum (PTZ), Berlin. The analysis contained the determination of the first ten eigenfrequencies which are important for the dynamic behaviour of the machine tool components. Here, the effects of design changes as well as the effect of an integration of the CRP-bandages were determined. Using FE-analysis, the first eigenfrequency of the reference spindle housing was determined to be 260 Hz at a distance of the guide wagons of 350 mm and fully extended spindle. This value was used as a reference for the evaluations of the newly designed spindle housing, Figure 2.

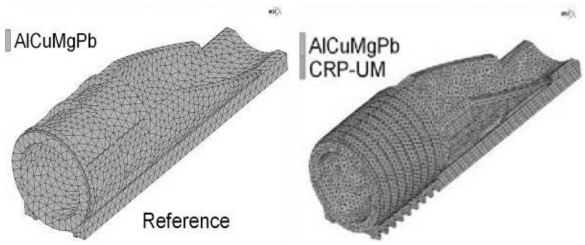


Figure 2: Spindle housings used for dynamic analysis.

With adjustments to the design it was possible to improve the eigenfrequencies substantially. Especially with the use of CRP-bandages it was possible to gain an increasing of the first eigenfrequency from 423 Hz to 663 Hz.

Responsible for this effect is firstly the additional strength of the spindle housing and secondly damping effects of the CRP-bandages.

For a controlled heating of the unidirectional CRP-bandage, an adaptronic sensor-actor-coupling is used that initiates heat depending on the temperature- and displacement data at the spindle housing into the CRP-bandage and therefore leads to a contraction of the bandage. The controller is realized as a model-based virtual instrument (vi) in National Instruments LapVIEW®. For the generation of the relevant output signals, temperature differences and thermal expansion data are detected. By combining this data the active heating of single bandages is controlled. As thermal displacement cannot be reduced without cooling, only thermally induced angular displacements resulting from an inhomogeneous radial displacement can be compensated. The resulting larger linear displacement of the spindle housing, however, can be compensated more easily by a machine tool control [1]. Figure 3 displays the significant reduction of the displacement in the radial direction for a monolithic aluminium tube. The clear reduction of the radial shift at the bearings additionally leads to a reduction of the angular displacement at the tool-centre-point in dependence of the bearing offset. This reduction has direct positive influence on the machining accuracy of the machine tool.

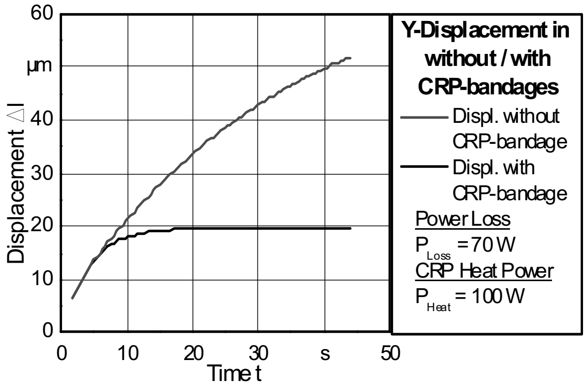


Figure 3: Radial displacement of aluminium-tube without and with use of CRP-bandages.

Based on the low heat conductivity of the CRP-material, a significantly higher heating power is necessary for the heating of the CRP-bandage. In addition, a heating of the bandage in the interior is achieved by integrating heat filaments into the structure. By integrating heat filaments into the CRP-bandages a homogeneous temperature field in the component is generated in order to avoid an angular shift at the spindle tip.

3 EXPERIMENTAL ANALYSIS OF THE DEMONSTRATOR PROTOTYPE

During this project a demonstrator prototype was designed to do experimental analyses of the effectiveness of CRP-bandages for the reduction of thermal strain. The initial experiments that were made with the demonstrator were without CRP-bandages. With later experiments the bandages were then included in the assembly, Figure 4.

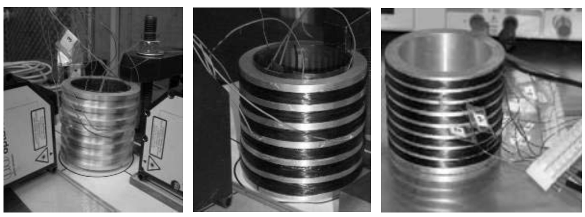


Figure 4: Demonstrators for experimental analysis.

A comparison of the experimental data with the analytical data from the FEA shows a very good correlation both in the setup without CRP-bandages and with CRP-bandages, Figure 5 and Figure 6.

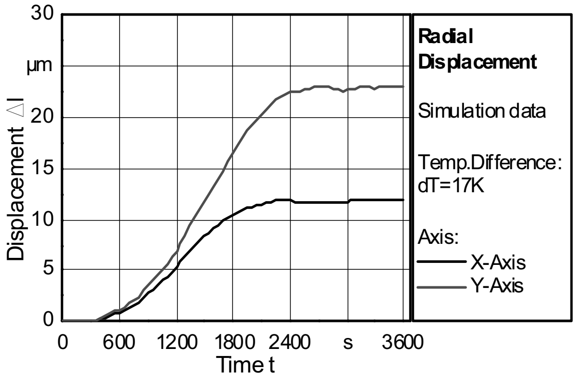


Figure 5: Simulation data of demonstrator without CRP-bandages.

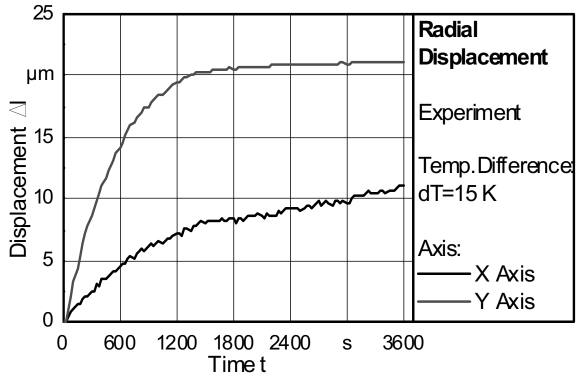


Figure 6: Experimental data of demonstrator without CRP-bandages.

The effect of reducing the thermal displacement by using CRP-bandages is shown by comparing Figure 6 and Figure 7. The experimental data acquired during the tests of the setup including CRP-bandages show a clear reduction of the thermal displacement. Here, it is important to mention, that the heating of the bandages was done passively using the heat generated at the inside of the demonstrator. The compensatory effect can therefore be enhanced by applying an external heating of the CRP-bandages. This is currently worked upon within this project.

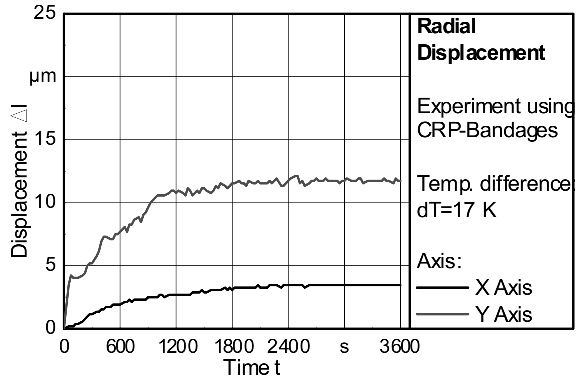


Figure 7: Experimental data of demonstrator using CRP-Bandages.

The experimental data generated in this project clearly shows the potential of CRP-material to reduce a thermally induced angular displacement of aluminium spindle housings. The lightweight material also helps to improve the dynamic behaviour because of its good damping characteristics. The newly designed spindle housing examined in this project already shows the mode of operation and the compensation capabilities which can be achieved with CRP-bandages [3].

A compensation of the thermal displacement through the use of CRP-bandages is possible and offers advantages regarding the dynamic behaviour with a regulated heating. Therefore, it is possible to react to occuring displacements of the spindle housing and to increase the accuracy of the machine tool. The draft for the adaptronic regulation of the developed compensation structures is represented in Figure 8.

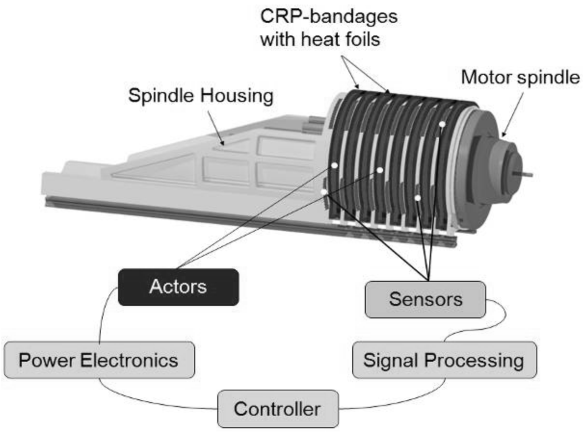


Figure 8: Concept for the fully integrated adaptronic spindle housing.

4 SUMMARY

The knowledge gathered in this project underlines the feasibility of a regulated compensation of thermal displacements in machine tool spindles with the use of CRP-structures. Herewith the use of high modulus carbon fibre is suitable because of the mechanical and thermal properties, which are superior to all competing fibre types for this application. A radial taping of the spindle housing, however, is problematic and requires a new construction of the entire spindle housing. Based on the low thermal expansion coefficients a retroactive assembly of the compensation structure is not possible, the carbon fibre must be directly taped on the spindle housing. A further reason for this is the uniform load transmission into the structure. In the case of the compensation structure integrated into the spindle housing, the use of additional load transmission elements has to be minimized, because of the reduction of the efficiency of the CRP-bandage. This fact was considered in the development of the spindle housing.

In order to realize the required heating power for the CRP-bandages, the use of heat foils located outside in connection with heat filaments brought in into the bandage is necessary. An additional cooling of the spindle housing is not necessary in this case. If applying the compensation with radially arranged CRP-bandages it is important to provide a good load transmission of the bandage into the structure. Due to the contraction of the bandage causing tangential loads, the normal force required for the compensation of the thermal expansion of the spindle housing are naturally small. A further extension of the compensation of thermal displacements with CRP is, therefore, to design suitable load transmission components which can make use of the linear thermal contraction of the CRP-structure directly, Figure 9 [4].

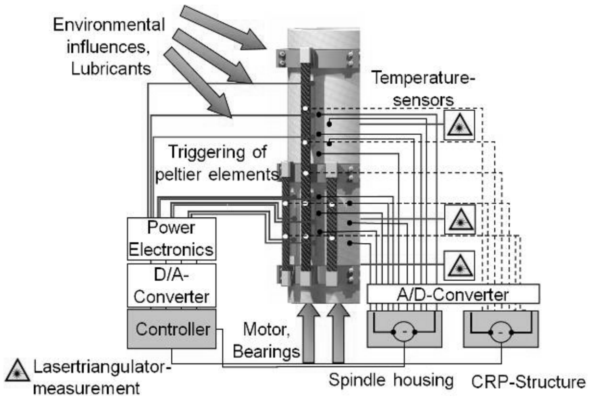


Figure 9: Concept for the fully integrated adaptronic spindle housing with linear CRP-rods.

Furthermore, this setup enables the usage of peltier elements for the simultaneous cooling of the spindle housing and heating of the CRP-structure as shown in Figure 11.

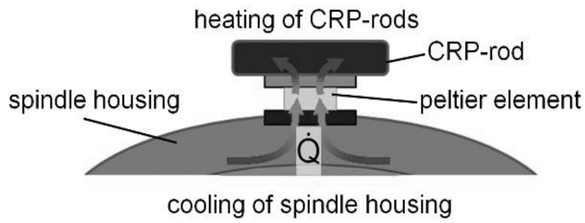


Figure 10: Working principle of simultaneous heating and cooling using peltier elements.

5 ACKNOWLEDGMENTS

The authors would like to thank the German Research Foundation (DFG) for the financial support of the project Uh 100/58-2 within the framework of SPP 1156.

6 REFERENCES

[1] Uhlmann, E., Marcks, P., 2006, Adaptronische und steuerungstechnische Kompensation thermischer Verlagerungen an Werkzeugmaschinenspindeln, Futur, 3.

[2] Ehrenstein, G., 2006, Faserverbund-Kunststoffe. Werkstoffe, Verarbeitung, Eigenschaften, Hanser Fachbuch Verlag.

[3] Uhlmann, E., Marcks, P., 2007, Adaptronic Compensation of Thermal Strain at Machine Tool Spindles using CRP-Bandages, Second Manufacturing Engineering Society International Conference (MESIC).

[4] Uhlmann, E., Marcks, P., Mense, C., 2007, Milling machine evolution in the area of conflict between efficiency, accuracy and social ecology, 12th International Seminar on High Technology UNIMEP.

Image Noise Reduction Using Wide Range Tuned Mass Damper for Scanning Electron Microscope

Mitsuru Hamochi [1], Makoto Aoshima[2] and Shinji Wakui [3]

[1] Electron Optics Division, JEOL ltd., Japan

[2] Metrology Inspection Division, JEOL ltd., Japan

[3] Department of Electrical and Electronic Engineering, Tokyo University of Agriculture and Technology, Japan

Abstract

In a scanning electron microscope, outside acoustic noise causes image noise that distorts observations of the specimen. This paper investigates vibration damping against the eigenmodes of the base plate. To restrict these modes, a tuned mass damper (TMD) using viscoelastic material was designed. The natural frequency and damping coefficient of the TMD were tuned by dividing the width of the materials. It was effective to set the ratio of the width and thickness smaller than certain value, for higher damping coefficient and wider frequency range. The TMD decreased the vibration from 250Hz to 1.3kHz, and reduced image noise remarkably.

Keywords:

scanning electron microscope; acoustic noise; tuned mass damper; viscoelastic; mode shape

1 INTRODUCTION

Scanning electron microscope (SEM) is used in wide range applications from the observation of microorganisms to material analysis. Recently, the resolution of the SEM has reached to 1 million magnifying power. SEM is often used in offices or clean rooms, which have relatively high levels of floor vibration and sound pressure. These cause small vibrations in the SEM. These vibrations flexure the SEM structure, and affects its resolution. It is possible to reduce the translation of the floor vibration with its isolator. However, it is difficult to reduce the force input by the acoustic noise. This noise excites eigenmodes of each structure, and increases the vibration amplitude. The eigenmodes of the specimen stage are most sensitive against the resolution of SEM, however it was difficult to change the vibration modes of the stage. Thus, the structural mode shaping of the base plate had conducted to reduce the transmissibility of the vibration [1], [3]. It was effective to place the specimen chamber on the vibration nodes at 1st and 2nd twisting modes of the base plate, although at the higher modes, this place was no longer the nodes. Thus, reduction of the vibrations at higher modes was desired.

In this paper, a tuned mass damper (TMD) is used to reduce the vibration. We focused on the largest amplitude in each eigenmode of the base plate. Next, we developed a TMD effective in the wide frequency range [2] using “shape effect” of the elastic materials. The TMD using elastic materials had been well known, however it had not been mentioned about the detailed design technique [5], and the effect had been considered relatively small in general [6]. Therefore, we discuss how to design this kind of TMD. To confirm the performance, we installed the TMD to the base plate and remarkable reduction of SEM image noise under acoustic noise excitation was shown.

2 THE SEM AND EXCITATION OF IMAGE NOISE

2.1 SEM

Figure 1 shows SEM principles. An electron beam is generated by an electron gun and concentrated through a condenser lens, and is directed to certain point on the specimen. The intensity of the secondary electron emission is measured by a detector. Two-axis scanning of the electron beam forms the image of the specimen.

2.2 Image Noise Excitation by Transmission of the Acoustic Noise and Vibration

The floor vibration and the acoustic noise vibrate the SEM at small amplitude. This causes the relative displacement between the electron beam and specimen, and fluctuates the SEM image. This fluctuation in the SEM images that we refer to as image noise is observable when it is greater than the resolution of the SEM. The countermeasure of the image noise with the electron beam deflection was also studied [4], however, this paper is concerned about the mechanical

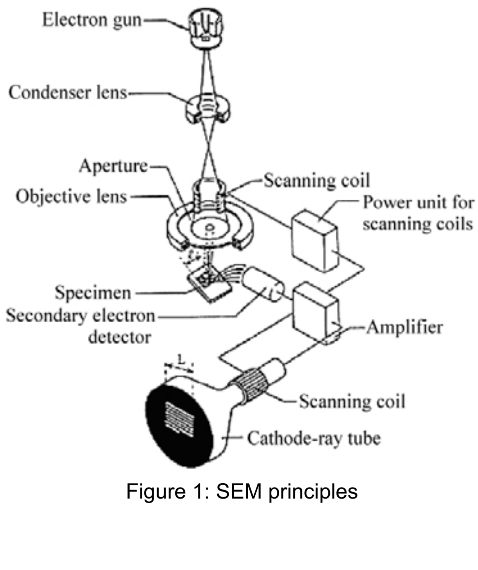


Figure 1: SEM principles

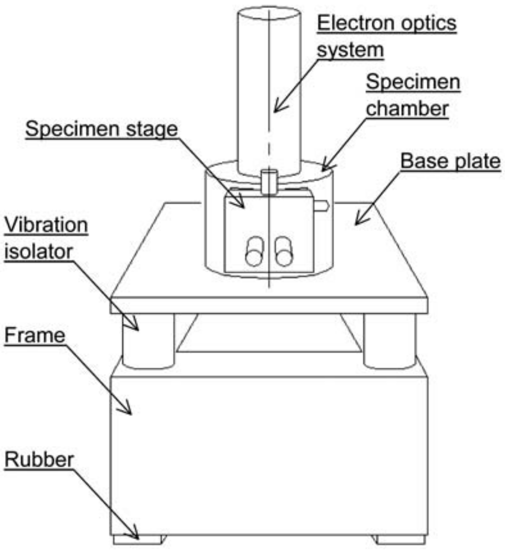


Figure 2: Main structure of SEM

damping of the structure. Figure 2 shows the main structure of SEM. The electron optics system that generates directs the electron beam is located on the specimen chamber. The specimen stage that holds and moves the specimen is also mounted on the specimen chamber, which is fixed to the base plate. These structures ride on the vibration isolators that reduce the transmission of the floor vibration. The isolators are placed on a frame that is set on the floor through rubber feet. Figure 3 shows the factor analysis of the transmission path of the vibration excited by floor vibration and acoustic noise [3]. The arrows show the transmission paths of vibration and dashed line show the input paths of acoustic noise. Acoustic noise impinges on all surfaces of the SEM and excites each local eigenmode of the structures. These vibrations transmit to the specimen. In this paper, we focus on the vibration exited by acoustic noise. Thus, the explanation concerned with the floor vibration and the factors not shown in Figure 2 will be omitted. The base plate is one of the sensitive structures on SEM image noise caused by acoustic noise. As mentioned above, the vibration transmissibility was reduced with the mode shaping of the base plate: that avoid the overlap of the frequencies with the

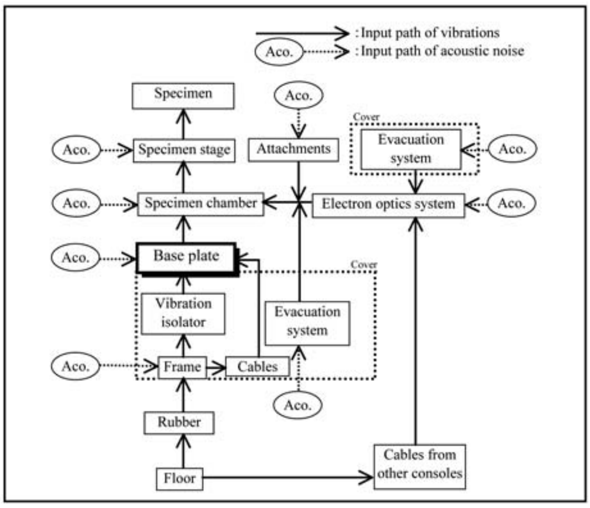


Figure 3: Transmission model of SEM vibration [3]

eigenmodes of the specimen stage, and 1st and 2nd modes are both twisting that has vibration nodes around the specimen chamber [1]. In addition, we attempted to reduce the image noise by the restriction of higher modes. In Chapter 3, it will be discussed in detail about the mode shapes of the base plate and the design of TMD.

3 MODE SHAPES OF THE BASE PLATE AND DESIGN OF TMD

3.1 Mode Shapes of the Base Plate

Figure 4 shows the structure of the base plate. The double octagonal rib is welded below the octagonal plate. With this rib, both 1st and 2nd modes show twisting. Fig. 5 shows the mode shapes of the base plate calculated with FEM analysis. The blue areas have smallest amplitude, and the red areas have largest amplitude at their modes. The main direction of the vibration is perpendicular to the plate. Focusing on the 4-trapezoidal areas at the edge of the plate, they have relatively large amplitudes at the all modes. On the contrary, relatively small amplitudes are shown at the areas inside of the rib frame due to larger stiffness.

3.2 Design of TMD

A multi-layered structure with thin layer of viscoelastic material to add the damping to the plate is well known [7]. However, it needs another binding part that has enough stiffness, and the mass is remarkably increased. A damping structure that needs no another binding part is also known, though its effect is smaller. TMD is often used to reduce the vibration at tuned narrow frequency range, although it is impossible to reduce the several frequency peaks at one time. To restrict the several peaks, multi parallel TMDs [8] are used, which are tuned to each peaks, and they need more space and difficult to be applied. If TMD that works sufficiently at wider frequency rage was developed, it would be possible to restrict several frequency peaks at one time. Figure 6 shows the structure of the developed TMD. Refer to Figure 5, TMDs are placed on 4-edges of the base plate that have relatively large amplitudes at all vibration modes. TMD is composed with a steel mass and viscoelastic materials. Above mentioned, the directions of the vibrations at the 4-edges are mainly perpendicular to the plate. Thus, viscoelastic materials are pressed and expanded vertically, not horizontally. The energy consumption of the viscoelastic material depends on the amplitude of the vibration displacement, and is not concerned with the velocity like at the oil damper [6]. Strictly speaking, the damping coefficient

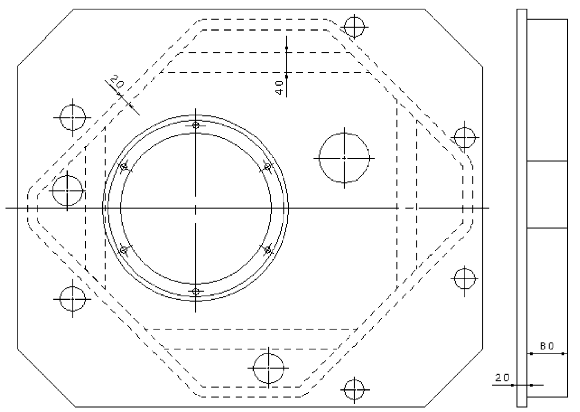


Figure 4: Structure of the base plate

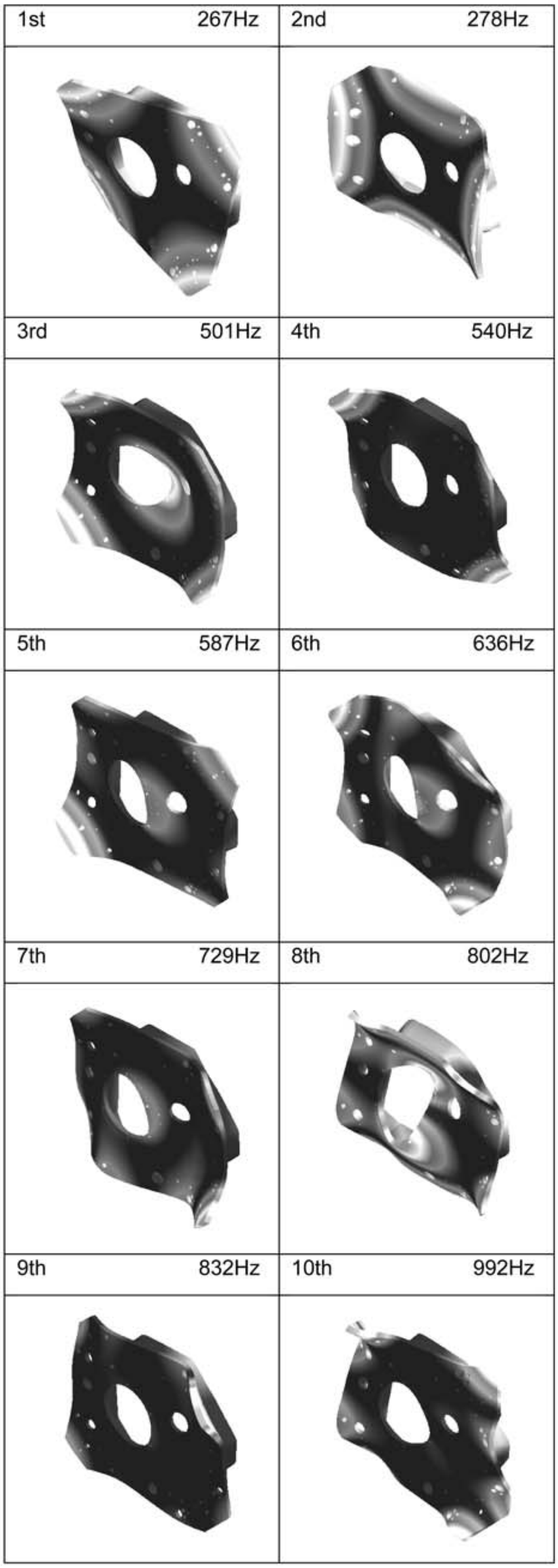


Figure 5: Mode shapes of the base plate

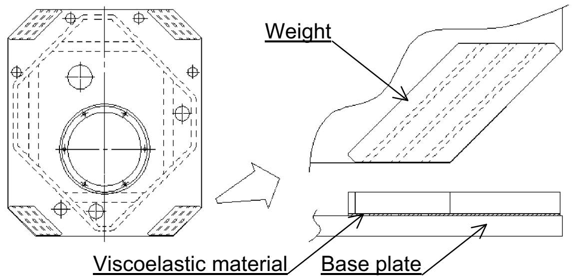


Figure 6: Structure of TMD

depends on the frequency, and the properties are expressed with classic Maxwell or Kelvin-Voigt model, or fractional derivative model [5]. However, the frequency range studied in this paper is within one order from 200Hz to 1 kHz, and the damping coefficient does not differ in large. While, one degree-of-freedom vibration model is constructed with the base plate and TMD. The natural frequency of TMD should be adequately adjusted to restrict the vibrations of the base plate. For example, to decrease the natural frequency of TMD, the area of the viscoelastic material would be smaller, or the mass would be heavier. However, there is the practical design restriction to increase the height of the mass, or to change into heavier materials. On the contrary, the area of the viscoelastic material should not be smaller, because its damping property is used. Thus, in this paper, the "shape effect" was used [9]. The formula of Hattori and Takei [10] about the infinity length rectangular material could be applied to the rectangular elastomer shown in Figure 7. The nondimensional description of this is shown as Equation 1. This formula is possibly used when the ratio of the width *a* mm against the length *b* mm is under 1/3 or over 3.

$$E^{(0)}{}_{ap} / G = 4 + 3.290 S^2 \quad (1)$$

In Equation 1, $E^{(0)}{}_{ap}$ is the pseudo initial Young's modulus Pa, G is the modulus of transverse elasticity Pa, $S=a/2h$ when the height of the material is h mm. From Equation 1, $E^{(0)}{}_{ap}$ is increased in proportion to a^2. Therefore, the spring constant, in other words, the natural frequency of TMD could be decreased by setting a smaller without decreasing the total area of the viscoelastic materials. This will be experimentally confirmed in Section 3.3.

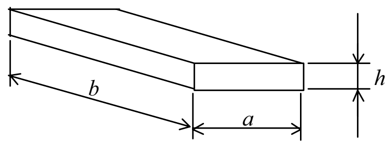


Figure 7: Shape of the viscoelastic material

3.3 Results of the Experiments

Natural Frequency and Damping Coefficient of TMD

The developed TMD should restrict several eigenmodes of the base plate. To realize this property, the vibration transmissibility of TMDs were measured when the width: *a* was divided into 1/2, 1/4, 1/8, 1/16 respectively, although the total area of the viscoelastic material was the same. The measurements were conducted with two accelerometers by impulse force input to the base plate. The height of the viscoelastic materials: *h* was 1, 2, 3 mm each, and total width

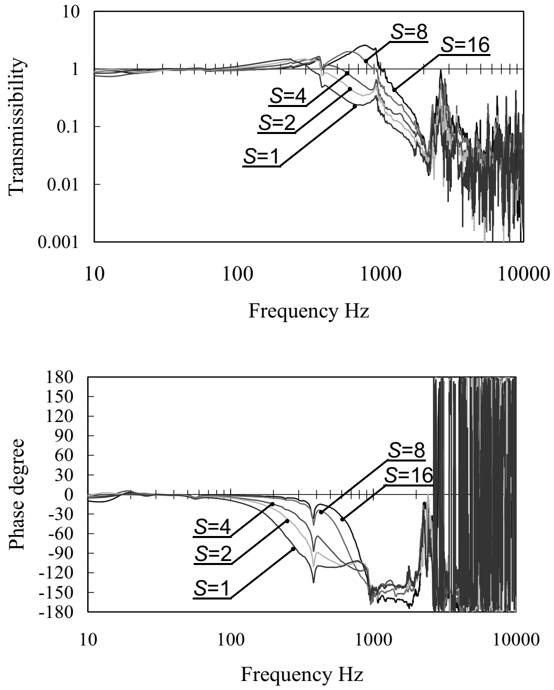


Figure 8: FRFs of the damper as a function of *S* (*h*=2mm)

of the materials was 64 mm. Figure 8 shows the frequency responses of TMDs when *h* =2 mm. As shown in this figure, the natural frequency of TMD showed lower as the *S* (explained in Section 3.2) was set up to lower value by changing the width *a*, although the total area of the materials were the same. The "shape effect" was shown in quality. Moreover, when *S* was set up to the value 16 and 8, the magnitude of the resonance was larger, and the phase shift was more rapid against the frequency as increasing *S*, comparing to the cases when *S* were less 4. The same trend was shown when the height *h* was changed. Thus, the "shape effect" on the damping coefficient was also observed. In the case when *S* is larger than certain value, the damping coefficient of TMD shows smaller value and the active frequency range will be narrower [5]. Therefore, it was not adequate to our objective. Figure 9 shows the variation of the natural frequency of TMD by the estimation in comparison with the measured values. The estimation has conducted by setting the natural frequency of TMD to 1, when the viscoelastic material was divided into 16 parts. The dashed lines show the estimated values with Equation 1, and solid lines shows the measured values. When *h*=3 mm, the normalized natural frequencies were fit at *S* =0.67, 1.3, although they were larger than estimation when *S* were bigger than 1.3. When *h* were lower than 3, the differences from estimations were larger. As the thickness of the material was thinner, and as the number of division was smaller, the spring constant of TMD was larger and the deflection of the viscoelastic materials would be smaller. Therefore, the conditions of the contact areas of the viscoelastic materials were less uniform, and it might be one of the occasions for this phenomenon. In these results, it must be better for wide frequency range damping to set *S* less than certain value and stabilize the properties of TMD; low response at the

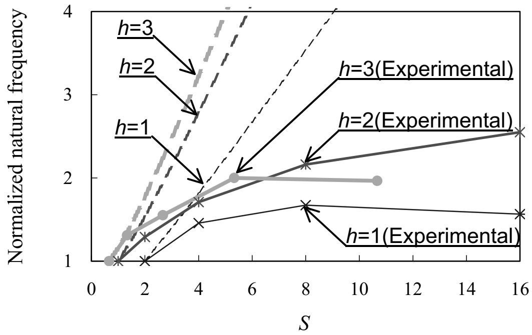


Figure 9: Shape effect on natural frequency of TMD

resonance frequency, and slow phase shift against the frequency.

To compare the frequency properties of damping, one TMD was set on the edge of the base plate. Under random acoustic exposure from the front of SEM, the vertical vibration on the edge was measured with the accelerometer. In this paper, we defined "vibration-absorbing coefficient". This is the ratio of the difference between the acceleration spectra without and with TMD against the spectrum without TMD. For example, when the acceleration shows the value 0 with TMD, the vibration-absorbing coefficient is 1. When TMD increased the acceleration, the value would show negative value. Careful attention should be paid to the reliability of the value when the acceleration shows small value without TMD because it contains differential value. Using vibration-absorbing coefficient makes it easier to compare the frequency properties of TMDs. In general, the performance index of TMD was the gain at the objective frequency. However, in practical applications, the structure has several vibration frequencies to be restricted, and the values show diversity and difference in conditions. Therefore, it is important to be useful at wider frequency range. There would be no need to improve the reduction rate more, when the vibration-absorbing coefficient was close to 1 at the objective frequency. Figure 10 shows the vibration-absorbing coefficient at *h*=2 mm and *S* was 1,4,16 each. When *S* was 16, low coefficient at 250 Hz was shown, because the active frequency range of TMD was too high. On the contrary, when *S* was 1, the active frequency range moved to lower and the vibration-absorbing coefficient at higher frequency were

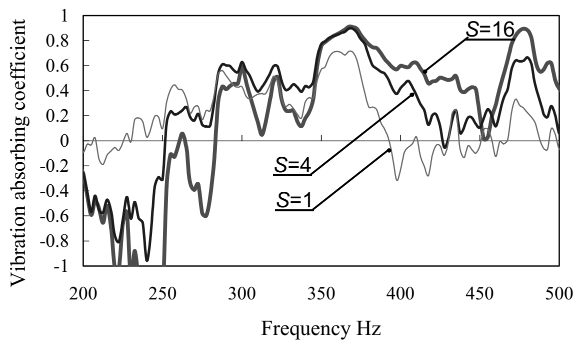


Figure 10: Vibration-absorbing coefficient with 1 damper under acoustic noise excitation as a function of *S* (thickness:2mm)

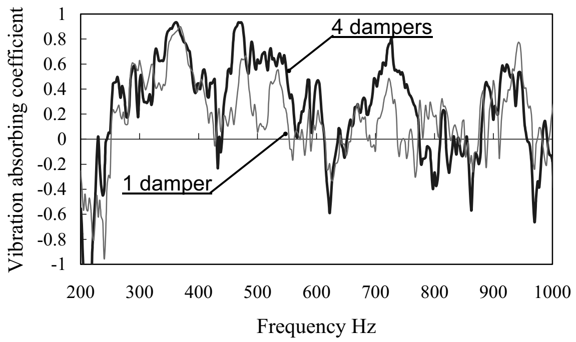


Figure 11: Vibration-absorbing coefficient with 1 and 4 dampers under acoustic noise excitation (thickness:2mm, *S*=4)

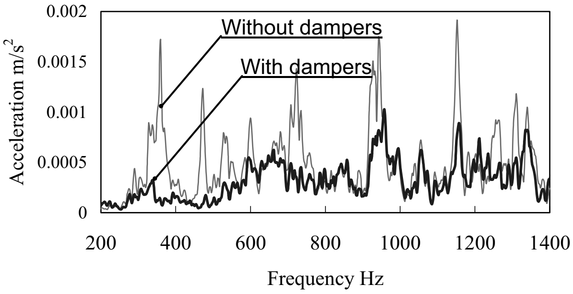


Figure 12: Reduction of the acceleration on the base plate with 4-dampers

decreased. Thus, TMD effective from 250 to 500 Hz (*h*=2 mm, *S*=4) seemed to be optimum. After that, 4-TMDs were installed to SEM. Figure 11 shows the vibration absorbing coefficient in comparison with 1 TMD. Larger coefficients were shown with 4 TMDs at from 200 Hz to 1 kHz. Figure 12 shows the comparison of the acceleration spectra between with and without TMDs. Remarkable wide range reduction of the vibration peaks from 250 Hz to 1.3 kHz was shown. The amplitude at 360 Hz without TMDs was 0.34 nm as displacement. Thus, the vibration in nm order was reduced. This wide range damping effect shows the same or more usefulness at lower cost in comparison with using expensive high damping materials. Moreover, the increase of the mass is only 5% of the base plate.

Comparison of the Image Noise under Acoustic Excitation

Under 1/3 octave band pass filtered random acoustic noise exposure, SEM images were captured and compared with and without TMDs. The frequency range was set from 100 Hz to1 kHz and equivalent sound pressure revel on flat scale was 85 dB. As a result, remarkable image noise reductions were shown at 315 and 400 Hz band. Figure 13 shows the comparison of the images under 315 Hz band noise acoustic exposures. There was no increase of image noise at any frequency band. Because 315 Hz was most sensitive frequency band for this model, it was effective to increase the allowance level of acoustic noise. Although the accelerations at higher frequencies than 500 Hz were attenuated, no image noise reductions were shown. The reason might be that the main path of the vibration transmissibility was not from the base plate.

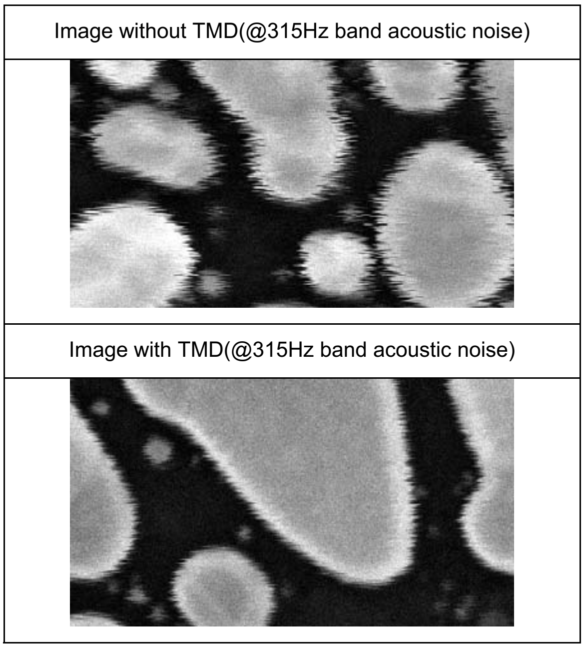


Figure 13: SEM images under 1/3 octave band acoustic noise (315Hz band)

4 CONCLUSIONS

Concerning about the SEM, vibrations of the base plate excited by acoustic noise cause image noise. As a countermeasure, TMD using viscoelastic material was applied to the base plate. The results were summarized as follows:

1. A TMD effective from 250 Hz to 1.3 kHz at the small vibration amplitude in nm order has been developed.
2. Using “shape effect”, the natural frequency of TMD was adjusted to the optimum without changing total area of the viscoelastic materials.
3. “Shape effect” on damping coefficient was also shown. This is useful for wide effective range of TMD.
4. The reduction of SEM image noise under 1/3 octave band pass filtered random acoustic noise exposure has been confirmed at 315 and 400 Hz band.

This TMD [2] has been already applied to the equipment on the market. Remained subjects are: study about the frequency properties of TMDs using other various viscoelastic materials, and research to design *S* without try-and-error method.

REFERENCES

[1] JEOL ltd. (Hamochi, M.), 2004, Base plate for Isolation of Vibration, JP, 3581257.

[2] JEOL ltd. (Hamochi, M.), 2005, Base plate for Isolation of Vibration, JP, 3741558.

[3] Hamochi, M., Wakui, S., 2007, Image Noise Reduction Using Vibration Mode Shaping for Scanning Electron Microscope, Journal of JSPE, 73/6:665-670.

[4] Hamochi, M., Oohashi, S., Suga, M., Wakui, S., 2007, Image Noise Reduction Using Electron Beam Deflection for Scanning Electron Microscope”, Journal of JSPE, 73/9:1080-1084.

[5] Jones, D. I. G., 2003, Handbook of Viscoelastic Vibration Damping, Maruzen & Wiley, 50-51.

[6] Kunieda, M., 2001, Practical Study of Mechanical Vibration, Rikougaku- sya, 80-81.

[7] The Japan Society of Mechanical Engineers, 1998, New Technology fusion series Vol.4 Technology for vibration damping, Youkendou, 126-128.

[8] I・T・O, 2000, Handbook for Practical Technology of Acoustic Noise and Vibration for Engineer and Designer, N・T・S, 538.

[9] Togawa, H., 1975, Rubber for Vibration Isolator , Gendai-kougaku-sya, 97.

[10] Hattori, R., Takei, K., 1950, the Journal of SRIJ, 23:194-200.

Planning

A Model for Adaptively Generating Assembly Instructions Using State-based Graphs

Michael F. Zaeh[1], Mathey Wiesbeck[1]

[1] Institute for Machine Tools and Industrial Management, Technische Universität München, Munich, Germany

Abstract

Traditional systems for digital assistance in manual assembly, e.g. optical displays at the work place, are inherently suboptimal for providing efficient and ergonomically feasible worker guidance. The display of sequential instructions does not offer an increase in productivity beyond a certain degree. Little situational support and the resulting deterministic guidance lead to a reduced acceptance by the worker. A solution to this discrepancy is seen in adaptive and cognitive systems of worker guidance. In this context, the paper presents a process model for adaptively generating assembly instructions. It is part of an integrated framework for human worker observation and guidance based on state charts.

Keywords:

Assembly, Information, Flexibility

1 INTRODUCTION

Skilled human workers turn mechanical workshops into today's most flexible and widely applicable form of production. This flexibility and generality however come at very high production costs. These restrict the use of mechanical workshops to building prototypes and a limited range of specialized and valuable products. The source of flexibility in this case can be easily identified: the cognitive capabilities of the human workers that operate it. Humans can perceive their environment, plan actions, learn and adapt behaviors and they can interact in multiple ways with their surroundings. Moreover, they can do this robustly despite changing contexts and situations. The realization of comparable cognitive capabilities in technical systems therefore bears an immense potential for the creation of industrial automation systems that are able to overcome today's boundaries.

2 COGNITION AND COGNITIVE TECHNICAL SYSTEMS

Cognition is in the scope of investigation of several scientific disciplines. Among these are cognitive psychology, cognitive sciences, cognitive engineering, cognitive ergonomics and cognitive systems engineering. While robots have learnt to walk, navigate, communicate, divide tasks, behave socially and play robot soccer [1], only a few examples for the application of artificial cognition in manufacturing exist. Therefore, several authors emphasize the need for cognitive systems to overcome deficiencies in automation [2], human-robot-interaction [3] and planning [4].

Cognitive Technical Systems are hereby equipped with artificial sensors and actuators, are integrated and embedded into physical systems and act in a physical world. They differ from other technical systems in that they perform cognitive control and have cognitive capabilities. Cognitive control comprises reflexive and habitual behavior in accordance with long-term intentions. Cognitive capabilities such as perception, reasoning, learning and planning turn technical systems into ones that 'know what they are doing'. More specifically, a cognitive technical system becomes a technical system 'that can reason using substantial amounts of appropriately represented knowledge, learns from its experience so that it performs better tomorrow than it does today, explains itself and can be told what to do, is aware of its own capabilities and reflects on its own behavior, and responds robustly to surprise'. Technical systems being cognitive in this sense will be easier to interact and cooperate with and they will be more robust, flexible and efficient.

The present research is carried out within the CoTeSys Cluster of Excellence at Munich, Germany. It aims for a considerable utilization and establishment of cognitive capabilities in technical systems.

3 STATE OF THE ART

Currently, a well-experienced work force on the shop floor is required to dynamically adapt to the current state of the process [5]. Consequently, complex capital products, e.g. machine tools, are manufactured in continuous manual assembly by teams of expert technicians. The reason for this originates from the failure to succeed in a detailed preliminary planning and usage of such to drive an guidance system because of the high dissimilarity and complexity of the goods [6]. Conventional, similar systems lack the ability of cognition, leading to serious deficiencies when acting in complex environments, especially in the context of human-machine-interaction [7].

Known systems for worker guidance operate on manually derived and statically compiled assembly plans [8][9].

Furthermore, literature shows mainly technical implementations of guidance systems based on Augmented Reality (AR). However, a successful application has to support the planning entity methodically sound. Only in this way, new technologies for efficient worker guidance will be usable [10].

4 ADAPTIVE INFORMATION PROVISIONING IN MANUAL ASSEMBLY

4.1 Motivation

Several authors have attempted to describe and model the process of decision-making in the field of cognition [11][12], while others have characterized its role in information processing [13] and as part of a larger cognitive architecture [14]. However, these works focus mainly on modeling concepts for interpreting the cognitive processes of humans to trigger respective actions. A framework for modeling the cooperational behaviors in production environments is lacking. Hence, the integration of cognitive machines in currently humanly dominated, continuous manual assembly environments is the core issue of the subproject ACIPE (Adaptive Cognitive Interaction in Production Environments).

A crucial goal is to develop concepts and technical implementations essential for the Cognitive Factory, which will allow for the adaptive generation of assembly instructions. Adaptive in this context meaning the integration of factors of the production environment and the state of the human for determination of the current set of instructions, as well as using the provided sensory information to accommodate the output to the human worker accordingly. This knowledge will allow for interactive multimodal guidance and support of the worker in manual and semi-automatic assembly.

4.2 State-based Graphs for Generation of Assembly Instructions

Adaptive Provisioning of assembly-related information

Representative digital systems for assistance in manual assembly situations in their current concepts are laid out suboptimally. These systems do not offer an efficient and at the same time ergonomical worker guidance [16]. Indeed, a reduction of the mental strain as well as an increased performance for individual tasks can be exhibited [17]. However, the utilization of Augmented Reality (AR) systems in the context of workflow related series of instructions does not lead to the expected increase in productivity [18][19].

The occurring phenomenon of “attention tunneling” describes an overly stimulation of the receptiveness of a human worker. As a consequence, a distraction from relevant and important influences of the physical environment arises. This incident is in part caused by the purely deterministic nature of assembly planning. Moreover, the generation of assembly instructions happens without the comprehension of the actual production environment. The little situational support of the worker results in a rigid guidance and leads to the phenomenon mentioned above. Consequently, there is minor to no acceptance for such a system on the part of the worker [20]. A key to solving these discrepancies is seen in the application of adaptive and cognitive methodologies for the generation of assembly tasks. These require environmentally-dependant and situation-dependently triggered paths on state-based graphs.

Application of State-based Graphs for the Generation of Assembly Instructions

The basis for the generation of task instructions is an st-graph (see Figure 1). A st-graph G is a acyclic directed graph G with exactly one source vertex *s* and exactly one sink vertex *t*, which is embedded in the plane such that s and t are on the boundary of the external face [21]. In contrast, undirected graphs merely connect the vertices, without any consideration for direction. As such, the latter is not suitable to deliver a set of instructions for reaching the state of a desirably assembled product from a given state. An acyclic graph is defined as a graph with no path that starts and ends at the same vertex [22]. In its current concept, the graph shall not allow disassembly. Therefore, the acyclic property of the graph is mandatory. The abstraction of the formal definition above is shown in Figure 1.

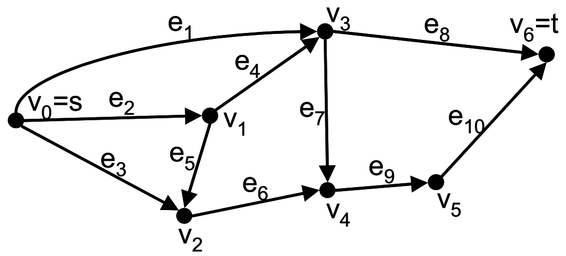


Figure 1: A (planar) st-graph with vertices v and edges e.

In accordance with Figure 2, the source vertex *s* represents the initial state of the product to be assembled (state z_{start}) and the sink vertex *t* represents the target state of the fully assembled product (state z_{target}). The edges of the graph (e. g. $A_{start,I}$) symbolize assembly task instructions. The execution of such by the worker will transfer the work piece in focus from one state to another (e. g. z_{start} to z_I).

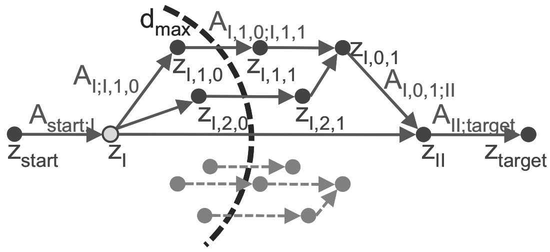


Figure 2: Schematics of a state-based assembly graph.

Every product’s set of instructions is defined by an st-graph. The product-specific graphs are deducted from construction- and assembly-related information. The path from the initial state z_{start} to the target state z_{target} leads across the respective major intermediate states z_{start} to $z_{target-1}$ (z_{II}). The minor intermediate states shown in Figure 2 ($z_{I,1,0}$, $z_{I,1,1}$,...) are reachable via an increased degree of detail of the assembly task instructions. If a vertex has more than one outgoing edge $A_{i,j}$, then alternative assembly sequences or alternative parts can be disposed of.

The graph will not allow for cycles, under the presupposition that a disassembly of products to be built shall not be possible. This simplification entails, that the target state of the fully assembled product (state z_{target}) is reachable on a path from every state of the graph. Therefore, there are no states which would hinder a further assembly towards the target state. A future extension of the presented concept might well account for disassembly and is in the scope of the current research activities.

Assembly Task Generation

The task instruction to be displayed is determined by an optimization algorithm regarding shortest paths on the graph. The course of action for the task generation is as follows:

- Data about the current state of the product is delivered by the *environment model* (scene representation).
- The received data is matched against that of known product states within the existing graph.
- If an exact match is found, this is the vertex of origin for the further calculations. If no exact match can be determined, then an unknown state of the product is entered into the graph.
- A suitable shortest path algorithm is applied on the graph starting from the vertex z_i of a known state of the product, running a path to the target state z_{target}.
- The information attached to the outgoing edge $A_{i,i+1}$ on the shortest path from the vertex z_i to the vertex of the target state z_{target} corresponds to the assembly task information to be displayed.
- After the information was transmitted to the *cognitive human model*, executed by the worker and a varying current state is determined by the *environment model*, the matching of the state (now z_{i+1}) is executed again.

As noted, if no exact match of the currently acquired product state exists, then a new state is adjoined to the assembly graph. However, the vertex representing this state does not have any outgoing or ingoing edges. It is consequently not connected to the operating graph of the product to be assembled. Methods for deriving assembly task instructions from such a state are under investigation. These will create a feasible edge from the current vertex to a vertex on the graph from which a path to the vertex of the target state z_{target} can be determined.

It is easily seen from the above elaboration that unforeseen and varying task executions of the human worker lead to a non-deterministic path of actions towards the fully assembled product.

Figure 3: Worker at the adaptive work place.

Integration of Information of the Production Environment

Information and factors of the production environment influencing the manual assembly work place in focus are integrated into the state-based graph by an extended definition of the edges (corresponding to the task instructions). The edges of the graph are weighted subject to data of the process-relevant environment of the human worker. This weighting reflects the costs incurring for the execution of an assembly task. Possible factors for this include e. g.:

- Quantification of the elemental task according to a predetermined time system such as methods-time-measurement (MTM) or Work-Factor: This allows for an accurate modeling of single tasks a well as of the assembly plan as a whole in regard to an approved system of productivity measurement.
- Quantification of the elemental task according to prior observation and stored knowledge: Integration of the performance of individual workers offers a means for determining feasible instructions to display on a personal basis.
- Availability of parts to be assembled within an assembly task: On the one hand, the utilization of this data can trigger the selection of a path on the graph to reach the target state of the fully assembled product using alternative components. On the other hand, information for the assembly of an alternative product variant based on the progress up to now can be derived by matching the product's current state to other products' state-based graphs.

Integration of Information regarding the human worker

Information regarding the mental as well as the physical state of the worker is integrated into the task retrieval process. The scope of integration depends on the availability of recognizer technologies and their capabilities. A feasible approach therefore is currently developed in the course of a co-project [23]. The work of the present paper implies the results of the aforementioned project.

Utilization of information about the worker current mental and physical state is done in a supportive, yet unobtrusive, way. After the shortest path has been determined (see Assembly Task Generation), relevant information about the worker (e. g. receptiveness, mental strain or secureness) is merged into a parameter d_{max}. The value of d_{max} and its input variables are generated on the basis of neuro-psychological research results achieved in the course of the CoTeSys cluster [24]. In respect to the task generation within the graph, d_{max} restricts the maximum step-length on the shortest path towards the target state z_{target} (Figure 2). As such it represents the degree of detail of the task outputted to the relevant model.

4.3 Overall Framework / System Context

Among research done regarding the present paper, three major models have been defined. The implementation and integration of these concepts is currently in progress (e.g. work place as seen in Figure 3).

First of all there is the *process model* outlined in this paper. It represents the knowledge about the work plan and contains the required steps to complete the building process of a work piece. To aquire state information relevnt to the current work step, an *environment model* is needed. It is provided with different kinds of sensor data. These are used to give the system status signals about the assembly step the worker currently is at and how far the work piece has been completed. To know what kind of information the worker currently needs to successfully complete her/his work, a mental representation of him is also required. This is realized in the *cognitive human model* gained by experimental data combined with knowledge about ergonomic factors. Figure 4 shows central aspects of the models in keywords.

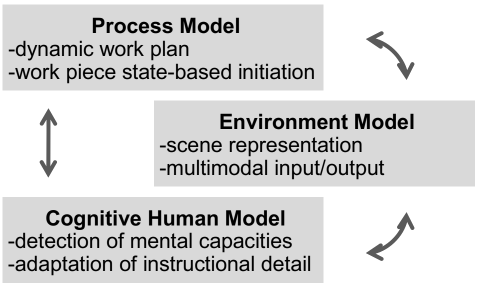


Figure 4: Key features of the framework's individual models.

5 SUMMARY

It is the goal of project ACIPE to bring aspects of cognition and adaptability into guidance systems of manual assembly work places. In this paper, a process model utilizing state-based graphs to generate assembly task instructions adapted to the human worker and its context was presented. The necessity and feasibility thereof was shown and the approach and its workflow were further detailed. Moreover the process model's integration into a framework for the adaptive generation of assembly task instructions was elaborated.

6 ACKNOWLEDGMENTS

We would like to thank Professor Rigoll, Professor Bubb, Professor Müller and the coordinator of the research cluster Professor Buss and their respective institutions for their scientific support. In particular, we would like to acknowledge the funding support of the Deutsche Forschungsgemeinschaft (DFG).

7 REFERENCES

[1] Mataric, M.J., 1998, Behavior-based robotics as tool for synthesis of artificial behavior and analysis of natural behavior, Trends in Cognitive Sciences, 2:82-87.

[2] Putzer, H., Onken, R., 2003, COSA - A generic cognitive system architecture based on a cognitive model of human behavior, Cognition, Technology & Work, 5:140-151.

[3] Hoc, J.M., 2001, Towards a cognitive approach to human-machine cooperation in dynamic situations, International Journal of Human-Computer Studies, 54: 509-540.

[4] Shalin, V.L., 2005, The roles of humans and computers in distributed planning for dynamic domains, Cognition, Technology & Work, 7:198-211.

[5] Kammüller, M, 2006, Synchrone Produktion im Werkzeugbau, Münchener Kolloquium, Utz, Munich, 103-108.

[6] Thaler, K., 1993, Regelbasierte Verfahren zur Montageablaufplanung in der Serienfertigung, Springer, Berlin.

[7] Längle, T., Wörn, H., 2001, Human-Robot Cooperation Using Multi-Agent-Systems, Journal of Intelligent and Robotic Systems 32/2:143-160.

[8] Reiners, D., Stricker, S., Klinker, G., Müller, S., 1999, Augmented Reality for Construction Tasks: Doorlock Assembly. Augmented Reality Placing Artificial Objects in Real Scenes, Natick, MA, 47-60.

[9] Haringer, M., Regenbrecht, H., 2002, A Pragmatic Approach to Augmented Reality Authoring, Proc of the International Symposium on Mixed and Augmented Reality, Darmstadt, 237.

[10] Patron, C., 2004, Konzept für den Einsatz von Augmented Reality in der Montageplanung,Utz, Munich.

[11] Hammond, K.R., 2000, Judgments under stress, Oxford University Press, New York, NY.

[12] Speed, A., Forsythe, J.C., 2002, Human emulation technology to aid the warfighter: Advances in computational augmentation of human cognition, Whitepaper, Sandia National Laboratories.

[13] Deutsch, J.A., Deutsch, D., 1963, Attention: some theoretical considerations, Psychological Review, 70:80-90.

[14] Leiden, K., Laughery, K.R., Keller, J., French, J., Warwick, W., Wood, S.D., 2001, A Review of Human Performance Models for the Prediction of Human Error. NASA, System-Wide Accident Prevention Program, Ames Research Center.

[15] Neufeld, R.W.J., 1999, Dynamic Differentials of stress and coping, Psychological Review, 106:385-397.

[16] Reinhart, G., Patron, C., 2003, Integrating Augmented Reality in the Assembly Domain - Fundamentals, Benefits and Applications, Annals of CIRP 52/1:5-8.

[17] Tang, A., Owen C., Biocca, F., Mou, W., 2003, Comparative Effectiveness of Augmented Reality in Object Assembly, Proc of the SIGCHI Conference on Human factors in computing systems, Ft. Lauderdale, FL, 73-80.

[18] Dopping-Hepenstal, L.L., 1981, Head-up displays: The integrity of flight information, IEEE Proc Part F, Communication, Radar and Signal Processing 128/7: 440-442.

[19] Yeh, M., Wickens, C.D., 2000, Attention and Trust Biases in the Design of Augmented Reality Displays, Technical report, Aviation Research Lab.

[20] Livingston, M.A., 2005, Evaluating Human Factors in Augmented Reality Systems, IEEE Computer Graphics and Applications 25/6:6-9.

[21] Roberto, T., 1996, On-line Planar Graph Embedding, Journal of Algorithms, 21/2:201-239.

[22] Black, P.E. (ed.), 2004, Dictionary of Algorithms and Data Structures, U.S. National Institute of Standards and Technology, Gaithersburg, MD.

[23] Wallhoff, F., Ablaßmeier, M., Bannat, A., Buchta, S., Rauschert, A., Rigoll, G., Wiesbeck, M., 2007, Adaptive Human-Machine Interfaces in Cognitive Production Environments, Proc of the International Conference on Multimedia & Expo, Beijing, China, 2246-2249.

[24] Stork, S., Stößel, C., Müller, H. J., Wiesbeck, M., Zaeh, M. F.; Schubö, A., 2007, A Neuroergonomic Approach for the Investigation of Cognitive Processes in Interactive Assembly Environments, Proc of the 16th IEEE International Symposium on Robot and Human Interactive Communication, Jeju Island, Korea.

Evaluating Assembly Instruction Methods in Cell Production System by Physiological Parameters and Subjective Indices

Nuttapol Pongthanya[1], Feng Duan[2], Jeffrey Too Chuan Tan[2], Kei Watanabe[2],
Ye Zhang[2], Masao Sugi[2], Hiroshi Yokoi[2], Tamio Arai[2]

[1] Department of Systems Innovation, Faculty of Engineering, The University of Tokyo, Japan

[2] Department of Precision Engineering, School of Engineering, The University of Tokyo, Japan

Abstract

Providing assembly information to support human workers has been verified to be an effective way to improve the assembly performance. However, the relationship between the assembly information format and the human workers' fatigue has not been investigated clearly. This paper is aimed to evaluate assembly instruction provided by information supporting system in fatigue aspect by using physiological measurement. The task performance and subjective rating measurements are taken into account to verify the physiological responses. The experimental results show that image and text instruction have more potential to convey the assembly information than voice instruction in this setting.

Keywords:

Cell Production System; Assembly Instruction; Mental Fatigue

1 INTRODUCTION

The traditional automated manufacturing line system is designed to produce single specific products by special machines on a mass basis. However, line production system is not flexible to adapt to the fast changing consumers' demands. Therefore, cell production system is purposed in order to overcome this demerit [1]. In the cell production system, a single human operator manually assembles each product from start to finish [2], [3]. Multi-skilled human workers amplify the flexibility of the cell production system to meet the requirements of manufacturing diversified products and quantity.

Many studies on cell production design have been done to improve collaboration between human operator and support machines [1] [3]. Assembly information support is another important research in industrial manufacturing design. In 1991, Gery had developed Electronic Performance Support System (EPSS) as an electronic information platform to capture, store and distribute knowledge from one operator to others [4]. Based on this idea, Kasvi et al. introduced the Information Support System (ISS) which can provide information based on operator's need [5]. However, most of the research objectives concentrate on the improvement of manufacturing productivity. Since human operator plays an important role in the cell production system, the fatigue condition of the human operator cannot be neglected.

Fatigue can be classified into two categories: which are physical and mental fatigue [6]. Physical fatigue causes the reduction of performance in muscular system. On the other hand, mental fatigue is observed in a sense of weariness and reduced alertness [6]. Fortunately with the development of automation system, most physical demanding jobs are diminished from manual assembly processes. Hence, this research work mainly concentrates on the mental fatigue in the production operation.

The main objective of this paper is to evaluate the effectiveness of the assembly instructions in the physiological aspect. The relation between different types of instruction and the level of fatigue experienced by workers is also presented in this paper.

In order to evaluate mental fatigue, physiological parameters, task performances and subjective indices are utilized. The methodologies, experiment setup and the formats of assembly instruction used in this study are discussed in Section 2. Section 3 explains the experimental results. Finally, the conclusion and future work are given in Section 4.

2 METHODOLOGY

2.1 Experiment Setup

Tasks

In this study, in order to examine the relation between mental fatigue and task performances, a mental workload experiment is conducted. This experiment consists of two different tasks: a pick-and-place task and a colour-count task. The pick-and-place task is chosen because of its similarity to a connector-assembly, which is a usual operation in assembly process.

For the pick-and-place task, the subjects were instructed by a predefined assembly instruction to select a coloured spherical magnet and place it on the specific location as shown in Figure 1. For as the colour-count task, the subjects calculate the total number of magnets for each colour they placed during the pick-and-place task.

Figure 3 demonstrates the experimental procedure which starts from fifteen-minute equipment preparation following by six successive sessions of 100-trial task set. Totally, the subjects are required to perform this pick-and-place task for 600 trials with five-minute break after each task session is finished.

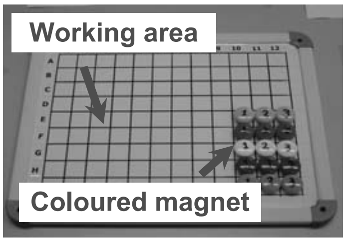
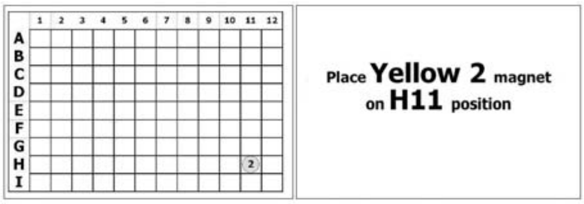


Figure 1: Working area.



(a) (b)

Figure 2: Assembly instructions.

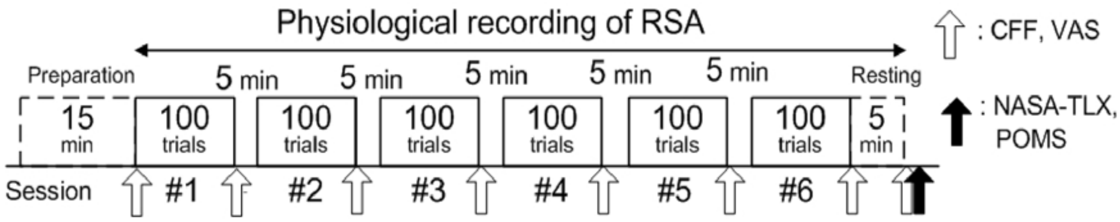


Figure 3: Experimental procedure.

Instruction types

Three different types of instruction: text, image and voice are used in this study. For image instruction experiment, the simulated image of working area is displayed on the screen of the monitor as shown in Figure 2(a). The similar setup is used in text instruction experiment. Bigger font size is used to highlight the important words as illustrated in Figure 2(b). Finally, the same sentence from text instruction is recorded into audio and played back through the speaker in the voice instruction experiment.

Subjects

Five healthy males (with ages 23±1.3) participated in this study. The experiment is conducted in two different conditions, which are actual-condition (A-condition) and control-condition (C-condition). The subjects remain sitting posture throughout the experiment.

In A-condition, the subjects have to execute both a pick-and-place task and following with a colour-count task at the same time for each instruction format (referred to: 'A-image', 'A-text' and 'A-voice'). On the other hand, only a pick-and-place task is executed during the C-condition experiment (referred to: 'C-image', 'C-text' and 'C-voice').

2.2 Mental fatigue evaluation

Interest in human fatigue has been demonstrated a long time ago and actual documentation dates as far back as World War I [6]. Many researchers who study on ergonomic field are interested in the quantitative measurement of mental fatigue. Thus, several methodologies to evaluate mental fatigue are developed in various aspects. The methods used in this paper can be categorized into task performances, subjective indices and physiological parameters.

Task performances

Task performances are calculated by the number of mistake in a pick-and-place task and the response time used in each trial.

Subjective indices

Subjective indices consist of direct or indirect queries to the individual perception of the workload encounter during any task [6]. The simplest way to estimate the mental workload is consult the human operator directly about the mental stress suffered during the task. In order to obtain subjective measurement, three following subjective indices are employed: NASA-task load index (NASA-TLX); Visual Analogue Scale (VAS) and Profiles of Mood States (POMS).

NASA-TLX classifies the degree of mental workload into six important factors which are mental demand, physical demand, temporal demand, performance, effort, and frustration [7].

VAS is the pain index of the subjects feel. POMS is used to measure mood status of the subject after performing the mental workload.

In this experiment, the subjects rate VAS during every rest period (eight times), whereas, the NASA-TLX and POMS are only conducted once at the end of experiment.

Physiological parameters

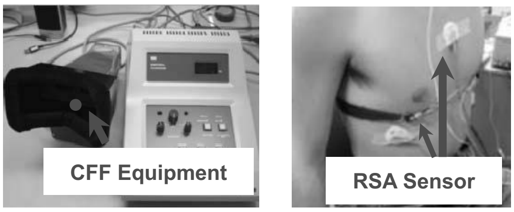
Physiological parameters are used to objectively evaluate mental fatigue in numerical figure. These kinds of measurement seem to be the most promising method for the development of a real time vital sign monitoring system.

In this experiment, the following two physiological parameters is collected during experiment to indicate mental fatigue: Critical Flicker Fusion frequency (CFF) (Figure 4 (a)); and Respiratory Sinus Arrhythmia (RSA) (Figure 4 (b)).

Critical Flicker Fusion frequency (CFF) is the frequency of an intermittent light stimulus, when it appears to be completely steady to the subject. This rate, expressed in Hz, is commonly referred to as the "threshold frequency". The CFF is widely used in the study of human mental fatigue, because it can be obtained quickly and easily. The CFF is relevant to occupational health and safety and could be interpreted as the working performance of human workers. Hence, this technique is commonly applied to analysis the level of mental fatigue.

Heart Rate Variability (HRV) is used to observe physiological condition of human operator. The high-frequency (HF) component of HRV is respiratory modulation of the heart rate. This HF component is physiologically known as Respiratory Sinus Arrhythmia (RSA) [8]. RSA refers to the slowing down of heart rate during expiration and speeding up of heart rate during inspiration. In many medical studies, RSA represents the large fluctuation of HRV that is moderated by cardiac vagal activity. From the physiological point of view, the amplitude of RSA is used for estimating mental workload. As reported by Kotani [8], the amplitude of RSA decreased as the mental workload level increased.

In this paper, this approach is utilized to evaluate the level of mental fatigue caused by different types of assembly instruction.The RSA is obtained throughout the experiment, while the CFF is measured at the beginning of each break.



(a) (b)

Figure 4: Physiological change monitoring equipment.

3 RESULT AND DISCUSSION

3.1 Task performance

Figure 5 (a) shows the average response time in each task session. Although the workload level of mental workload tasks remains constant, all subjects can execute the same task in shorter response time. This implies that the tasks are relatively simple; therefore the subjects can improve their performance in a short period of time. Therefore, the response to the same tasks became faster as the operation continued.

The image and text instruction show almost similar results while the voice instruction seems to require longer time to complete the same task.

In comparison of A-condition and C-condition, the additional mental stress caused by a colour-count task which retarded the operation was obviously observed in all formats. It can be observed that the correct rate statistically decreased in voice instruction experiment for both A-condition and C-condition.

However, the correct rate does not significantly change in image and text instruction experiment, as illustrated in Figure 5 (b). There is no distinctive difference of correct rate between A-condition and C-condition in all instruction formats.

3.2 Subjective indices

Figure 5 (c) shows the results of VAS rating score which was obtained throughout the experiment.

In this chart, the difference between A-condition and C-condition is clearly observed. The subjects seems to rate mental fatigue in the same level as the experiment proceeded. Nevertheless, there is no significant difference of image, text and voice instruction when all instructions were studied under the same condition.

The average of NASA-TLX total score in each instruction is: 64.8 ± 17.0, 70.8 ± 14.7 and 77.7 ± 16.9 for image, text and voice instruction respectively. On the other hand, the average of POMS score is: 13.8 ± 8.5 for image instruction; 13.6 ± 6.8 for text instruction and 16.8 ± 6.5 for voice instruction. The results are concerned to be relatively high where the average score for non-fatigue condition of POMS test is around 9.3 ± 6.2 [9]. In subjective rating, image and text instruction are more comfortable for delivering the assembly information to the subjects when compare to voice instruction.

3.3 Physiological parameters

Figure 5 (d) shows the change of CFF in percentage value. Opposed to other evaluation techniques, the CFF indicates that the image and text instruction cause higher mental workload than voice instruction. These results seem to have been caused by using a monitor to display the assembly information during the image and text instruction experiment, which directly affects CFF [10] [11].

The change of RSA value is illustrated in Figure 5(e). Although the decrease of RSA can be observed in some subject, the significant correlation has not been found. Figure 5 (f) is the change of RSA value that obtained from a subject. The decrease of amplitude of RSA can be observed in voice instruction experiment; however, the same inclination has not been found in image and text instruction experiment.

4 CONCLUSION AND FUTURE WORK

In order to improve the information support in cell production system, different types of assembly instruction and their effects on mental fatigue is evaluated in this study.

Two mental workload experiments are conducted to evaluate the mental fatigue by using several evaluation techniques.

The results show that image and text instruction have more potential to provide the assembly information than voice instruction. This relation can be observed in task performance measurement and subjective indices measurement. However, the same inclination has not been found in physiological measurement. In the CFF measurement, image and text instruction seemed to cause more mental fatigue than voice instruction.

Besides the improvement of assembly information support system, this study is conducted as a fundamental research to develop a real time vital sign monitoring system. Hence, the mental fatigue evaluation technique that is practical to an actual production system is the most important issue and challenging problem for future research.

5 ACKNOWLEDGEMENTS

This project for strategic development of advanced robotics elemental technologies is supported by NEDO. We extend our sincere thanks to NEDO. We also express our thanks to FANUC for their excellent cooperation and technical supports to the present research.

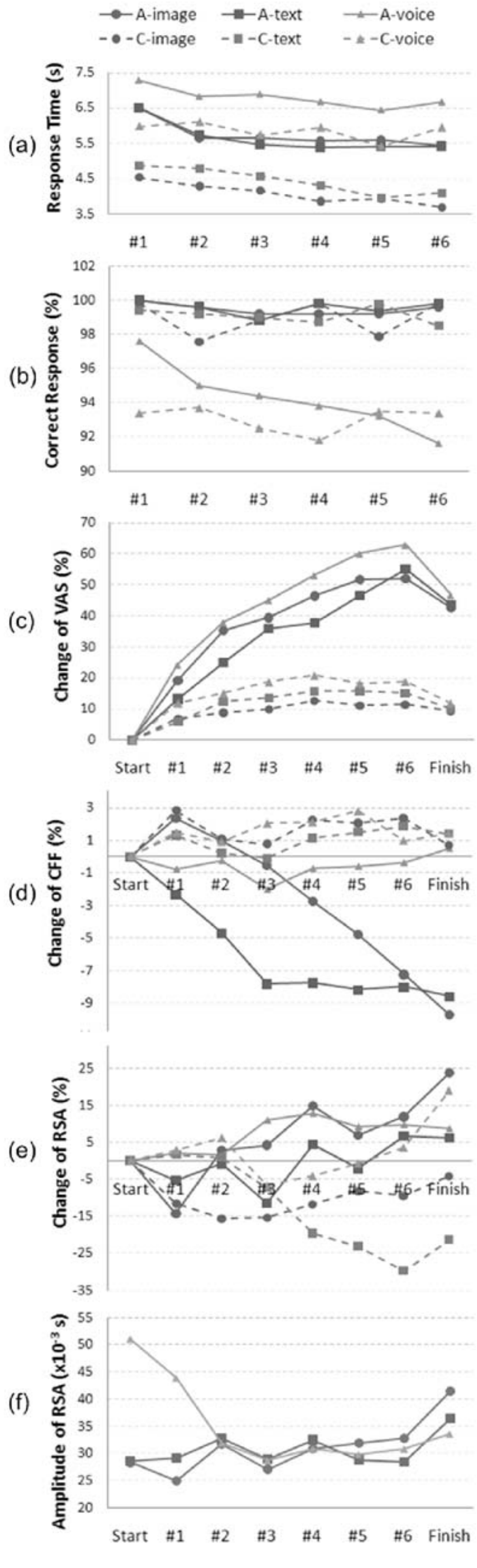


Figure 5: Experiment results.

6 REFERENCES

[1] Sugi, M., Nikaido, M., Tamura, Y., Ota, J., Arai, T., Kotani, K., Takamasu, K., Shin, S., Suzuki, H., Sato, Y., 2005, Motion Control of Self-Moving Trays for Human Supporting Production Cell "Attentive Workbench", Proceedings of ICRA, 4080-4085.

[2] Seki, S., 2003, One by One Production in the "Digital Yatai" - Practical Use of 3D-CAD Data in the Fabrication, Journal of the JSME, 106/1013:32-36.

[3] Vasilash, G., 1999, Cell for Assembly, Automotive Manufacturing & Production.

[4] Gery, G., 1991, Electronic Performance Support Systems, how and why to remake the workplace through the strategic application of technology, Tolland, MA: Gery Performance Press.

[5] Kasvi, J. J. J., Vartiainen, M., Pulkkis, A., Nieminen, M., 2000, The role of information support systems in the joint optimization of work systems, Human Factors and Ergonomics in Manufacturing, 10/2:193-221.

[6] Okogbaa, O.G., 1994, On the Investigation of the Neurophysiological Correlates of Knowledge Worker Mental Fatigue using the EEG signal, Applied Erogonomics, 25/6:355-365.

[7] Yamada, S., 2007, Effects of Long Term Mental Arithmetic on Physiological Parameters, Subjective Indices and Task Performances, The Japan UOEH Association of Health Sciences, 29/1:27-38.

[8] Kotani, K., 2007, Respiratory-phase Domain Analysis of Heart Rate Variability Can Accurately Estimate Cardiac Vagal Activity during a Mental Arithmetic Task, Methods of Information in Medicine, 3:376-385.

[9] Mcnair, D.M., Lorr, M., Droppleman, L.F., 1992, Manual for the Profiles of Mood States, Educational and Industrial Testing Service.

[10] Aaron, R.S., 2005, Visual Experience can substantially alter Critical Flicker Fusion Thresholds, Human Psychopharmacology Clinical and Experimental, 20:55-60.

[11] Murata, K., Araki, S., 1996, Accumulation of VDT Work-Related Visual Fatigue Assessed by Visual Evoked Potential, Near Point Distance and Critical Flicker Fusion, Industrial Health, 34:61-69.

Assembly Work Settings Enabling Proactivity – Information Requirements

Jessica Bruch[1, 2], Johan Karltun[1] and Kerstin Dencker[3, 4]

[1] Department of Industrial Engineering and Management, School of Engineering, Jönköping University, Sweden

[2] Department of Product and Production Development, Chalmers University of Technology, Sweden

[3] Department of Production Engineering, Royal Institute of Technology, Sweden

[4] Industrial Research & Development Corp., Swerea IVF, Sweden

Abstract

Information is a critical factor to support a proactive behaviour of operators in complex work settings characterized by flexible levels of automation and need for knowledge-based decision making. In this conceptual paper the authors define proactive behaviour as the ability of operators to control a situation by taking action in advance. Information requirements that enable proactivity and different control behaviour are identified. Moreover, several demands on the information support system are outlined. Further, the paper presents some implications for management as a result of the new work role of the operator regarding decision making, planning, and control.

Keywords:

Uncertainty; Proactive behaviour; Information support

1 INTRODUCTION

Turbulent markets and powerful customers require assembly systems to deal with frequent changes in the market demand as well as to adapt quickly to new products and variants. Managing the dynamics of the market requires decreased response time in the assembly system in order to remain competitive. To be able to act quicker it is important to reduce the complexity of the control system.

Fleig and Schneider [1] claim that the implementation of automated systems has failed to integrate planning and control tasks effectively and efficiently. The implementer of automated systems did not recognize the importance to distribute planning and control tasks to human operators on the shop floor, who 'experience' the production process at first hand and influence it by their actions [1]. Moreover, since real preplanning is only possible in a few exceptional cases [1] it is important to take care of appropriate human participation during planning and assembling. If a company provides a satisfying education and facilitates life-long learning for their employees, they will be the most flexible resource of the company [2].

In working situations characterized by uncertainty, need for knowledge-based behaviour and work role aspects that cannot be formalized, proactive behaviour can contribute to competitive advantages (e.g. [3], [4], [5], [6], [7]). Thus, assembly work settings enabling proactivity with a focus on active participation of the human operators may be favourable to achieve flexibility, reconfigurability, changeability and evolvability of the assembly system. The combination of a need for knowledge-based behaviour, information and automation characterize proactive assembly systems. Hence, levels of automation, information and operators' competence were identified as three main drivers in a proactive assembly system in the ProAct project within which this paper was written [8]. ProAct is a collaborative effort involving six Swedish manufacturing businesses and three Swedish technical universities.

In this paper we define proactive behaviour as the ability of operators to control a situation by taking action and effectuating changes of the work situation in advance in order to create a favourable outcome. Necessary planning and decisions are made by the operators and they have the authority and autonomy to perform needed actions.

The system behaviour of a proactive work setting should thus have the ability to act before a situation becomes a source of confrontation or crisis. Systems with proactivity can to some extent foresee and adapt to dynamic changing environments in real time. In doing so, sustainable proactivity necessitates two different time perspectives; one concerning acting proactively during operation and the other concerning maintaining the proactive ability over time.

Such an assembly work setting needs improved information exchange and more acting operators who are able to make decisions previously associated with middle management level. This implies that proactive assembly work settings require operators that are able to exercise their own judgement, take own initiatives and anticipate future problems in order to avoid exceptional situations in production or work. It is not any longer sufficient for the operators to perform a narrow assembly task satisfyingly. Work instruction sheets cannot represent all possible situations operators will face. Furthermore, the actual activity performed by the operator in real work situations will vary from the prescribed task according to activity analysis theory [9].

Therefore, operators need to have a certain degree of freedom in a proactive assembly work setting to contribute to effectiveness in uncertain situations. Proactive behaviour of the operators will support both short and long range development of proactive assembly systems. As Crant, describes, "Proactive behaviour can be a high-leverage concept rather than just another management fad, and can result in increased organizational effectiveness" [5, p. 435]

The benefit of proactive work behaviour of the operators may be for example that the operators will develop knowledge about a future need for change in assembly and they will identify an alternative route of action. Operators who will obtain information about upcoming production demands will have the time needed to collect more information, think about the best alternative and choose the right method to assemble.

The paper is conceptual, and aims at clarifying the role of information in a proactive assembly work setting. The literature study conducted is the basis for further research at the different case companies in order to empirically establish and validate the attributes of a proactive assembly work setting.

In this paper we identify information requirements that enable the proactive behaviour of the operators. We also present different demands on the information support system. Further, the paper outlines some implications for management as results of the new work role of the proactive operators regarding decision making, planning and control.

2 THEORETICAL BACKGROUND

2.1 Information requirements

The term "information" has been defined in many ways. In this paper information is defined as "the collection of data, which, when presented in a particular manner at an appropriate time, improves the knowledge of the person receiving it in such a way that he/she is better able to undertake a particular activity or make a particular decision" [10, p.4].

If workers on the shop floor are responsible of tasks including planning, control and execution of assembly they need adequate information. Each activity needs information, or collection of data, requirements that must be fulfilled to be able to complete the activity [11]. These information requirements as well as the value of information are also affected by contextual (when is it used and in which situation) and personal (who uses it) characteristics [12], [13]. If information requirements may be either informative or irrelevant to a user depends on both the performed activity as well as the individual cognitive processes [10]. Thus, the perception of "right information" is of importance with respect to the performance of the operators.

The information presented to operators should inform them about a reliable up to date picture of the situation but also present future production demands. In line with Byström [12] and Fjällström [13] information is understood in the role of an abstract tool that enables, or is initiated to enable the fulfilment of activities. Thus, information is viewed from the operators' perspective.

Information that supports event handling of the operators has to be pragmatic [13], [14]. Pragmatic information, which consists of two complementary dimensions: novelty (Erstmaligkeit), and confirmation (Bestätigung) [14], adds knowledge to the operators. Pragmatic information will support the operators action since it will make a difference to what the operators already know (by novelty) but at the same time also include pre-knowledge of the matter (confirmation) [14]. Neither 100% novelty nor 100% confirmation will make any different to the operators knowledge. 100% novelty does not contribute to understanding since the operators are not able to relate the information to any meaning and 100% confirmation does not comprise any new information at all for operators. Further, for information to effectively fulfil the users needs there are six qualitative criteria that are required if information should be useful: relevance, timeliness, accuracy, accessibility, comprehensiveness and format [15].

However, there exists differences between less experienced and experienced operators (e.g. [16], [17]). Thunberg [17] claims that the operators' level of expertise controls the information processed and the experienced workload. Both the type and the amount of information perceived differ between less experienced and experienced operators [17]. The skill and expertise of the operator affect the human behaviour, which can be categorised in Rasmussen's three levels of performance: skill, rule-, and knowledge-based (SRK) behaviour [16]. Each level of the SRK taxonomy defines different ways of representing the constraints in the environment and thus different level of cognitive control of the operator, see Figure 1. The skill-based behaviour (SBB) takes place without conscious attention or control and consists of smooth, automated, and highly integrated patterns. Rule-based behaviour (RBB) is characterised by pattern matching with stored rules derived from previous successful experiences. At the RBB level people are aware of their cognitive activities, and hence, can express the applied rules. Finally, the knowledge-based behaviour (KBB) is required in unfamiliar situations and demands a conscious, focal attention of the operator. In these situations previous experience is no longer valid and a solution must be improvised by functional reasoning.

The differentiation of human performance in SRK based behaviour means also differences in the way in which workers perceive and interpret information from the environment. To understand the relationship between the level of human performance and the way in which information is interpreted Rasmussen [16] introduced the concepts of signals, signs, and symbols. These three different types of information represent three ways in which workers can interpret information. During SBB, information is typically interpreted as signals. Signals have a strong perceptual basis because they are continuous quantitative indicators of the time–space behaviour of the environment. Signs, on the other hand, refer to the RBB of the operator. Signs are characterized as arbitrary but familiar perceptual cues in the environment. Information has to be perceived as symbols to be meaningful for KBB. Symbols are reasonable formal structures that represent the functional properties of the environment.

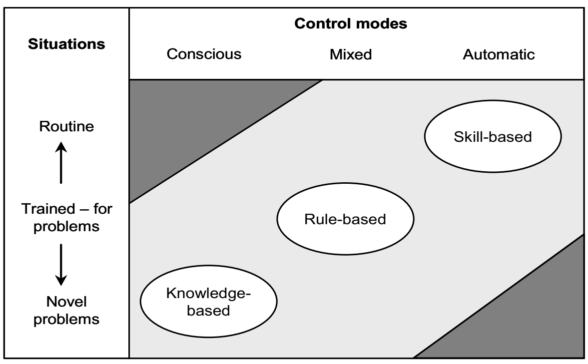


Figure 1: Control modes during different problem situations in production (From Reason and Hobbs [18]).

However, whether information is interpreted as signals, signs, or symbols relies on the context in which information is perceived and not only on the form in which the information is presented according to Rasmussen [16]. This means, that the very same object observed may be interpreted as a signal, sign, or symbol, depending on the intentions, expectations and expertise of the operator.

In order to analyse the information requirements for a range of activities, which mean different difficulties to different operators, it is useful to categorize activities according to the complexity of the activities (Figure 2). The activities that the operators in assembly systems have to handle in a proactive way are of diverse character. The activities needed to be performed by the operators are either planned or unplanned as well as how to handle the activity is either known or unknown. By supporting SBB and RBB in familiar tasks, more cognitive resources may be devoted to KBB. Unknown handling of unplanned activities is not considered to be a relevant option for defining information requirements.

To be able to control a situation the operators need to have knowledge about the cause-effect relationships that can be used for control. That is, in order to act in the right way, the operators not only need to know the result of the action, but rather the operators also have to understand the consequences of their acting. Petersen [19] enhanced the idea of von Wright [20] that it is important to make a distinction between what the operators are doing and what changes they brings about. Usually, the operators will perform control actions in order to bring about the desirable changes in the controlled system. However, the desired system change may just be a consequence of the doing and not the immediate result of action due to causal relations inherent in the controlled system. This means that the action performed by the operators will result in a system change that causes another system change (Figure 3).

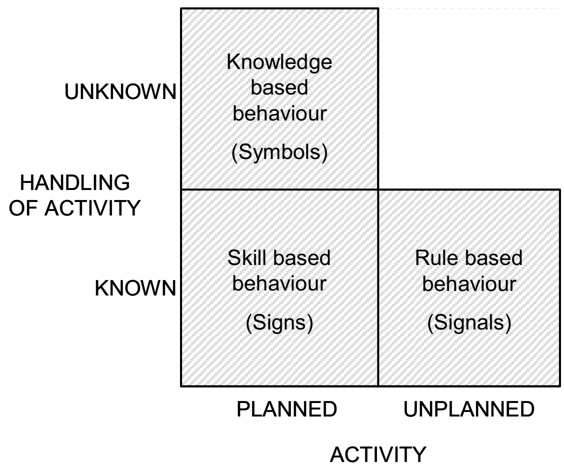


Figure 2: The relationship between complexity and performance (Modified from Fjällström [13]).

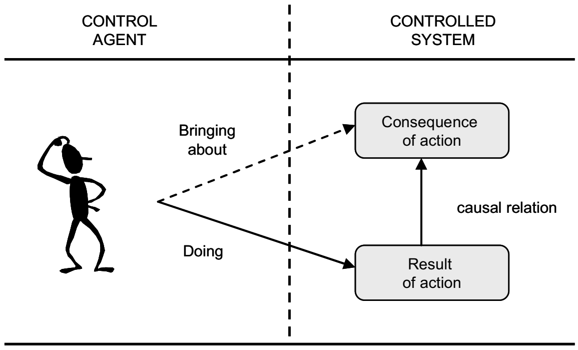


Figure 3: The doing and bringing about aspects of control actions (From Petersen [19]).

In a complex system like a proactive assembly system the potential actions of the operators are specified by constraints. The constraints can be classified as five different layers of constraints, namely: work domain, control tasks, strategies, social organization and cooperation, and worker competencies. These constraints are behaviour shaping [21] because they define the boundaries of possible actions for the operators. By identifying the constraints on operators' actions it is possible to embed them in the information system. Human operators then have the flexibility to improvise their action within the remaining space of action opportunities that are acceptable work strategies (Figure 4). According to Vicente [22] the constraint space will provide the flexibility that is required:

- To handle context-conditioned variability.
- To handle the intrinsic variability.
- To handle individual differences between the operators.

Hence, the approach based on constraints illustrated in Figure 4 helps operators to solve unanticipated situations and to follow their own choices (by choosing the preferred trajectory), while at the same time, fulfil the demands of the activity (by staying within the constraint boundaries) [22]. However, flexibility does not mean that everything goes. Therefore, it is important that the constraints remain identifiable, independent of how the operators choose to act.

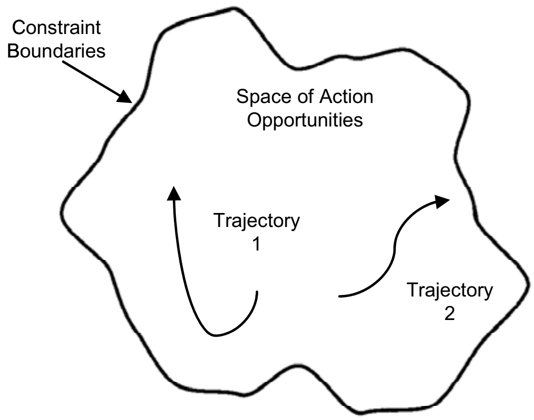


Figure 4: The available opportunities for actions by the operators are specified by behaviour-shaping constraints (Modified from Vicente [22]).

The study of Parker et al. [7] state that if a company wants to have a proactive work behaviour they should support a flexible role orientation (to define one's role broadly so that one feel responsibility for longer term goals), role breadth self efficacy (one's judgement about one's capability to perform particular tasks) and job autonomy of the operators. To facilitate the development of these characteristics, information needs to promote the flexible role orientation, to build employees' self efficacy and to encourage employees learning and development.

Moreover, the processing of information in organizations shall reduce uncertainty and equivocality [23]. The processing of rich information, which changes understanding within a time interval and thus enable debate, clarification and enactment will reduce equivocality. Additionally, organizations have to provide the right amount of information to facilitate a reduction of uncertainty. Yet, the challenge is to process sufficient information, that is not too little or too large amounts of data, in order to avoid frustration, confusion and time consuming activities (e.g. [24], [25]).

In general, the smallest amount of information required for the operators to keep the system stable is defined by Ashby's "Law of Requisite Variability" [26]. This means that to control a situation and to perform according to requirements, the variety of the control responses must as a minimum be as large as the variety of system changes that need to be compensated. Thus, the law of requisite variety has two important consequences:

1. The amount of information available determines the amount of appropriate actions that can be performed.
2. To be able to control a situation the variety of the controller must be equal or greater than the variety in the system being controlled.

Because the performance of the operators is influenced by the utilized information [27] and the information flow in an organization is seen as one of the predominant factors to higher levels performance [28] it is necessary to apply these information requirements in designing a system interface that controls the collecting, evaluating, organizing, and distributing of information. For example information should be organized in such a way that the operators have access to relevant information at the required time.

2.2 Information support system

The information support system is used to provide information and knowledge that supports different work activities [29]. The general aim is to design information support that enable human operators and technologies to work together in a more flexible and mutual control system which in turn functions reliably in complex work environments. Or to put it in the words of Hollnagel [30, p.221], the objective is to "provide the right information at the right time and in the right way".

Providing operators with a complete, accurate, and up to date picture of the situation will be the challenge as the work setting of the operators become more complex and demanding. Kasvi et al. [29] identified four parts of an information support system:

1. A source of information, supporting efficient, good quality, and safe completion of work activities.
2. Support is available on demand, in context with the task supported.
3. Information is accessed spontaneously and the order of access is controlled by the end user of information.
4. In addition to providing information, the system interactively supports the collecting, creation, and synthesis of the experienced-based knowledge of the members of the operative organization.

The actual form and functionality of an information support system depends on the work and people supported [29]. One important challenge is to develop an information system that simultaneously supports the use of existing knowledge and the creation of new knowledge as in proactive work settings. Therefore, the information support system has to include both reader-users (recipient of information) and author-users (ability to change and personalise the content of information) [31].

The SRK framework developed by Rasmussen [16] will support designers to combine information requirements for a system and aspects of human cognition. Designers can apply the SRK framework to determine how information should be presented to the operators to take advantage of their perception and psychomotor abilities [32].

It is important to have a broader functionalist perspective to the development of the support systems. The system does not only consist of the human operators and the machines, it also includes the work domain. It is important to understand that the human operators and the machines are integral parts of the work system (e.g. [31], [33], [34]) The importance of this perspective can be illustrated by Simons [35] parable of the ant. To understand the underlying rationality of the behaviour of the ant and its path along a beach it is necessary to see the path in the context of the beach. When the ant's path, on the other hand, is considered isolated from the beach it appears complex and the underlying rationality is not obvious. In consequence, if we want to learn about the ant's behaviour we need a description that recognizes the constraint arising from both the beach and the ant.

The interaction between human operators and their work has to be the focus in the design process. For the design of support systems, the above mentioned constraints, which define the boundaries of action opportunities, have to be analysed [22, 36]. The operators then should have the possibility to choose from different alternatives as well as the operators have different ways to gather information to make more informed choices. Moreover, the information system should support the operators to handle the limits of local decision making. That is, it should control that operators will only make decisions within the outer boundaries [22].

Information should furthermore help to avoid sub optimization of the proactive behaviour because operators are only parts of the system and need to cooperate. In interdependent systems, the behaviour of the individual has an impact not only on the effectiveness of that individual, but also in the effectiveness of others, including groups, teams, and the organization as a whole [3]

2.3 Implications for management

Since work in assembly systems is based on standardization, proactive behaviour implies to deviate from the standardized processes. Moreover, the opportunities of proactive behaviour of the operators may be limited due to little autonomy and control. Thus, to support proactive behaviour of the operators, organizations and individuals need to change. The traditional view of the operator as someone who is exposed to decisions by production management must be redefined towards a view of the operator as an actor and as a learning and collaborating individual. This poses problems of organizational solutions like raising the authority of the operators and the needed support from the organization. It also puts new demands on the technical solutions of the production system and the information system [37].

When operators have to make decisions on their own, it has to be considered how to handle the relation between local decision making of the operators and decision making provided by production management. The concept of proactivity always bears the risk of unexpected and unpredicted (and unwanted) outcomes because of the independent judgement and initiative of the employees [4]. As demanded, employees may go beyond given task descriptions.

Moreover, other organizational factors have to be considered when developing a proactive workforce. First operators need to adopt a more customer-strategic focused orientation since it will not be enough for shop floor employees to restrict their efforts to maximizing volume [38]. Organizations have to develop initiatives that enhance the understanding of modern principles (like increasing flexibility, continuous improvement, etc.). This will help employees to understand and create acceptance for principles that derive from broader strategic objectives.

Second, Freese and Fay [6] defines that employees need the right knowledge and skills, know when it is better to "give up", and the degrees of autonomy have to have limits. Cambbell [4] suggest four strategies of which only the first three are in line to our concept in order to avoid the so called 'initiative paradox' (employees are expected to use independent judgement and initiative, and simultaneously expected to think and act like their bosses):

1. Goal alignment requires substantial alignment between the goals and interests of the organization and the goals and interests of the individual; the alignment of interests minimizes the likelihood of undesired, unexpected results.
2. Communication of boundaries, careful communication of the kind of initiative desired, and the limits surrounding these, initiative and judgement are encouraged only in circumscribed job or work situations.
3. Information sharing, minimizing unshared expectations by providing trusted employees with the same information and frame of reference that the managers use in running the work unit.

This is in line to Bateman and Crant [39] and Vicente [22] who suggest that the organization should make the core activities clear and define the outer constraints. After that, it is important to give the operators considerable degrees of freedom for decision-making.

Furthermore, research on control behaviour suggests that human control requires perception of being in control. A high degree of autonomy will increase the controllability of the task [7]. This is in line with Karasek's demand and control model [40]. High demands will contribute to a proactive workforce only if combined with high levels of job control and social support. Thus, facilitating proactivity may require structural changes that devolve authority to the people on the shop floor – changes that can be quite difficult to achieve.

3 SUMMARY

It is widely accepted that a proactive behaviour is advantageous in today's workplaces. Yet, operators need information that enables them to anticipate and plan for expected changes. This paper sought to address information requirements in a work setting enabling proactivity, by examining demands on information, the information support system and management.

Our findings are that there is a need to develop an appropriate information support system that considers both the operators' cognitive abilities and demands on information. Since in the end, the ability to present meaning, rather than information is the most important factor to the operators. Moreover, we have shown that operators proactivity will not only put demands on information and the information support system but also on how to handle a much more delegated decision making and still avoiding sub optimization.

4 ACKNOWLEDGEMENTS

The research is financially supported by the Swedish Agency for Innovation Systems (VINNOVA) as it is a part of the ProAct project (http://proact.iip.kth.se/). Moreover, the authors would like to thank the companies participating in the ProAct project as well as ProAct colleagues for valuable comments.

5 REFERENCES

[1] Fleig, J., Schneider, R., 1998, The Link to the 'Real World': Information, Knowledge and Experience for Decision Making, in Shop Floor Control–A Systems Perspective, Edited by Scherer, E., Springer Verlag, Zürich, 221-43.

[2] Bley, H., Reinhart, G., Seliger, G., Bernardi, M., Korne, T., 2004, Appropriate Human Involvement in Assembly and Disassembly, CIRP Annals-Manufacturing Technology, 53/2:487-509.

[3] Griffin, M.A., Neal, A., Parker, S.K., 2007, A New Model of Role Performance: Positive Behavior in Uncertain and Interdependent Contexts, Academy of Management Journal, 50/2:327-347.

[4] Campbell, D.J., 2000, The proactive employee: Managing workplace initiative, The Academy of Management Executive, 14/3:52-66.

[5] Crant, J.M., 2000, Proactive behavior in organizations, Journal of Management, 26/3:435-462.

[6] Frese, M., Fay, D., 2001, Personal Initiative: An Active Performance Concept for Work in the 21st Century, in Research in Organizational Behavior, Edited by Staw, B.M., Sutton, E.L., Elsevier Science, Amsterdam, 133-187.

[7] Parker, S.K., Williams, H.M., Turner, N., 2006, Modeling the Antecedents of Proactive Behavior at Work, Journal of Applied Psychology, 91/3:636-652.

[8] Dencker, K., Stahre, J., Grondahl, P., Martensson, L., Lundholm, T., Johansson, C., 2007, An Approach to Proactive Assembly Systems -Towards competitive assembly systems, Proceedings of IEEE International Symposium on Assembly and Manufacturing (ISAM), 294-299.

[9] Daniellou, F., 2005, The French-speaking ergonomists' approach to work activity: cross-influences of field intervention and conceptual models, Theoretical Issues in Ergonomics Science, 6/5:409-427.

[10] Galliers, R., 1987, Information analysis: selected readings, Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA.

[11] Wersig, G., 1973, Informationssoziologie: Hinweise zu einem informationswissenschaftlichen Teilbereich, Athenäum Fischer, Frankfurt a.M.

[12] Byström, K., 1999, Task Complexity, Information Types and Information Sources, University of Tampere, Tampere, (Academic Dissertation).

[13] Fjällström, S., 2007, The Role of Information in Production Ramp-up Situation, Chalmers University of Technology, Göteborg, (Ph.D. Thesis).

[14] Weizsäcker, E.v., 1974, Erstmaligkeit und Bestätigung als Komponenten der Pragmatischen Information, in Offene Systeme I Beiträge zur Zeitstruktur, Entropie und Evolution, Edited by Weizsäcker, E.v., Ernst Klett Verlag, Stuttgart, 82-113.

[15] Kehoe, D.F., Little, D., Lyons, A.C., 1992, Measuring a company IQ, Proceedings of Third International Conference on Factory 2000 - Competitive Performance Through Advanced Technology, London, 173-178.

[16] Rasmussen, J., 1983, Skills, Rules, Knowledge, Signals, Signs and Symbols and other Distinctions Human Performance Models, IEEE Transactions on Man, Systems and Cybernetics, SMC-13/3:257-266.

[17] Thunberg, A., 2006, A cognitive approach to the design of alarm systems for nuclear power plant control rooms, Chalmers University of Technology, Göteborg, (Licentiate Thesis).

[18] Reason, J.T., Hobbs, A., 2003, Managing Maintenance Error: A Practical Guide, Ashgate Publishing, Hampshire.

[19] Petersen, J., 2004, Control situations in supervisory control, Cognition, Technology & Work, 6/4:266-274.

[20] von Wright, G.H., 1971, Explanation and understanding, Cornell University Press, Ithaca, NY.

[21] Rasmussen, J., Pejtersen, A.M., Goodstein, L.P., 1994, Cognitive systems engineering, John Wiley & Sons, Inc., New York.

[22] Vicente, K.J., 1999, Cognitive Work Analysis: Toward Safe, Productive, and Healthy Computer-Based Work, Lawrence Erlbaum Associates, Mahwah.

[23] Daft, R.L., Lengel, R.H., 1986, Organizational Information Requirements, Media Richness and Structural Design, Management Science, 32/5:554-571.

[24] Almgren, H., 1999, Pilot production and manufacturing start-up in the automotive industry: principles for improved performance, Chalmers University of Technology, Göteborg, (Ph.D. Thesis).

[25] Berglund, M., Skog, E., Öjmertz, B., 2001, Driftsättning av monteringsystem - Erfarenheter och synpunkter från svenska industriföretag, IVF-rapport 01009, IVF, Mölndal, Sweden (Swedish).

[26] Ashby, W.R., 1956, An Introduction to Cybernetics, Chapman & Hall Ltd, London.

[27] Drury, C.G., Prabhu, P., 1996, Information requirements of aircraft inspection: framework and analysis, International Journal of Human-Computer Studies, 45/6:679-695.

[28] Goldman, S.L., Nagel, R.N., Preiss, K., 1995, Agile Competitors and Virtual Organizations - Strategies for Enriching the Customers, Van Nostrand Reinhold, New York.

[29] Kasvi, J.J.J., Vartiainen, M., Pulkkis, A., Nieminen, M., 2000, The Role of Information Support systems in the Joint Optimization of Work Systems, Human Factors and Ergonomics in Manufacturing, 10/2:193-221.

[30] Hollnagel, E., 1988, Information and reasoning in intelligent decision support systems, in Cognitive engineering in complex dynamic worlds, Edited by Hollnagel, E., Mancini, G.,Woods, D.D., Academic Press, London.

[31] Kasvi, J.J.J., 2003, Knowledge Support in Learning Operative Organisations, Helsinki University of Technology, Helsinki.

[32] Vicente, K., 1999, Ecological interface design: Supporting operator adaptation, continuous learning, and distributed, collaborative work, Proceedings of Human Centered Processes Conference, Brest, 93-97.

[33] Flach, J.M., Dominguez, C.O., 1995, Use-centered design, Ergonomics in Design, 3/3:19-24.

[34] Rasmussen, J., Vicente, K.J., 1989, Coping with human errors through system design: Implications for ecological interface design, International Journal of Man-Machine Studies, 31/5:517-534.

[35] Simon, H.A., 1969, The Science of the Artificial, MIT Press, Cambridge, MA.

[36] Rasmussen, J., 1991, Modelling Distributed Decision Making, in Distributed Decision Making: Cognitive Models for Cooperative Work, Edited by Rasmussen, J., Brehmer, B.,Leplat, J., John Wiley & Sons Ltd, New York, 111-142.

[37] Westlander, G., 1999, Fokus på människan i forskning om verksamhetsutveckling, in Forskningsperspektiven: Forskares syn på samspelet mellan Människa-Teknik-Organisation, Edited by Ahlin, J., Nutek, Stockholm, 20-33.

[38] Parker, S.K., Wall, T.D., Jackson, P.R., 1997, "That's not my job": Developing flexible employee work orientations, Academy of Management Journal, 40/4:899-929.

[39] Bateman, T.S., Crant, J.M., 1999, Proactive behavior: Meaning, impact, recommendations, Business Horizons, 42/3:63-70.

[40] Karasek Jr., R.A., 1979, Job Demands, Job Decision Latitude, and Mental Strain: Implications for Job Redesign, Administrative Science Quarterly, 24/2:285-308.

Assembly Information System for Operational Support in Cell Production

Jeffrey Too Chuan Tan[1], Feng Duan[1], Ye Zhang[1], Kei Watanabe[1],
Nuttapol Pongthanya[2], Masao Sugi[1], Hiroshi Yokoi[1] and Tamio Arai[1]
[1] Department of Precision Engineering, School of Engineering, The University of Tokyo, Japan
[2] Department of Systems Innovation, Faculty of Engineering, The University of Tokyo, Japan

Abstract
In cell production, the output depends heavily on the performance of the worker. Assembly information support is useful to support the worker. The objective of this study is to develop a framework to organize assembly information to support production operation. It is illustrated by an application of a production operation in laboratory simulation. Assembly motion of skilled worker is extracted and the assembly process is decomposed into operational units, which linked with support information to set up the information database. New assembly process with information support can then be generated and implemented with multimodal system to assist production operation.

Keywords:
Cell Production; Assembly Process Planning; Task Analysis; Multimodal User Interface

1 INTRODUCTION

1.1 Towards Better Cell Production

The changing manufacturing requirements from conventional mass production to diversify product design with flexible production quantity, has greatly excited industrial interest in cell production. Also known as cellular manufacturing, this system consists of human worker as the center of the working cell to assemble the product [1]. In a production dell, a highly and multiple skilled worker is required because the manual assembly jobs are normally impractical for automation system. Moreover, the skilled worker is costly and requires frequent job training as the production jobs vary. Therefore, many studies [2] [3] have been conducted to improve working skill in production systems. On the other hands, efficient production cell system design is important to ensure good collaboration between human worker and machines. Physical support [1] [4] such as parts providing can greatly shorten assembly completion time. Besides tangible support, information support is another key research in industrial manufacturing.

1.2 Information Support in Production

In 1991, Gloria Gery had introduced Electronics Performance Support Systems (EPSS) as an electronic environment that provides various types of information to employee to improve job performance [5]. Finnish research team had developed Interactive Task Support Systems (ITSS) to support operative tasks (assembly work tasks) and Information Support System (ISS) that supports different work activities [6] [7] [8]. The importance of information support in production is apparent. However, most of the research works focus more on the overall system functionality rather than the content itself – Information. In this study, the main concern is to support human worker for assembly operation in cell production. The suitability form of information to the corresponding operation task and the effectiveness of the information transfer to the human worker are the core findings of this work. Hence, the objective of this work is to develop a framework to organize assembly information to support production operation. The objective is incorporated with two distinctive approaches: job analysis by Hierarchical Task Analysis (HTA) method and multimodal based operational support system. In this paper, the framework is illustrated by an application in an actual cell production operation in laboratory simulation. The selected cable assembly task and working cell setup are explained in Section 2 together with introduction to the principal system. Section 3 discusses the analysis of cable assembly operation using HTA approach. Section 4 illustrates the assembly information organization in term of development and database structure. The assembly flow design and multimodal assembly support are explained in Section 5. Finally, Section 6 concludes this work with suggestions for further improvement.

2 SYSTEM BACKGROUND

2.1 Assembly Task

An existing cell production assembly operation is selected in this study to ensure the practicability of the developed framework. The operation job is to assemble parts and cables into the assembled product in Figure 1. However, the assembly tasks in this study are limited to cables assembly only (excluding cable binding, terminals and plate assembly).

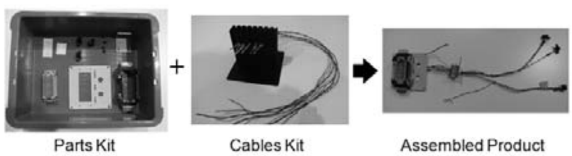


Figure 1: Assembly components and assembled product.

The selected working cell setup (Figure 2) consists of a workbench and a set of twin six degree-of-freedom manipulators to supply assembly parts to the human worker. The manipulators are located in between the workbench and parts tray rack. Parts are picked up from the blue trays on the rack by bin picking method and placed into the kit tray in front of the workbench.

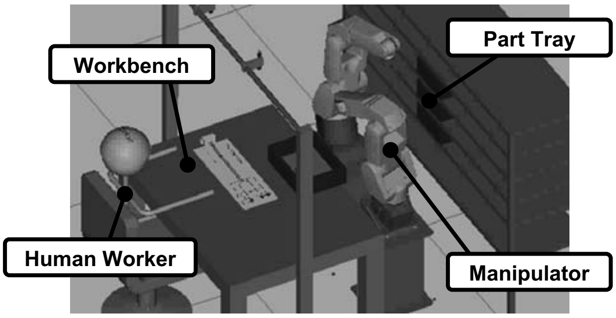


Figure 2: Working cell setup.

2.2 Principle Support System

The framework developed in this work belongs to a larger support system that includes expert worker motion measurement for skill study, Multimodal Assembly Support System (MASS) and also human vital sign monitoring system. Figure 3 shows the entire principle support system flow. The entire system serves for two main purposes: operation information support and operation teaching support. These two supports are key enhancement studies for human worker performance in cell production system. In this work, operation information support is illustrated.

In Figure 3, the elements in red box are the framework components of this work. The process that develops information support in assembly operation involves four main stages: Operational motion and task analysis, information database development, assembly flow design and multimodal assembly support. As shown in Figure 3, the first stage involves two main analyses, expert worker operation motion study and operation task analysis. In this stage, detailed study is conducted on the operation motion by expert worker and also the assembly tasks of given job for skill extraction and task modeling. Operation motion measurement and study are done in another work in the project [9]. However, this input is not included into design consideration in this work but only operation task analysis is concerned. Details on task analysis are discussed in the next section in this paper. The second stage involves assembly information processing and development. Information database is also developed in a collaborative approach to enable team development. The information is then used to construct into operation support in assembly flow design. In the final stage, the operation support is implemented in production with MASS.

3 TASK ANALYSIS

3.1 Task Analysis and Hierarchical Diagram

Task Analysis (TA) is a study to describe activities that must be carried out to achieve a specific goal [10]. It provides a scientific management to model tasks especially in user's task domain which is an important factor in system design [11]. In this work, the first stage of work focuses on the study of the given assembly operation in cell production. HTA which is widely used in various applications for system ergonomic and performance improvements [12], is an excellent task analysis technique for task modeling in this work. The extendibility sub-goal hierarchical structure and the ergonomics approach well addressed the design requirements of this system.

3.2 Cable Assembly Operation

The main task of the given assembly job in this study is the assembly of five cables on a specific marking board. The assembly operation is being modeled in hierarchical diagram as shown in Figure 4. In the diagram, the main goal of the job is 'Assemble cables on marking board' (Super-ordinate goal 0). Under the main goal, there are four sub-goals: 'Secure cable contact on connector', 'Temporary fix cable end', 'Set connector on marking board' and 'Form cable on marking board' (Sub-goal 1, 2, 3 and 4). Plan 0 addresses the working flow of these four sub-goals. Sub-goal 1, 3 and 4 are further

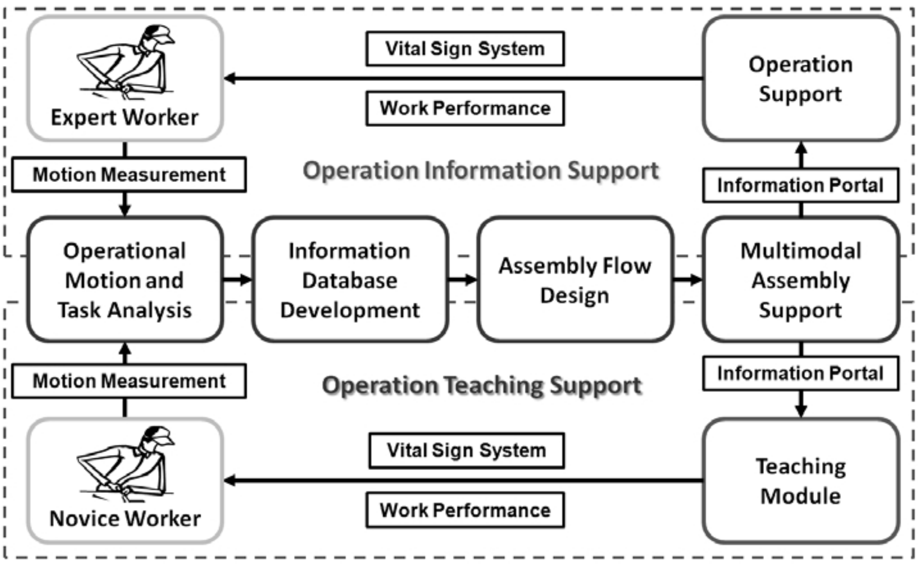


Figure 3: Principle support system flow.

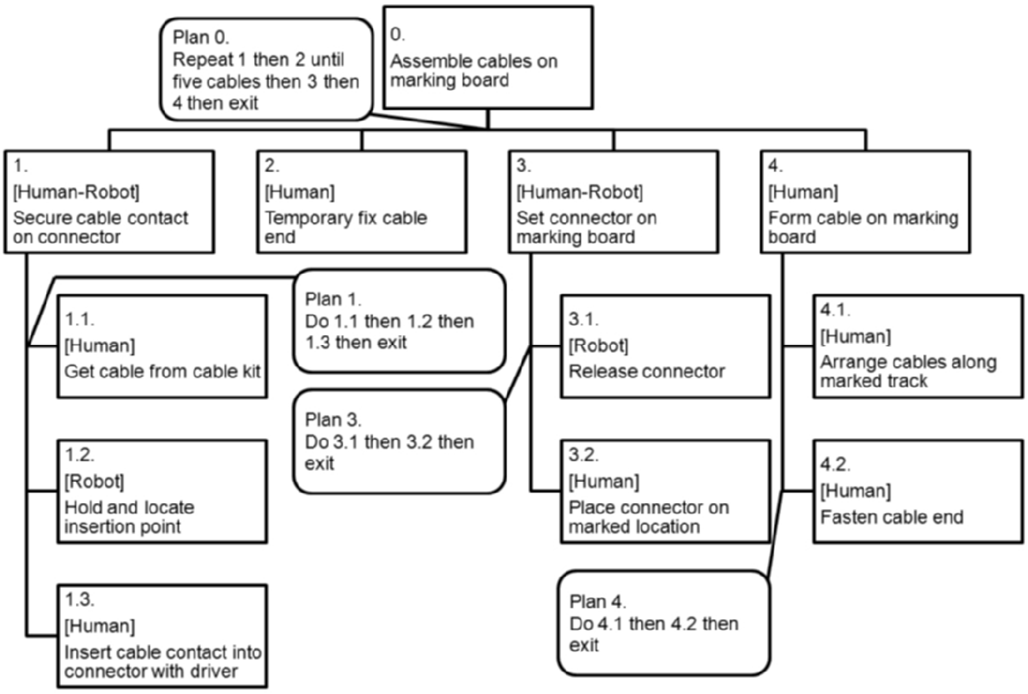


Figure 4: Hierarchical diagram for the cable assembly operation.

branched out to own sub-divisions with respective plans for operation flow. The hierarchical analysis stops at the third layer as the sub-goals on this layer are considered 'fit-for-purpose' [12]. One special parameter has been inserted in this diagram to indicate the working role of tasks between human worker and robot manipulators to address the man-machine collaboration in this work.

4 INFORMATION SYSTEM

4.1 Assembly Information

Based on HTA study, the entire assembly job has been properly modeled into discrete task segments. The next step is to develop the task segment 'descriptions' into assembly information, which will be supplied later to the human worker for assembly operation. These 'descriptions' can be developed in various media formats which traditionally are appeared in text format as in conventional operation manual. However, the studies of human cognitive science and workload theories show that different media formats can influence human cognitive and working performance. Based on Multiple Resource Theory (MRT) by Wickens [13], the usage of multiple media formats can improve the effectiveness of the information supply to the human worker in assembly work. Hence, in this work, the assembly information is developed in text, picture, audio and video formats extracted from expert worker operation recording, operation manual, training materials, operation documentation and other resources. The same reason also leads to the application of multimodal based assembly support system in the next section.

4.2 Information Database

A systematic database structure is essential to handle the vast amount of multi formats assembly information. Raw assembly data from various sources are being processed and edited into the desired forms and saved into digital file system. All the files are stored in online document center on Microsoft Office SharePoint Server platform. To further enhance the information organization, a Task-Information Mapping tool is being developed on Eclipse platform [14] (Figure 5). This tool provides a development environment to map task with all the related information files. This mapping offers great assistance to the multimodal assembly presentation development in the next section.

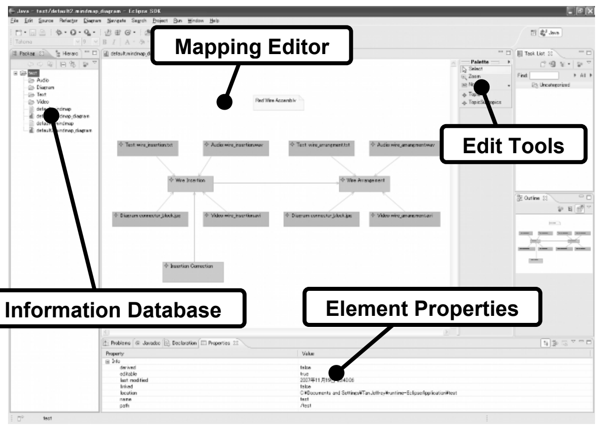


Figure 5: Task-Information Mapping tool.

4.3 Collaborative Development

The information database is managed by Microsoft Office SharePoint Server system. This system enables collaborative development through its online portal system. In the document center, team members can manage the files remotely. Apart from technical development, the information also can be accessible by other parties to serve as an online production information portal.

5 ASSEMBLY SUPPORT

5.1 Multimodal Assembly Support System (MASS)

The final step in this work is the assembly flow design and multimodal assembly presentation development before input into operation. These developments are conducted in Microsoft Office PowerPoint environment. PowerPoint provides a very flexible and user-friendly development environment. Based on the HTA study, the assembly flow is edited into PowerPoint slide presentation and the assembly information are inserted in multimodal approach to improve the information supply effectiveness [15]. Figure 6 shows a sample of multimodal assembly presentation for information support in assembly operation.

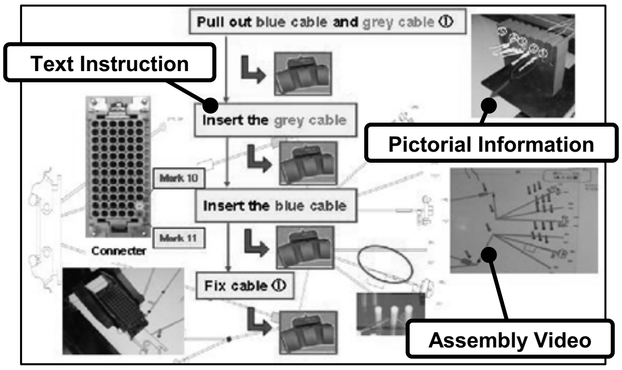


Figure 6: Multimodal Assembly Presentation.

5.2 Laboratory Simulation

Based on the proposed framework, the assembly information support is completed for the given cable assembly operation and implemented in the laboratory simulation. The manipulators are being programmed based on the assembly flow and the man-machine system is collaborated well in the laboratory simulation. The assembly operation is being able to carry out without any major mistake by non skilled worker with better working performance in the production cell with the assistance of the information support.

6 CONCLUSIONS AND FUTURE WORK

In this work, a framework that organizes assembly information to support production operation is developed. Key points of the development:

- The framework is successfully implemented in a cable assembly laboratory simulation.
- HTA method has effectively modeled man-machine collaborative assembly job.
- Information database, a task-information mapping tool and a collaborative development environment are built to manage the assembly information.
- Multimodal based operational support system has enhanced the working performance during assembly operation.

The integration of task analysis technique and multimodal interface to create operational support in term of assembly information is feasible with great potential. The study should continue to investigate the effectiveness of such support system in term of working performance and find out the design parameters. The accuracy of task analysis in assembly operation and the relationship between task and information are also valuable research work to further enhance working performance in cell production.

7 ACKNOWLEDGMENTS

This work is part of 'Strategic Development of Advanced Robotics Elemental Technologies' project supported by NEDO. Highest gratitude is expressed to NEDO and research partner, FANUC for the excellent cooperation and technical support.

8 REFERENCES

[1] Vasilash, G., 1999, Cell for Assembly, Automotive Manufacturing & Production.

[2] Yamada, Y., Konosu, H., Morizono, T., Umetani, Y., 1999, Proposal of Skill-Assist: A System of Assisting Human Workers by Reflecting Their Skills in Positioning Tasks, IEEE SMC, 4:11-16.

[3] Pathomaree, N., Charoenseang, S., 2005, Augmented Reality for Skill Transfer in Assembly Task, ROMAN, 500-504.

[4] Sugi, M., Nikaido, M., Tamura, Y., Ota, J., Arai, T., Kotani, K., Takamasu, K., Shin, S., Suzuki, H., Sato, Y., 2005, Motion Control of Self-Moving Trays for Human Supporting Production Cell "Attentive Workbench", Proceedings of ICRA, 4080-4085.

[5] Gery, G., 1991, Electronic Performance Support Systems, How and Why to Remake the Workplace through the Strategic Application of Technology, Gery Performance Press, Tolland, MA.

[6] Kasvi, J. J. J., Vartiainen, M., Pulkkis, A., Nieminen, M., 2000, The Role of Information Support Systems in the Joint Optimization of Work Systems, Human Factors and Ergonomics in Manufacturing, 10/2:193-221.

[7] Kasvi, J. J. J., Vartiainen, M., Pulkkis, A., Nieminen, M., 1996, Information Tools for the Shop-floor, AI & Soc, 10/1: 26-38.

[8] Repokari, L., Nieminen, M., Hailikari, M., Kasvi, J., Vartiainen, M., Pulkkis, A., Kari, I., 2000, Different Modalities in Assembly Support System User Interface, ICME, 2:921-924.

[9] Duan, F., Tan, J. T. C., Zhang, Y., Watanabe, K., Pongthanya, N., Sugi, M., Yokoi, H., Arai, T., 2007, Analyze Assembly Skills using a Motion Simulator, Proceedings of IEEE ROBIO, 1428-1433.

[10] Hollnagel, E., 2006, Chapter 14 Task Analysis: Why, What, and How, Handbook of Human Factors and Ergonomics, John Wiley & Sons, Inc.

[11] Johnson, H., Johnson, P., 1991, Task Knowledge Structures: Psychological Basis and Integration into System Design, Acta Psychologica, 78:3-26.

[12] Stanton N. A., 2006, Hierarchical Task Analysis: Developments, Applications, and Extensions, Applied Ergonomics, 37/1:55-79.

[13] Wickens, C. D., 2002, Multiple Resources and Performance Prediction, Theoretical Issues in Ergonomics Science, 3/2:159-177.

[14] Eclipse – an open development platform, 2007, http://www.eclipse.org/.

[15] Sarter, N. B., 2006, Multimodal Information Presentation: Design Guidance and Research Challenges, International Journal of Industrial Ergonomics, 36/5:439-445.

Multimedia based Assembly Supporting System for Cell Production

Feng Duan[1], Masahiro Morioka[3], Jeffrey Too Chuan Tan[1], Ye Zhang[1], Kei Watanabe[1], Nuttapol Pongthanya[2], Masao Sugi[1], Hiroshi Yokoi[1], Ryou Nihei[3], Shinsuke Sakakibara[3], Tamio Arai[1]

[1] Department of Precision Engineering, School of Engineering, The University of Tokyo, Japan

[2] Department of Systems Innovation, Faculty of Engineering, The University of Tokyo, Japan

[3] 1st Robot Laboratory, FANUC, Japan

Abstract

Multiple skilled human operators enable the cell production system to meet the requirement of the diversified products and production quantity flexibly. How to improve the assembly performance of the operators and reduce their assembly burden are two important factors. To solve these problems, a multimedia based assembly supporting system (MASTER) is proposed, which aims to support assembly operators from both information aspect and physical aspect. The results show that using this system, the assembly performance of the operators is improved. With the information and physical supports from this system, the assembly complexity and burden of the assembly operators are reduced.

Keywords:

Cell Production System; Multimedia; Assembly Support

1 INTRODUCTION

Automated manufacturing line system is traditional designed to produce single specific products without much flexibility. To follow the changing consumers' tastes, a manufacturing system with the name "Cell Production System" is proposed, in which a multiple skilled human operator manually assembles each product from start to finish [1], [2]. Multiple skilled human operators enable the cell production system to meet the requirements of diversified products and production quantity flexibly. However, due to the negative or zero growth of population in Japan, it is difficult to maintain the cell production system with enough multiple skilled human operators in the near future. How to improve the assembly performance of the operators and how to reduce their assembly burden are two important factors, which limits the further development of cell production system.

To solve these problems, Seki [1] invented a production cell called "Digital Yatai" which monitors assembly progress and presents the information about the next assembly process. Using a semi-transparent head mount display, Reinhart [2] developed augmented reality (AR) system to supply information to the human operator. These researches support the human operator from information aspect.

Vasilash [3] made another production cell called "Raven Super Cell", which consists of a horseshoe-shaped fixed table and a round-shaped movable worktable. To reduce the operator's physical burden and improve the assembly precision, Hayakawa [4] employed a manipulator to grasp assembly parts during the assembling process. These improved the assembly cell in physical supporting aspect.

Sugi [5] aimed to support human assembly operators from both information side and physical side, and designed an attentive workbench (AWB) system. In this system, a projector is employed to provide assembly information to the operator; a camera is used to detect the operator's finger pointing direction; and several self-moving trays are utilized to bring parts to the operator. Although AWB first achieved the aim to support operator from both information aspect and physical aspect, the direct supporting parts are just a projector and several self-moving trays, which are very general but can not meet the real manufacturing requirements.

In this paper, a multimedia based assembly supporting system (MASTER) is induced, which aims to satisfy the real manufacture's requirement and support assembly operators from both information aspect and physical aspect.

The rest of this paper is organized as follows. Section 2 briefly describes the whole system and its simulator. Section 3 introduces two manipulators which are used to provide assembly parts to the human operator. In section 4, the multimedia based assembly workbench and corresponding devices are discussed. Operator's fatigue monitor system and safety strategy are presented in section 5 and section 6 respectively. Finally, the conclusions and future work are given in section 7.

2 MASTER SYSTEM

2.1 Structure of the whole system

The whole multimedia based assembly supporting system is divided into physical supporting part and assembly information supporting part, shown in Figure 1. Physical supporting part is aimed to support human operators from physical aspect, and it is composited of two six degree-of-freedom manipulators: one is used to bring the assembly parts from the tray shelf to the workbench; the other one is utilized to grasp the assembly part to avoid wobbling during the assembly process. Assembly information supporting part is designed to assist human operators in information aspect. Therefore, a LCD TV, a speaker, a laser pointer etc. are employed to provide assembly information to guide human operator. Meanwhile, vital sensors are used to monitor the operator's physical conditions during the assembly process, and a light curtain is adopted to protect the operator from hurting by manipulators.

2.2 Simulator of the whole system

To reduce the design period, in this research, a simulator of the whole system is developed based on ROBOGUIDE [6] and OpenGL [7], shown in Figure 2. This simulator can not only reproduce the real manipulators' motions, but also can predict the collisions in the work space. Therefore, based on the simulation results, the real system can be constructed conveniently.

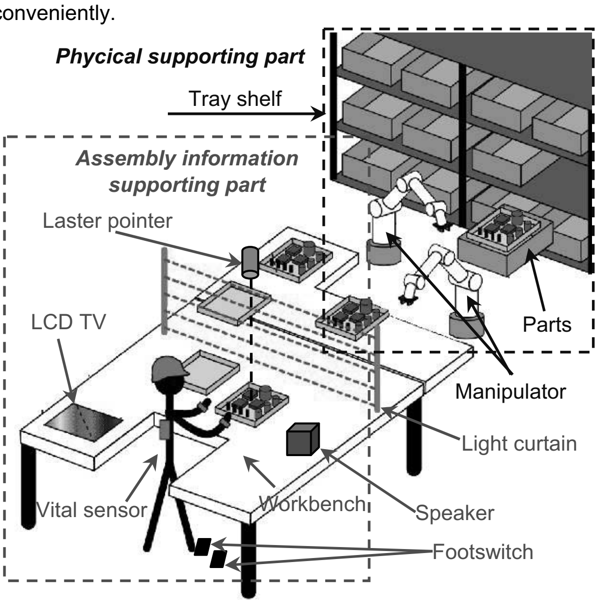


Figure 1: Structure of the whole system.

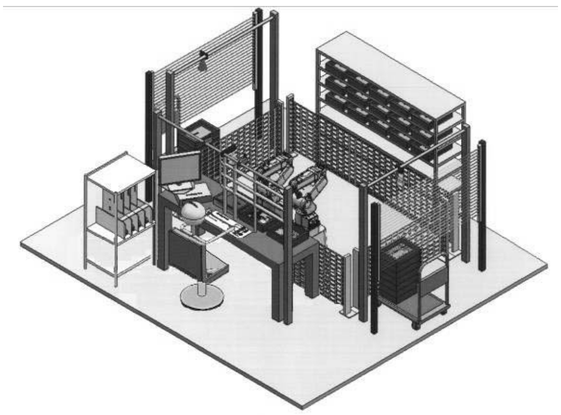


Figure 2: Simulator of the whole system.

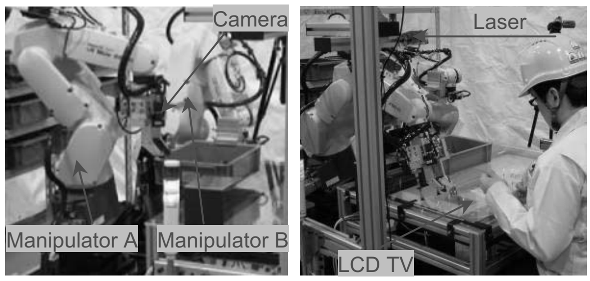
3 PHYSICAL SUPPORTING PART

In order to increase system's physical supporting effect, two six degree-of-freedom manipulators are used to bring assembly parts to the operator, shown in Figure 3 (a). A LED based camera is fixed at each manipulator to recognize the assembly part. After the manipulator recognizing the assembly part, based on the predesigned assembly part database, the manipulator can pick up and supply the assembly part to the human operator according to his or her requirement. In Figure 3 (a), the manipulator B grasps and draws the trays to assist the manipulator A; depending on the grasp clamp or the suck device fixed at the end of the manipulator A, it can either pick up or suck up the assembly parts according to their characteristics; and then supply them to the human operator. Meanwhile, to increase the assembly precision and reduce the human operator's assembly burden, the manipulator A can grasp the required assembly part to avoid its wobbling during the assembly process, and the human operator executes the assembly task based on the manipulator's assistance, shown in Figure 3 (b). Human operator can control the movement of these two manipulators through footswitch according to the assembly requirements, and can stop the manipulators by emergent button when accident happens. These strategies enable the manipulators to support operators in physical aspect effectively and safely.

4 ASSEMBLY INFORMATION SUPPORTING PART

4.1 Multimedia based assembly workbench

In Sugi's opinion, providing assembly information to the human operators based on their intentions and conditions is an effective way to reduce the assembly complication and improve the productivity [5]. Therefore, in AWB system [5], a projector is employed to provide assembly information to the human operator. In this research, taking the advantages of AWB system, the multimedia based assembly information supporting system is developed, shown in Figure 1.



(a) (b)

Figure 3: Manipulators and assembly workbench.

Voice equipment

To let the operator understand the assembly information easily, speaker and wireless Bluetooth earphone are adopted to guide the operator by voice information.

Laser pointer

Showing the assembly position to the operator is an effective way to reduce the assembly mistake. In this case, a laser pointer is used to point out the assembly position. The laser pointer is fixed at the left top of the workbench, and the manipulator A is used to hold the assembly part, illustrated in Figure 3 (b). When the operator finishes one assembly step and changes to another step, the manipulator A will change the position of the assembly part to let the laser pointer point to the next assembly position.

LCD TV

Although the projector can provide information to the operator, however, its light can be shaded easily, especially by the operator's own hands. Furthermore, projector's light can injure human eyes, such as cataract illness. Hence, a

LCD TV is adopted to provide assembly information. The LCD TV is laid at the top of assembly workbench and covered by a transparent assembly board, in Figure 3 (b). The whole assembly scheme is divided into several simple assembly steps, and the corresponding assembly information is written in PowerPoint slides [8]. During the assembly process, these PowerPoint slides are inputted into the LCD TV and switched by footswitch.

Footswitch

During the assembly process, it is difficult for the operator to switch the PowerPoint slides by hands. Therefore, footswitch is employed, shown in Figure 1. There are two kinds of footswitches: footswitch A has three buttons, and footswitch B has only one button. Just pushing the different buttons of the footswitch A, the operator can change the PowerPoint slides either forward or backward. Pushing the button of footswitch B, the operator can control the manipulators to supply the necessary assembly parts to the operator, or let the manipulator A change the position and orientation of the assembly part during the assembly process.

Assembly supporting information

To reduce the assembly burden, the assembly supporting information is provided to the human operators. The whole assembly scheme is divided into several simple assembly steps, and the corresponding assembly information is written in PowerPoint slides [8], shown in Figure 4. In each PowerPoint slide, the assembly parts and assembly tools are illustrated by pictures; the assembly positions are pointed out by colourful marks; following the assembly flow chart, the videos of the experienced operators' assembly motions will appear to guide the novices to execute the assembly tasks. To make the operator understand the assembly process easily, the colours of the words in the slides are same with the real assembly parts' colours, such as "blue cable", "grey cable" in Figure 4. During the assembly process, these PowerPoint slides are output into the LCD TV and switched by the footswitch during the assembly process.

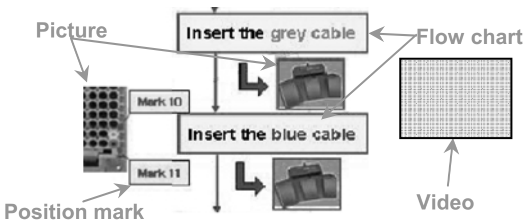


Figure 4: Assembly information of one step.

4.2 Assembly information supporting database

In this multimedia based assembly supporting system, the assembly information files are classified into document files, audio files and video files. The assembly guidance is concisely written in the document files; the voice guidance of each assembly step is record in the audio files; after the experienced operator's standard motions are record and analyzed into primitive assembly motions, they are saved into video files. In Jeffrey's research [9], he set up an assembly information support database to preserve all these assembly information files and provide them to the human operator depending on the situation. This database contains training data and assembling data: training data are designed for novice, and the assembly information files contain assembly details; assembling data are aimed to guide the experienced operators by the assembly sequence, which do not contain assembly details. As a consequence, this system may promote both the experienced operators and the novices to enter the workforce.

5 OPERATOR'S FATIGUE MONITOR SYSTEM

Since the human operator is the most important part in the cell production system, in order to create a comfortable working environment, operator's fatigue conditions should be monitored. Operator's fatigue conditions can be classified into muscle fatigue, eye fatigue and mental fatigue [10]. In this system, respiratory sinus arrhythmia (RSA) method [11] is used to monitor operator's muscle fatigue conditions. During the assembly process, eye fatigue also influences the yield rate and productivity. To improve the assembly supporting process and reduce eye fatigue, the critical flicker fusion frequency (CFF) method is adopted to measure the eye fatigue. Different from other production system, besides assembly task, the operator has to deal with multimedia information; therefore operator's mental fatigue cannot be ignored. According to Pongthanya's research [10], the following parameters are collected during the assembly process as the information to indicate mental fatigue: critical flicker fusion frequency (CFF), electroencephalogram (EEG), skin potential level (SPL), electromyography (EMG), respiration and subjective fatigue scores [11]. Based on the experiment results, the effect of the assembly supporting system can be improved.

6 SAFETY STRATEGY

During the assembly process, several practicable safety strategies are designed to avoid accident and protect human operators.

Working sequence

In the assembly system, two manipulators are utilized to support the human operator to execute the assembly tasks in physical aspect. Without special safety strategy, human operator is very possible to be hurt by the manipulators. Effective working sequence is one of the effective ways to reduce the possibility of collision between the human operator and the manipulators. Either the manipulator cannot move towards the human operator when the operator executes the assembly task. Due to this special working sequence, collision possibility can be partly reduced.

Footswitch

To realize the proposed working sequence, footswitch is used to control the manipulators, illustrated in Figure 1. When the human operator finishes one assembly step, he or she should push the button of footswitch B, and then manipulators will provide the assembly parts for the next assembly step to the operator.

Light curtain

To improve the safety of the system, light curtains are fixed between the human operator and the manipulators. In order to accelerate the development period and reduce the cost, this experiment made in the simulator first, illustrated in Figure 5. When the manipulators move too close to the human operators and cross the light curtain, the power of the manipulator will be cut down by the light curtain. As a consequence, the manipulators will stop.

Emergent button

When an accident happens, the human operator can just push the emergent button at the right side of assembly workbench to stop the whole system, shown in Figure 5. After the accident has been solved, the operator can push the reset button to restart the assembly system.

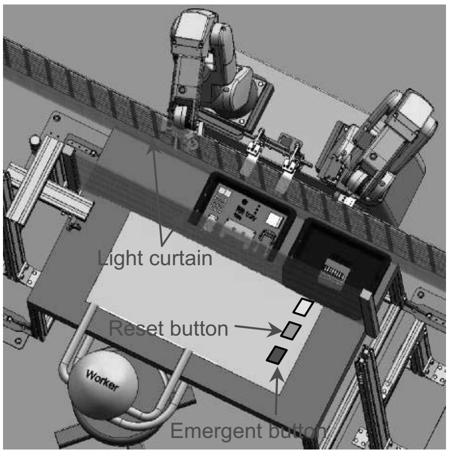


Figure 5: Safety strategy of the system.

Inherent safety theory

Although several safety strategies are adopted, there is no guarantee that the collision between manipulator and human operator will never happen. Thus, the manipulators should be ameliorated according to inherent safety theory [12] to reduce the injury of human operator. The sharp-angled brims of the manipulators are changed into obtuse-angled brims; the force and speed of the manipulators are reduced as low as possible while meeting the assembly requirement; the gravity center of the manipulators are descended to avoid them falling down.

7 CONCLUSION AND OUTLOOK

In this paper, taking the advantages of several previous researches but mediating their disadvantages, the multimedia based assembly supporting system for cell production is proposed. Depending on the MASTER system, the human operator can obtain effective physical support and assembly information support. Based on the assembly information database, this system is capable to meet the assembly and training requirements of both the experienced operators and the novices. Besides developing the real system, in order to reduce the design period and the development cost, a simulator of the whole system is created. To protect human operator from harm, several safety strategies and equipments are discussed. According to inherent safety theory, two manipulators are ameliorated, which will reduce the injury of human operators even they are collided by the manipulators.

In the future, the operator's fatigue conditions and intentions should be monitored during the assembly process, which will improve the operator's working comfort and the assembly efficiency.

8 ACKNOWLEDGMENTS

This project for strategic development of advanced robotics elemental technologies is supported by NEDO. We extend our sincere thanks to NEDO for their supports to the present research.

9 REFERENCES

[1] Seki, S., 2003, One by One Production in the "Digital Yatai" − Practical Use of 3D-CAD Data in the Fabrication, Journal of the JSME, 106/1013:32-36.

[2] Reinhart, C., Patron, C., 2003, Integrating Augmented Reality in the Assembly Domain − Fundamentals, Benefits and Applications, Annals of the CIRP, 52/1:5-8.

[3] Vasilash, G., 1999, Cell for Assembly, Automotive Manufacturing & Production.

[4] Hayakawa, Y., Kitagishi, I., Sugano, S., 1998, Human Intention Based Physical Support Robot in Assembling Work, Proceedings of IROS, 930-935.

[5] Sugi, M., Nikaido, M., Tamura, Y., Ota, J., Arai, T., Kotani, K., Takamasu, K., Shin, S., Suzuki, H., Sato, Y., 2005, Motion Control of Self-Moving Trays for Human Supporting Production Cell "Attentive Workbench", Proceedings of ICRA, 4080-4085.

[6] FANUC ROBOGUIDE, Robot System Animation Tool, http://www.fanuc.co.jp/en/product/robot/roboguide/index.html.

[7] Neider, J., Davis, T., Woo, M., 1993, OpenGL Programming Guide, Addison − Wesley Publishing Company.

[8] Zhang, Y., Duan, F., Tan, J.-T.-C., Watanabe, K., Pongthanya, N., Sugi, M., Yokoi, H., Arai, T., 2008, A Study of Design Factors for Information Supporting System in Cell Production, 41st CIRP MS.

[9] Tan, J.-T.-C., Duan, F., Zhang, Y., Watanabe, K., Pongthanya, N., Sugi, M., Yokoi, H., Arai, T., 2008, Assembly Information System for Operational Support in Cell Production, 41st CIRP MS.

[10] Pongthanya, N., Duan, F., Tan, J.-T.-C., Zhang, Y., Watanabe, K., Sugi, M., Yokoi, H., Arai, T., 2008, Evaluating Assembly Instruction Methods in Cell Production System by Physiological Parameters and Subjective Indices, 41st CIRP MS.

[11] Kotani, K., Furukawa, K., Takamasu, K., 2003, Accurate Extraction of Respiratory Sinus Arrhythmia with Respect to Respiratory Information, Transactions of the JSMBE, 41/3:197-204.

[12] Ikeda, H., Saito, T., 2005, Proposal of Inherently Safe Design Method and Safe Design Indexes for Human-Collaborative Robots, Specific Research Reports of the NIIS-SRR-NO.33:5-13.

Development of a Computer Aided Procedure to Control Division of Labour Based Disassembly Systems

Gert Zülch[1], Jan Hrdina[1] and Rainer Schwarz[2]

[1] ifab-Institute of Human and Industrial Engineering, University of Karlsruhe, Germany

[2] Faculty of Mechanical Engineering, formerly ifab-Institute of Human and Industrial Engineering, University of Karlsruhe, Germany

Abstract

In contrast to manufacturing and assembly systems, disassembly systems are marked by non-deterministic work sequences. The control of disassembly systems which adequately reacts to stochastically induced plan deviations are, therefore, of great importance for an efficient operation of disassembly systems. Thus, special control procedures, aligned to the requirements of disassembly, are required in order to manage disassembly systems appropriately. In order to improve disassembly control strategies, a simulation-aided control procedure for systems based on division of labour is under development to support the following tasks: Planning of the sequence of the disassembly orders, variation of the disassembly depth dependent upon the capacity utilization and variation of the disassembly operations according to the capacity utilization.

Keywords:

Disassembly Systems; Disassembly Control Strategies; Simulation

1 INTRODUCTION

1.1 Regulatory framework for disassembly systems

Through the implementation of the ElektroG act passed in 2005 [1], the objectives of the European directive 2002/96/EC (2003) [2] of closing material cycles, preserving resources and avoiding environmental impacts of contaminants, have been realized in German legislation [3, 4]. As a result, as of 2006, manufacturers of electrical and electronic devices are required to take back their used products for reuse or disposal. Manufacturers must thereby fulfil certain environmental standards while taking the state of technology into account. This means that, should a used product not be reusable, fluids and toxic substances for example, must be removed and individual components detached.

According to another decree, the AltfahrzeugV from 2003, the same applies to the automobile industry: "Automobile manufacturers are required to take back all used vehicles of their models from the last owner. The vehicle manufacturers must … take back used automobiles, once they have been returned to an accredited return depot or to a designated disassembly establishment, free of charge" (AltfahrzeugV 2003, § 3).

In recent years manufacturers have established recycling departments for this purpose or commissioned specialized companies with the recycling of the used vehicles. The goal of all this is to reintroduce the raw materials used into the production cycle as either secondary raw materials or secondary products and to dispose of recycling by-products in an environmentally sound manner.

1.2 Consequences for the design of disassembly systems

The current legal situation forces manufacturers to anticipate an increased return of used products over the next few years. This expected increase will in turn force manufacturers to optimize the disassembly of used products [5]. In doing so, certain environmental standards must be fulfilled while taking the state of technology into account. During the implementation of the legally prescribed recycling obligations, manufacturers must take care to align the recycling operations with ecological objectives, but also in an economical manner. In particular when reclaiming secondary products for the recovery of spare parts, an automated disassembly of mechanical components is extremely elaborate and usually only feasible for certain used products. Therefore, in the future, manual disassembly will still make up the bulk of disassembly operations.

A first step for this is the planning of a sufficiently flexible disassembly system for a given spectrum of used products (for more details see [6]). The main objective of controlling disassembly systems is to allow for and maintain efficient disassembly operations. A computer-aided control procedure must be able to plan a favourable dispatching sequence for disassembly orders for a specified planning period. It must then also be equipped to react adequately to possible influences (e.g. caused by disturbances) [6]. This calls for the development and integration of appropriate disassembly control strategies.

1.3 Demand for a disassembly control concept

In contrast to parts manufacturing and assembly, disassembly is characterized by non-deterministic work sequences [7]. The specific problems associated with disassembly arise e.g. as a result of disassembly programmes composed of a mix of used products, a varying order arrival patterns and stochastic disassembly operation times. This necessitates the use of a specialized control procedure adapted to the requirements of disassembly. Accordingly, a procedure for controlling disassembly systems with division of labour must support the following functions:

- planning of the sequence of disassembly orders;
- variation of the disassembly depth dependent upon the capacity utilization;

- modification of the disassembly operations dependent upon the capacity utilization;

2 MODELLING OF DISASSEMBLY SYSTEMS

For the investigation of possible control strategies, simulation has proved to be a useful tool for evaluating their effects prior to implementation. In order to model a disassembly system realistically in a simulation procedure, the following configuration elements are required:

- order structure and arrival pattern
  - different product types (characterized by their disassembly operations and the possibility to vary the disassembly depth);
  - order arrival pattern (characterized by the distribution of the inter-arrival times of orders or by a pre-defined order sequence);
  - variable batch sizes of disassembly orders;
- resources
  - disassembly workplaces and their specific equipment (e.g. disassembly tools, mechanized devices);
  - staff (i.e. the workers employed, their working time models, abilities and workplace assignments);

3 CONCEPT TO CONTROL DIVISION OF LABOUR BASED DISASSEMBLY SYSTEMS

Industrial disassembly systems are typically based on division of labour (for planning of disassembly systems consult e.g. [6, 7, 8, 9]). In addition to allow for a uniform utilization of the workplaces, minimization of the makespan of an order programme or the lead times of the individual orders and a maximization of output or profit in a given period of time are central objectives of disassembly controlling [7, 10]. The disassembly procedure should achieve this by dispatching orders in a favourable sequence during the given planning period.

In order to achieve the abovementioned goals, established strategies from the area of manufacturing [11, 12] as well as other control strategies are available. This often results in greater flexibility regarding the sequence of disassembly orders to be carried out than is the case within assembly systems. This can be traced back to the smaller number of technological sequence constraints, when compared to assembly, and to a larger selection of applicable disassembly procedures and tools. This flexibility in the sequence of disassembly operations can be used to control a disassembly programme in order to avoid sequence-related idle times.

In order to lessen fluctuations in the utilization of workplaces, the disassembly depth can be varied. Dependent upon the utilization of the individual workplace, one can decide, for example, whether the planned disassembly depth should be maintained or modified. If a workplace has a high utilization, for example, the disassembly depth can be reduced in order to decrease the work volume at this place. This shortened disassembly operation in turn reduces idle time which may occur at the subsequent workplaces due to lacking work pieces. The preceding workplace can then concentrate on the disassembly operations promising the greatest added value and skip those with lesser added value potential. Conversely, if a workplace has a low utilization, the disassembly depth can be increased. This would mean that disassembly operations which are not economical when the system has a normal utilization, but still have a certain added value potential, could then be carried out. This allows the work loads at the individual workplaces to be balanced.

A further possibility to eliminate sequence-related idle times during order processing is to avoid a strict assignment of disassembly operations to workplaces. Dependent upon the utilization, work operations can be transferred from one workplace to a neighbouring one, as long as this is technically possible. This possibility is given when the necessary disassembly tools are available there or the executed disassembly operations are comparably simple. Furthermore it affords, that the extracted parts or components can be collected and sorted at both workplaces.

In this context, the selection of appropriate disassembly methods is of great significance. Therefore, if the utilization at one workplace is too high, a non-destructive but time-consuming disassembly method can be abandoned in favour of a more destructive but time-efficient one, thus improving the utilization of the disassembly system. The fact that a change of disassembly methods can lead to limitations regarding the reuse or recycling of the extracted parts and components must not be disregarded.

The control concept must be able to take both fixed as well as variable assignments of workers to disassembly workplaces into account. Varying working times of the workers may also present a further constraint. An all-encompassing control concept must, therefore, consider not only the sequence planning and the situation-dependent disassembly depth, but must also contain functionalities of personnel control.

In addition to organizational considerations, product-specific particularities must also be taken into account. The disassembly control concept must regard the fact that each used product has been subject to a diversity of influences throughout its lifecycle (corrosion, wear, deformation etc.) and is thus in a different state when it comes to recycling. Therefore, stochastic influences (e.g. when a bond cannot be disconnected or only with great difficulty), which make planned disassembly sequences unfeasible, must be taken into account. The control concept thus requires a module initiating a change of disassembly method dependent upon, for example, a workplace's utilization or stochastically occurring disturbances due to worn parts.

In order to incorporate the abovementioned disassembly-specific controlling aspects into a encompassing concept and to consider stochastically occurring influences in the dynamic winding-up of orders, the control concept was integrated into the object-oriented simulation procedure *OSim* [13, 14] (cf. chapter 5 for more details), which was developed at the ifab-Institute of Human and Industrial Engineering of the University of Karlsruhe (Germany). Throughput diagrams, which are networks representing the temporal-logistical dependencies of the work operations in the disassembly of used products, are thereby used to model the various used products.

4 ASSESMENT OF THE CONTROL PROCEDURE

In order to assess whether a control procedure is suitable for a specific disassembly system or not, an appropriate assessment concept is required. Here, the concept based on degrees of goal achievement is used. A degree of goal achievement can take on a value between 0 % and 100 %, where 0 % represents the pessimistic and 100 % the optimistic value of this index. Following this concept, a reduction of the lead time would therefore lead to an increase in the goal achievement lead time degree (for details about the productions logistical and monetary assessment of simulation results, consult e.g. [15, 16, 17]). Using the concept of goal achievements allows for the aggregation of competitive goal criteria into a comprehensive figure. For this purpose, the following individual key figures may be considered in order to assess the control procedure:

- order-oriented degrees of goal achievements
    - lead times degrees;
    - waiting times of the staff;
    - idle times of workplaces and technical devices;
    - number of completed orders;
    - makespan of the order programme;
- disassembly-oriented degrees of goal achievements
    - disassembly depth;
    - disassembly handling costs;
    - profits from disassembly orders;
- personnel-oriented degrees of goal achievements
    - utilization of human resources;
    - degree of physical stress;
    - degree of time stress;

5 THE SIMULATION PROCEDURE *OSim-DPS*

The transfer of the described control and assessment concept into the object-oriented simulation procedure *OSim-DPS* (German abbreviation for "*O*bject *Sim*ulator for Disassembly Planning and Control"; [15], see Figure 1), which was developed at ifab-Institute, allows for a prospective and quantitative assessment of various disassembly control strategies. Furthermore, it is possible to derive configuration recommendations for operations and process organization.

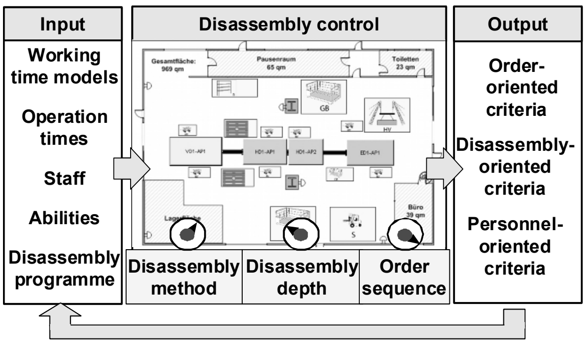


Figure 1: The simulation procedure *OSim-DPS*.

Figure 2 shows an excerpt of an example throughput diagram for the disassembly of computers of the make "Small Tower". So-called triggers are assigned to these kinds of throughput diagram which describe the arrival of used products into the disassembly system. The operation time of a disassembly operation can either take on a constant value (as shown here) or be subject to a statistical distribution (e.g. beta distribution).

In order to verify the simulation procedure *OSim-DPS*, pilot studies are performed. For this purpose, several disassembly systems are modelled based on real data and then assessed. If the deviation between the real and simulated data is minimal, the expanded simulation procedure can be regarded as valid [18].

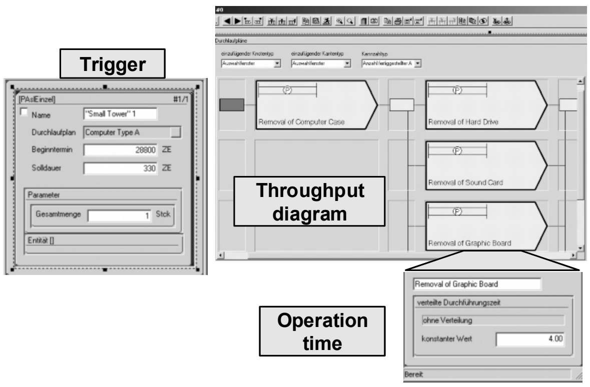


Figure 2: Typical throughput diagram and trigger for a computer disassembly.

In order to prove the effectiveness of a specific control procedure, results from the simulation are compared to empirically collected performance data from real disassembly systems (or the initial situation). In practice, the situation-dependent choice between different disassembly methods (e.g. destructive vs. non-destructive disassembly methods), for example, is of interest.

One of these pilot applications deals with the disassembly of PCs. PCs with various fittings (e.g. with or without a DVD drive) are thereby released into the disassembly system. One difficulty results from the fact that it is not known at the time of dispatching whether the disassembly plan can be adhered to (for example, has the hard drive already been removed by the former owner or not, which can only be ascertained after the housing has been opened). The first results show that the possibility of reallocating disassembly operations to alternate workplaces can increase personnel utilization in comparison with the originally planned sequence. This in turn increases the number of disassembly orders which can be processed in a given time period.

6 SUMMARY AND FURTHER RESEARCH

Due to legislative changes, recycling in the broader sense and within this context disassembly systems based on division of labour in the narrower sense become more and more important for goods-producing companies. In contrast to parts manufacturing and assembly, disassembly planning has to cope with more stochastic influences, which may make planned disassembly sequences obsolete. Therefore, an adequate tool for re-arranging order sequences and the assignment of operations to disassembly stations becomes necessary in order to obtain efficient disassembly processes. The presented paper dealt with the concept of a computer-aided procedure to control disassembly systems based on division of labour to attain these objectives.

Simulation studies are currently being carried out in cooperation with disassembly companies in order to validate the efficiency of the procedure. For this purpose, the effects of various disassembly control strategies will be examined with regards to their achievement of various production logistic and personnel-oriented goals.

In addition to the order- and product-oriented disassembly control, it is intended to develop strategies for personnel employment control and to integrate them into the computer-aided procedure. For example, an alternative to the flexible assignment of work operations to individual workplaces could be the flexible allocation of workers to workplaces. For this, various forms of partly autonomous groups are conceivable, in which the workers help each other when needed. Furthermore, workers from other departments can be called upon in order to eliminate short-term bottlenecks in the disassembly order processing. Moreover, the implementation of flexible working times can lead to a further form of personnel management. Simulation tools are very helpful to evaluate the effects of these organizational measures prior to their implementation in a specific disassembly system.

7 ACKNOWLEDGEMENTS

The authors would like to thank the German Research Association (Deutsche Forschungsgemeinschaft) for its support in the research work referred to in this paper. These endeavours are currently sponsored within a project dealing with the development of a computer-aided procedure to control disassembly systems based on division of labour ("Entwicklung eines rechnerunterstützten Verfahrens zur Steuerung von arbeitsteiligen Demontagesystemen").

8 REFERENCES

[1] ElektroG, 2005, Gesetz über das Inverkehrbringen, die Rücknahme und die umweltverträgliche Entsorgung von Elektro- und Elektronikgeräten (Elektro- und Elektonikgerätegesetz).

[2] 2002/96/EC, 2003, Directive 2002/96/EC of the European Parliament and of the Council of 27 January 2003 on waste electrical and electronic equipment (WEEE).

[3] BMU - Bundesministerium für Umwelt, Naturschutz und Reaktorsicherheit (edt.), 2008, Leitfaden Abfallwirtschaft und EU-Erweiterung, http://www.bmu.de/files/pdfs/allgemein/application/pdf/leitfaden_eu_erweiterung.pdf, 18.01.2008.

[4] Wolf, M., 2005, Vorbereitungen zur Rücknahme und Verwertung von Elektroschrott auf vollen Touren, Pressemitteilung 145/2005, Bundesministerium für Umwelt, Naturschutz und Reaktorsicherheit, Berlin, http://www.bmu.de/pressemitteilungen/pressemitteilungen_ab_01012005/pm/35607.php, 09.05.2006.

[5] Zülch, G., Müller R. , Schiller, E., 2000, Information System for Supporting Manual Disassembly Tasks, ICME 2000, The Eighth International Conference on Manufacturing Engineering, Causal Productions (edt.), Compact disk file:///D|/HTML/ME00S104.HTM.

[6] Zülch, G., Schwarz, R., 2006, Planning and Balancing of Disassembly Systems, Lean Business Systems and Beyond, Advanced Production Management Systems Conference 2006 of the IFIP Working Group 5.7, Preprints, Koch, Tomasz (edt.). University of Technology, Wroclaw, 325-330.

[7] Schiller, E., 1998, Ein Beitrag zur adaptiv-dynamischen Arbeitsplanung in der Demontage, Shaker, Aachen.

[8] Zülch, G., Schiller, E., Schneck, M., 1996, Adaptive Dynamic Process Plans - A Basis for a Disassembly Information System, ECO-Performance '96, Züst, R., Caduff, G., Frei, M. (edts.), Verlag Industrielle Organisation, Zürich, 245-253.

[9] Zülch, G., Schiller, E., Müller, R., 1997, A Disassembly Information System, Life Cycle Networks, Krause, F.-L., Seliger, G. (edts.), Chapman & Hall, London et al.: 400-412.

[10] Zülch, G., Schwarz, R., 2003, Demontagesystemplanung an einem Beispiel aus der Elektroindustrie, Good Practice. Ergonomie und Arbeitsgestaltung Landau, K. (edt.,). ergonomia Verlag, Stuttgart: 347-362.

[11] Zülch, G., 1989, Der strategische Steuerungsraum - Ein Ansatz zur Systematisierung von Strategien der Fertigungssteuerung, VDI-Zeitschrift, 131/5:58-65.

[12] Zülch, G., 1991, Strategien zur Steuerung von Fertigungssystemen, Fortschrittliche Betriebsführung und Industrial Engineering, 40/2: 52-58.

[13] Jonsson, U., 2000, Ein integriertes Objektmodell zur durchlaufplanorientierten Simulation von Produktionssystemen, Shaker, Aachen.

[14] Zülch, G., Fischer, J., Jonsson, U., 2000, An integrated object model for activity network based simulation, Proceedings of the 2000 Winter Simulation Conference, Joines, J. A., Barton, R. R., Kang, K., Fishwick, P. A. (edts.), Institute of Electrical and Electronics Engineers et al., Piscataway, NJ, 1: 371-380.

[15] Zülch, G., Grobel, T., Jonsson, U., 1995, Indicators for the Evaluation of Organizational Performance, Benchmarking - Theory and Practice, Rolstadås, A., (edt.), Chapman & Hall, London et al.: 311-321.

[16] Zülch, G., Brinkmeier, B., 1998, Designing organizational structures of production systems using a process-oriented approach, Okino, N., Tamura, H., Fujii, S. (edts.), Advances in Production Management Systems - Perspectives and future challenges, Chapman & Hall, London et al.: 183-194.

[17] Zülch, G., Grobel, T., 1996, Shaping the organization of order processing with the simulation tool FEMOS, International Journal of Production Economics, 46-47:251-260.

[18] Zülch, G., Bogus, T., Krüger, J. 1999, Decision theory based approach for modeling and simulating the activity preferences of workers, Axelsson, J., Bergmann, B., Eklund, J. (edts.), Proceedings of the International Conference on TQM and Human Factors - towards a successful integration, volume 2, Centre for Studies of Humans, Technology and Organization, Linköping: 211-216.

A New Method to Control Work-in-Process Inventory for High Manufacturing Productivity

Rei Hino, Norikazu Suzuki and Eiji Shamoto[1],
[1] Department of Mechanical Science and Engineering, Nagoya University, Nagoya, Japan

Abstract
A new method for controlling the WIP inventory is proposed in order to restrict the excessive production of un-finished products, while the manufacturing productivity is kept at a certain high value. Excessive WIPs decrease business profits and prevent the realization of agile manufacturing along with changes in markets, while insufficient number of WIPs reduce manufacturing productivity and flexibility to cope with malfunctions of production resources. Theoretical formulations of a scheduling problem without buffer is derived, and heuristic procedure is adopted to plan a feasible schedule. The proposed control methods are verified by computational simulations regarding job-shop type production system.

Key Words :
Work in Process, Buffer, Productivity, Mixed Integer Programming, Graph

1 INTRODUCTION

There are many important issues in manufacturing that allow businesses to overcome strict demands from drastically changing markets. Establishment of a control meth-od for works-in-process (WIPs) is one of these issues, and takes priority over the others. Generally speaking, we cannot help accepting a certain WIP in order to reduce idle machine times. However, excessive number of WIPs causes a long lead time for each product and introduces the risk holding unnecessary stock in the markets. On the other hand, an appropriate number of WIPs protects production from troubles such as machine failures and shortages of materials.

We propose a new method to control the number of WIPs in a production factory the perspective of scheduling procedures. Scheduling is one of the most important activities in the effective management of current manufacturing systems. The scheduling problem can be described simply, but it is difficult to optimize because there are many candidate solutions[1-2]. Therefore, a great deal of researchers has been carried out so as to establish a model and to solve for optimized schedules under various conditions[3-7]. Conventional scheduling problems are formulated around typical production resources, such as machine tools and robots. We extend the formulation to storage as a buffer in order to evaluate the status of WIPs in the planned schedule. An un-finished product, or the WIP, is stored in a buffer between two resources. Therefore, fundamental differences of roles of the the buffer and the resource are discussed and described from the perspective of scheduling. As a result, the status of the WIP in the buffer is identified as blocking of the resource, and the statuses of buffer and resource are integrated into common numerical formulations.

In the present paper, we first derive theoretical formulations of a scheduling problem for production resources without buffer to obtain an optimum solution. We then propose a new feasibility judgement method through the use of graph representation is proposed to plan the schedule by heuristic procedures under given buffers. Numerical simulations are executed to discuss the role of the buffers and to verify the validity of the proposed method.

2 SCHEDULING PROBLEM INCLUDING STATUS OF RESERVATION

The conventional scheduling problem is formulated under the condition of infinite buffer capacities. Therefore every production resource can begin any new operation immediately after accomplishment of the given operation. In this study, it is assumed that each production resource has to wait to remove the accomplished job from itself to an empty buffer. In other words, if there is no available buffer, no production resource can begin any new operations. The new formation of a scheduling problem is described in this section in order to consider the removal of the job from the production resource. The time from the accomplishment to removal of the job is identified as blocking of the resource. In this paper, a status of reservation by the buffer equals the blocking. The following symbols are used for formulation of the problem.

$j_{\alpha,i}^{\psi,v}$: vth operation processed by resource ψ and i th process for product α.

$\bar{j}_{\alpha,i}^{\psi,v}$: reservation for $j_{\alpha,i}^{\psi,v}$

s,f,p : start time, finish time and duration of operation

$\bar{s},\bar{f},\bar{p}$: start time, finish time and duration of reservation

In conventional scheduling, the start time of the job $j_{\alpha,i}^{\psi,v}$ are determined according to the processing order of jobs assigned to the resources.

$$s_{\alpha,i}^{\psi,v} = max(f^{\psi,v-1}, f_{\alpha,i-1}) \tag{1}$$

where, $j^{\psi,v-1}$ is the job processed by the same resource ψ in the $v-1$ th order, and $j_{\alpha,i-1}$ is the preceding job for product α. According to the result, the finish time of the job is calculated

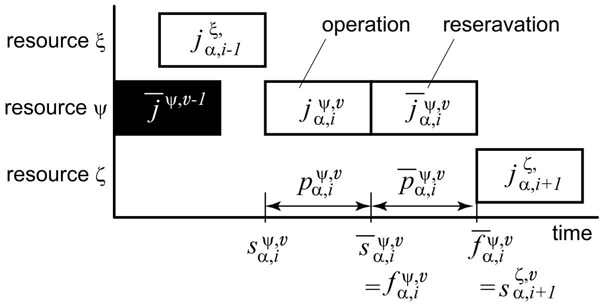


Figure 1: Operation and reservation of job represented on Gantt chart

by following equation,

$$f^{\psi,v}_{\alpha,i} = s^{\psi,v}_{\alpha,i} + p^{\psi,v}_{\alpha,i} \quad (2)$$

where, $p^{\psi,v}_{\alpha,i}$ is the time required for the operation; this value is given as one of the preconditions for the scheduling problem. Figure 1 shows the status of the operation and reservation of the job. In the proposed scheduling problem, the start and finish times of operation of the job are determined as follows,

$$s^{\psi,v}_{\alpha,i} = max(\bar{f}^{\psi,v-1}, f^{\xi,}_{\alpha,i-1}) \quad (3)$$

$$f^{\psi,v}_{\alpha,i} = s^{\psi,v}_{\alpha,i} + p^{\psi,v}_{\alpha,i} \quad (4)$$

As expressed in Equation (3), the start time of the operation is restricted by the finish time of the reservation $\bar{f}^{\psi,v-1}$. The production resource can remove the job if another succeeding production resource receives the job and begins operation of the job. The start time of the succeeding operation equals the finish time of the reservation on the production resources. The start and finish times of the reservation are determined as follows,

$$\bar{s}^{\psi,v}_{\alpha,i} = f^{\psi,v}_{\alpha,i} \quad (5)$$

$$\bar{f}^{\psi,v}_{\alpha,i} = s^{\zeta,}_{\alpha,i+1} \quad (6)$$

After the times are fixed, the duration of the reservation is calculated by the following equation.

$$\bar{p}^{\psi,v}_{\alpha,i} = \bar{f}^{\psi,v}_{\alpha,i} - \bar{s}^{\psi,v}_{\alpha,i} \quad (7)$$

3 FORMULATION OF A SCHEDULING PROBLEM BY MIXED INTEGER PROGRAMMING

To discuss significant features of the reservation, the proposed scheduling problem in the case of no buffer is formulated by mixed-integer programming. The proposed formulation for the problem is given as follows,

Minimize C_{max} (8)

Subject to

$$C_{max} \ge f^{\psi,}_{\alpha,n_\alpha} \quad (9)$$
$$s^{\psi,}_{\alpha,i} \ge 0$$

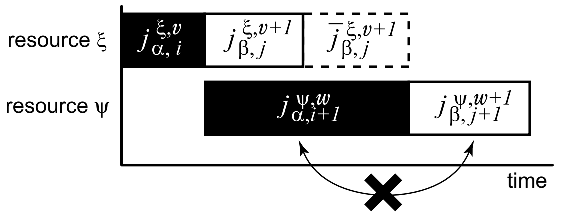


Figure 2: Incongruous exchange of operating order

$$f^{\psi,}_{\alpha,i} = s^{\psi,}_{\alpha,i} + p^{\psi,}_{\alpha,i} \quad (10)$$

$$s^{\psi,}_{\alpha,i+1} \ge f^{\xi,}_{\alpha,i} \quad (11)$$

$$s^{\psi,}_{\beta,j} \ge \bar{f}^{\psi,}_{\alpha,i} - M\cdot(1-x^{\psi}_{\alpha,\beta})$$
$$s^{\psi,}_{\alpha,i} \ge \bar{f}^{\psi,}_{\beta,j} - M\cdot x^{\psi}_{\alpha,\beta} \quad (12)$$
$$x^{\psi}_{\alpha,\beta} \in \{0,1\},\ \forall\psi\in R,\ \alpha\neq\beta$$
$$\forall\alpha,\beta\in P,\ \forall i\in Q_\alpha\ \forall j\in Q_\beta$$

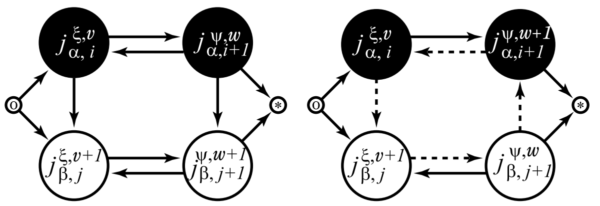
where, C_{max} is the makespan of the planned schedule. Sets P and R are for the products and resources respectively, set Q_α represents the set of the process of the product α, and n_α is the number of processes for product α. M is a large positive number and $x^{\psi}_{\alpha,\beta}$ is an integer, 0 or 1. If the job $j^{\psi,}_{\alpha,}$ precedes the job $j^{\psi,}_{\beta,}$ by the resource ψ, $x^{\psi}_{\alpha,\beta}$ is 0; otherwise it is 1. The virtual job $j_{\alpha,n_\alpha+1}$ is assumed to satisfy Equation (11), and $s_{\alpha,n_\alpha+1}$ equals the finish time of f_{α,n_α} the job j_{α,n_α}.

Equations (12) are different from the formulations for the conventional scheduling program. As expressed by the equations, every resource cannot begin operation of a new job after reservation of the preceding operation. The proposed formulation gives the optimal schedule of production resources under the condition of no buffer. However, the buffer is identified as the special production resource having a zero duration time for the operation. Therefore, the formulation can also be applied to a scheduling problem including finite buffers.

4 FEASIBILITY IDENTIFICATION BY USING GRAPH REPRESENTATION

In the present paper, we evaluate the planned schedule and modify it to make it more desirable by the addition or subtraction of buffers. There exist some prohibited job-processing orders in re-scheduling. Figure 2 shows an example of the prohibited job sequences. It shows a simple flow shop schedule for two products by two machines. The resource ξ cannot begin a job operation after the resource ψ receives the job. Therefore the resource ψ cannot begin operation of the job $j^{\psi}_{\beta,j+1}$ before the job $j^{\psi}_{\alpha,i+1}$, while the resource ξ operates the job $j^{\xi}_{\alpha,i}$ before the job $j^{\xi}_{\beta,j}$.

We propose a new graph representation to systematically identify the feasibility of the modified job-processing sequence. Figure 3 shows examples of the graph representation for the scheduling problem with buffer. The graph consists of nodes and two type of arcs, one conjunctive and the another disjunctive. The node represents the job operated production resources. The conjunctive arc represents processing sequences constrained by pre-condition of the scheduling problem. The disjunctive arc represents the job-processing sequences for one production



(a) feasible schedule (b) infeasible schedule

Figure 3: Proposed graph representation

resource. Both the conjunctive and disjunctive arcs have directions. In graph representation of the conventional schedule problem, the conjunctive arc is mono-directional, but in the proposed graph representation, the conjunctive arc is bi-directional. These results indicate that the finish time of the reservation of the job depends on the successive job, and that the start time of the job depends on the preceding job.

Figure 3(a) corresponds to the schedule shown in Fig.2 and the schedule is a feasible plan. On the other hand, Fig.3(b) shows a infeasible plan. As shown in the figure, the graph for an infeasible plan includes closed loop along the direction of the conjunctive and disjunctive arcs. Therefore, the feasibility of the planed schedule is identified by the proposed graph before the start and finish times of the all jobs are calculated by the Equations (3) to (7).

Figure 4 shows another example of feasibility identifications in a more complex schedule. Figure 4(a) shows a Gantt chart for four products from four production resources and the graph shown in Fig.4(b) corresponds to the planned schedule. After an exchange of the processing order for the job $j^{1}_{1,1}$ and $j^{1,}_{2,1}$, the graph is drawed as shown in Fig.4(c). Because the graph includes a closed loop, we can judge that the exchange is prohibited.

There are two exceptions of the proposed identification method. The Following graphs show feasible schedules despite the existence of the closed loop.

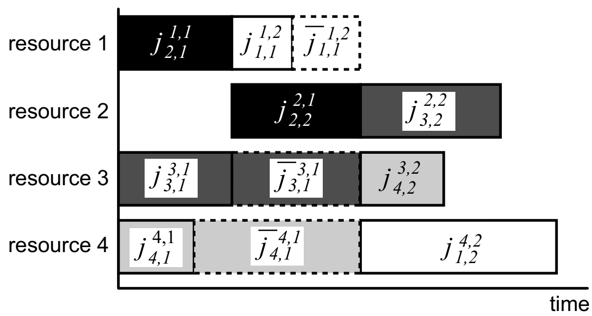
1) The closed loop contains no normal directed conjunctive arcs.
2) The closed loop contains a continuous reverse-directed conjunctive arc.

5 NUMERICAL STUDY OF WIP CONTROL

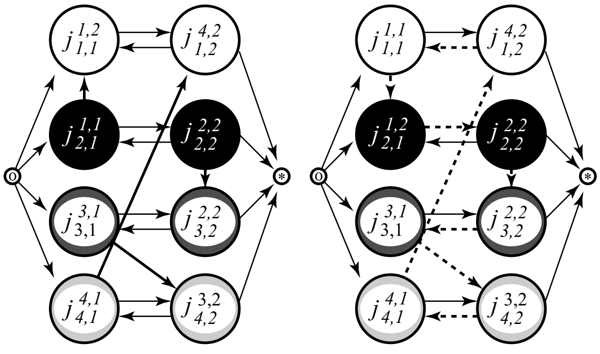
Computational simulations are executed to show the validity of the proposed theoretical formulations and an infeasible identification method. Table 1 shows the conditions well-known as Muth & Thompson's job shop scheduling problem for six products by six machines[1].

Table 1: 6 × 6 scheduling problem

No	Machine (Processing time)					
1	3(1)	1(3)	2(6)	4(7)	6(3)	5(6)
2	2(8)	3(5)	5(10)	6(10)	1(10)	4(4)
3	3(5)	4(4)	6(8)	1(9)	2(1)	5(7)
4	2(5)	1(5)	3(5)	4(3)	5(8)	6(9)
5	3(9)	2(3)	5(5)	6(4)	1(3)	4(1)
6	2(3)	4(3)	6(9)	1(10)	5(4)	3(1)



(a) an example of feasible schedule for products manufactured by four resources



(b) graph representation of feasible schedule shown in (a)

(c) graph representation of infeasible schedule caused by exchange of job processing order on resource 1

Figure 4: Examples of judgement of an infeasible schedule

Figure 5 shows an optimal schedule obtained by mixed integer programming under the condition of infinite buffer capacities, or conventional conditions. As shown in Fig.5, the total processing time, or makespan is the shortest, but many WIPs exist. Figure 6 shows the schedule with no buffer, and the schedule shows a longer makespan but does not require a buffer for the production. As such, the lead time of each product in the proposed problem is shorter than in the conventional problem.

We can know the status of the productions by evaluating the schedule without any buffers. For example, the second process for product 2, which is operated by machine 4, remains on machine 4 after the operation is complete, as shown in Fig.6, which means that machine 6 lacks processing performance because the third process of the product 6 is assigned to machine 6.

Figure 7 shows the modified schedule after the addition of one buffer for every machine. The modification is executed in a heuristic manner by randomly exchanging the operation order of two jobs. The proposed feasibility identification method can prevent the prohibited job operating sequences for searching an appropriate schedule. As shown in Fig.7, the additional buffers are used for improvement of the productivity and we can save the number of WIPs under the number of added buffers.

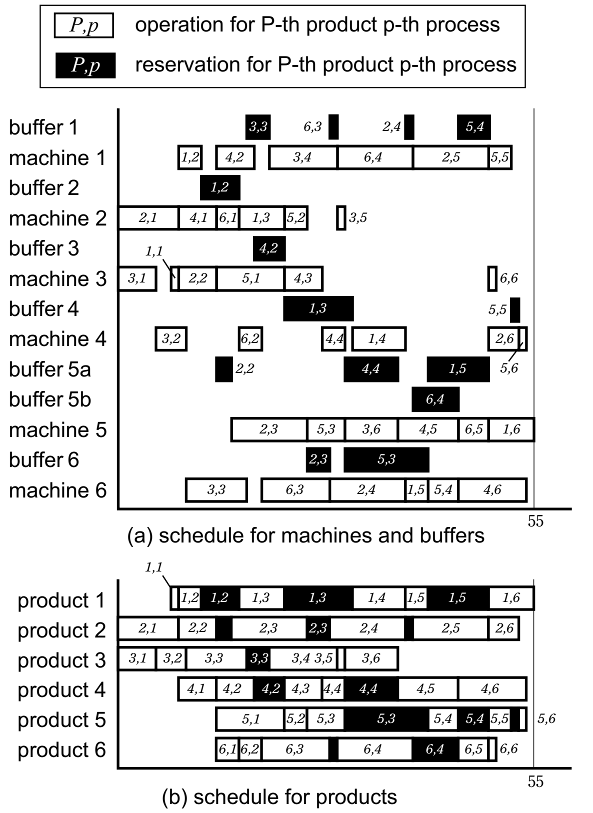


Figure 5: Optimal schedule for Muth & Thompson's 6 × 6 scheduling problem

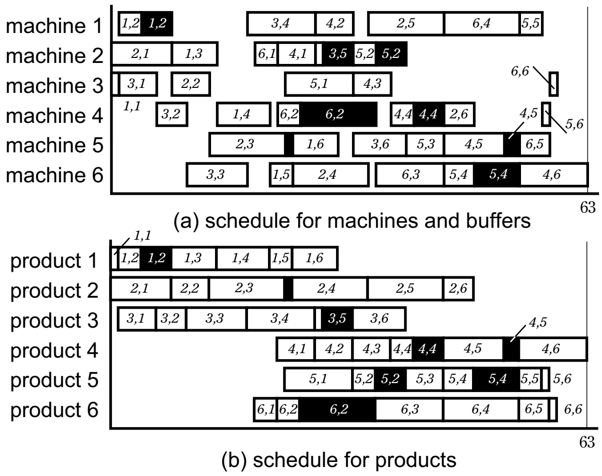


Figure 6: Optimal schedule without buffer

6 CONCLUSIONS

A new type of scheduling problem is defined to control the amount of work in a process, or WIP. A job, which is an essential element of the proposed scheduling problem, includes the status of reservation in addition to the operation by each production resource. The time of reservation is defined as the term between the accomplishment of the operation and the removal from the resource. In this paper, we propose theoretical formulations to plan an optimal schedule without buffer, and we also propose a new graph representation to achieve a heuristic schedule modification. The theoretical formulations are based on the mixed integer programming manner, and the optimal schedule is available to evaluate the performance of every production resource in manufacturing. The WIP means the semi-finished material on the buffer, and the number of buffers therefore restrains the number of WIPs.

Numerical simulations are executed to the general scheduling problem to verify the validity of the proposed methods. An optimal schedule is planed by the theoretical method, and then the heuristic method is applied in order to plan a feasible schedule based on the optimal schedule under the limited buffers.

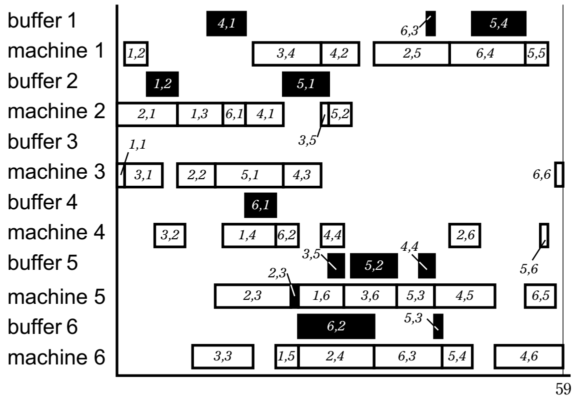


Figure 7: Near optimal schedule with finite buffer capacity

REFERENCES

[1] Muth, J, F. and Thompson, G, L., "Industrial Scheduling", Prentice Hall, Inc., Englewood Cliffs, New Jersey,1963.
[2] P. Brucker, "Scheduling Algorithms", Springer, 2001.
[3] Calrier, J. and Pinson, E., "An Algorithm for Solving the Job-shop Problem", Management Science, Vol.35, No.2, pp.164-176, 1989.
[4] R.Hino et. atl., Job Shop Scheduling Focusing on Role of Buffers, Transactions of the Japan Society of Mechanical Engineers, Series C, Vol.71, No.702 (2005) pp685-692.
[5] H.Tamaki et al., Solution of the Job-shop Scheduling Problem with Consideration of In-process Buffer Capacity, Transactions of the Society of Instrument and Control Engineers, Vol.31, No.8(1995), pp.1193–1201.
[6] G.Fu et.al., A Method for Solving Jobshop Scheduling Problems with Limited Common Buffers, Transactions of the Institute of Systems, Control and Information Engineers, Vol.17, No.3(2004), pp.113–121.
[7] Nicholas G. Hall, Chelliah Sriskandarajah, A Survey of Machine Scheduling Problems with Blocking and No-Wait in Process, Operations Research, Vol.44, No.3 (1996), pp.510-525.

An Evolutionary Algorithm for Vehicle Routing Problem with Real Life Constraints

Giuseppe Confessore[1], Graziano Galiano[1] and Giuseppe Stecca[1]

[1] Institute of Industrial Technologies and Automation, National Research Council, Italy

Abstract

Real life distribution problems present high degree of complexity mostly derived by the need to respect a variety of constraints. Moreover they are not considered by the classical models of the vehicle routing literature. In this paper we consider a vehicle routing problem with heterogeneous vehicle fleet with different capacity, multi-dimensional capacity constraints, order/vehicle, item/vehicle, and item/item compatibility constraints, different start and end locations for vehicles, and multiple time windows restrictions. We propose an evolutionary algorithm based on the combination of a genetic algorithm and local search heuristics. We investigated the performance of the implemented algorithm in the large-scale retail and in the waste collection industries.

Keywords:

Vehicle Routing Problem; Distribution Process; Evolutionary algorithm

1 INTRODUCTION

Real-life distribution processes are affected by high complexity due to elevated number of constraints to respect, to different optimisation criteria, to the need of practical extensions, and to need for responsiveness.

This work treats a dynamic real-life vehicle routing problem proposing an algorithm to solve a distribution problem in a manufacturing system. A basic distribution problem consists of a set of customers requiring the delivery of goods within given time windows and deliveries are done with capacitated vehicles departing from and returning to a depot. The goal consists in minimizing the number of necessary vehicles to effectuate the routing, or the total transportation cost (e.g. total distance covered by the set of vehicles). This is a vehicle routing problem that, although well known in the literature and very studied in its different variants, is often subject of strong simplifications not much suitable to the real world where the problem involves several real constraints.

In real-life problems we have a number of complexities that are not considered by the classical models found in the literature. In this paper we consider a generalised vehicle routing problem (CVRPTW) with a diversity of practical constraints. Among those are multi time window restrictions, a heterogeneous vehicle fleet with different capacity, travel times, temporal availability, travel costs, order/vehicle compatibility constraints, customers with multiple orders of pickup and delivery, different start and end locations for vehicles, route restrictions associated to orders, customers and vehicles, and drivers' working hours.

We propose an evolutionary approach based on the hybridisation of a genetic algorithm with insertion heuristic techniques. This algorithm was implemented and included as optimization module in a software product used by some companies in the transportation planning. We have had the opportunity to really validate the behavior of the module and to provide the results of the improvements achieved in three different industrial scenarios in which the software was installed: large-scale retail industry, waste collection and maritime transportation. In each scenario our algorithm has proven to be very efficient.

2 THE REAL LIFE PROBLEM

2.1 CVRPTW

The Capacitated Vehicle Routing Problem with Time Windows (CVRPTW) is a well known strongly *NP*-hard problem, and it is a generalization of the Capacitated Vehicle Routing Problem (CVRP). The CVRP consists of finding a collection of simple tours of minimum cost in a connected digraph starting from and ending to a common depot, such that each customer (i.e., a node of the digraph) is visited exactly by a tour, and the sum of the demands of the customers visited by a tour does not exceed the vehicle capacity. In general the objective is to find the minimum number k of tours; a secondary objective is often either to minimize the total travelled distance or the duration of the tours. In the CVRPTW, service at each customer must start within an associated time window and the vehicle must remain at the customer location during service. If a vehicle arrives before the customer is ready to begin the service, it waits.

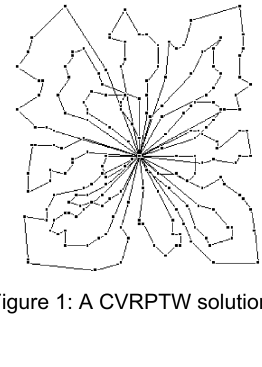


Figure 1: A CVRPTW solution.

The CVRPTW is one of the most studied variations of the VRP and recent surveys can be found in [1-5].

2.2 Practical constraints

A transportation order is specified by a customer location which has to be visited in a particular sequence by the same vehicle. The vehicles can have different capacities, as well as different travel times and travel costs between locations, and other compatibility characteristics which can constrain the possibility of assigning the transportation requests to certain vehicles. Instead of assuming that each vehicle becomes available at a central depot, each vehicle is given a start location where it becomes available at a specific time. All locations have to be visited within time windows different by type of transportation order.

A tour of a vehicle is a journey starting at the vehicles start location and ending at its final location, passing all other locations the vehicle has to visit in the correct sequence. For each tour we have additional constraints:

- Maximum number of deliveries and picks;
- Maximum time between subsequent of deliveries or picks;
- Maximum wait time before begin of service;
- Vehicle reuse for more tours but subject to restrictions to drivers' working hours.

Real-life problems often require rich models, in most of the literature on routing problems however, some simplifying assumptions are made. A discussion of real life vehicle routing can be found in [6-8].

2.3 Feasibility and unfeasibility

A tour is feasible if and only if for all request assigned to the tour compatibility constraints hold and at each location in the tour time window and capacity and additional restrictions hold. In our approach we consider both feasible and unfeasible solutions.

An unfeasible solution is penalized in proportion to the size of its constraint violations. The purpose of a penalty function formulation is to produce a representation of the problem that can be directly and naturally encoded as a genetic algorithm. Let *x* be a solution to a minimization constrained optimization problem. The objective of the problem is defined as:

$$\min_{x_i} z = Z(x) + P(x) \quad (1)$$

where $Z(x)$ is the objective function value produced by *x* and $P(x)$ is some total penalty associated with constraint violations at x.

Typically, the penalty imposed on an unfeasible solution will severely reduce the fitness of the solution in question, leading to quick elimination of the solution from the population. This may be undesirable, since unfeasible solutions may carry valuable information and may be useful in searching for optimal values. We use an evolutionary process evolving two subpopulations of solutions: the feasible subpopulation (consisting only of feasible solutions) and the unfeasible subpopulation (consisting only of unfeasible solutions). In the process, feasible solutions may produce unfeasible ones and unfeasible solutions may produce feasible ones.

It is evident in [9-11] that this technique has considerable merits.

At the end of the search process a set of no dominated solutions are found. The solving method is able to find a good solution with no violation and a set of efficient (no dominated) solutions in terms of travelling cost and customer service level.

3 AN EVOLUTIONARY APPROACH

3.1 Evolutionary algorithms

Genetic algorithm (GA) is an adaptive heuristic search method based on population genetics [12,13]. GA evolves a population of individuals encoded as chromosomes by creating new generations of offspring through an iterative process until some convergence criteria or conditions are met. The best chromosome generated is then decoded, providing the corresponding solution. At each iteration, the creation of a new generation of individuals involves primarily three major steps or phases: selection, recombination and mutation. The selection phase consists in choosing randomly two (parent) individuals from the population with a probability, in general, proportional to the fitness (goodness) of the individuals in order to emphasize genetic quality while maintaining genetic diversity. The recombination (i.e., reproduction or crossover) process makes use of genes of selected parents to produce offspring that will form part of the next generation. The mutation consists in randomly modifying gene(s) of a single chromosome (individual) at a time, to further explore the solution space and ensure or preserve genetic diversity. Both recombination and mutation operators are randomly applied with given probabilities.

Hybrid genetic algorithms combine the above scheme with heuristic methods for further improving solution quality.

There are also many applications of evolutionary techniques to the VRP and its variants. However, when applied alone, their success is limited. This led researchers to rely on hybrid approaches that combine the power of an evolutionary algorithm with the use of specific heuristics or to simplify the problem.

3.2 Description of our approach

Our hybrid genetic algorithm works on a population composed by a subpopulation of feasible solutions and a subpopulation of unfeasible solutions.

The schema of algorithm is as follow:

Definition: population *P* formed by *n* solutions and composed of $r \cdot n$ feasible solutions and $(1-r) \cdot n$ unfeasible solutions.

Initial Population: Fill the set *P* with solutions obtained by the randomized version of I1 Heuristics.

Selection: select solutions from *P* using a biased roulette wheel; a control mechanism is applied in order to maintain a prefixed rate *r* of feasible individuals on population *P*.

Crossover: The hybrid sequence based crossover (HSBX) is applied on two selected solutions.

Mutation Phase: Apply mutation operator *i* with a probability pm_i on a solution.

Return: the best solution.

The algorithm, starting from an initial population, progressively evolves the solutions by recombining feasible and unfeasible ones.

The considered selection operator is a fitness-proportional selector. Crossover and mutation operators are hybridized with insertion heuristics. In particular we use an extension of the SBX crossover operator.

Individuals are initially generated by a randomized version of the heuristic I1 proposed in [14]. The randomization is considered both in the seed computation to initialize a new tour and in the best feasible insertion position. Elitism strategy is implemented.

3.3 Crossover and mutation operators

We implemented a hybridisation of the Sequence-Based Crossover *SBX* described in [15] in which tours of parent individuals are merged. Given a pair (*X*, *Y*) of individuals, the crossover operator *HSBX*(*X*, *Y*) applied on the pair (*X*, *Y*) of solutions produces an offspring *X*′. We have also considered four different hybridised mutation operators. The description of these operators is in our previous work [16].

3.4 Fitness function

A solution *x* is specified by a pair <*z, p*> where *z* is the objective function value and *p* is the total penalty associated with constraint violations at *x*. Let x_1 and x_2 be two solutions, we have:

1. If $z_1 < z_2$ and $p_1 \le p_2$ then x_1 is better than x_2
2. If $z_1 > z_2$ and $p_1 \ge p_2$ then x_2 is better than x_1
3. If $z_1 < z_2$ and $p_1 > p_2$ then "which solution is the best?"
4. If $z_1 > z_2$ and $p_1 < p_2$ then "which solution is the best?"

For the case 3 and 4 we introduce the ranges d_1 and d_2, and the thresholds s_1 and s_2 considering the following rules to determine the best solution:

- If $|z_1 - z_2| \le d_1$: the solution with the lowest *p* value is the best;
- If $|z_1 - z_2| > d_1$ and $\max(p_1, p_2) \le s_1$: the solution with the lowest *z* value is the best;
- If $d_1 < |z_1 - z_2| \le d_2$ and $s_1 < \max(p_1, p_2) \le s_2$: the solution with the lowest *p* value is the best;
- If $|z_1 - z_2| > d_1$ and $\max(p_1, p_2) \le s_1$: the solution with the lowest *z* value is the best;
- If $\max(p_1, p_2) > s_2$: the solution with the lowest *p* value is the best.

For example in Figure 2 we have: the solution x_2 is better than x_1 , x_1 is better than x_3 and x_2 is better than x_3.

4 CONCLUSIONS

In this paper a vehicle routing problem with heterogeneous vehicle fleet with different capacity, multi-dimensional capacity constraints, order/vehicle, item/vehicle, and item/item compatibility constraints, different start and end locations for vehicles, and multiple time windows restrictions has been presented. The solving algorithm is a hybrid genetic algorithm with different hybrid crossover and mutation operators.

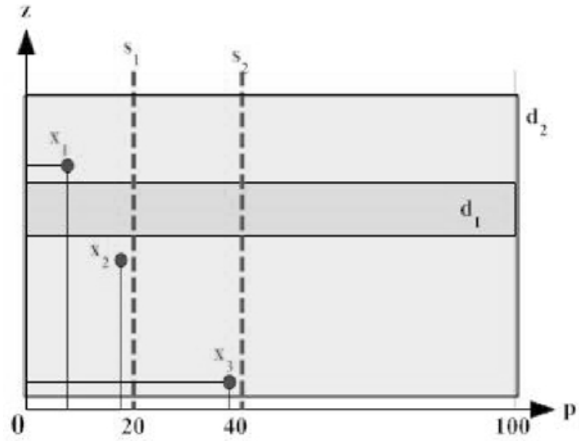


Figure 2: Example of selection of the best solution.

The hybrid genetic algorithm has been implemented in a software tool that allow easy configuration of objectives and constraints. It has been found that the proposed tool is effective and useful for more variants of Vehicle Routing Problem.

Computational experiments are performed on test cases derived from the real-life problem. They have shown that the algorithms perform well for problems with hundreds of vehicles and several hundreds of transportation requests. The combination of fast response times and the capability of handling the practical complexities allow the use of our algorithms in dynamic routing systems.

The solutions obtained by the proposed approach for various versions of the problem, in order to achieve effective use in real environments, are going to be presented in the future works.

5 REFERENCES

[1] Toth P. and D. Vigo (2000), *The Vehicle Routing Problem*, SIAM monographs on Discrete Mathematics and Applications.

[2] Bräysy O. and M. Gendreau (2003a), "Vehicle Routing Problem with Time Windows, Part I: Route Construction and Local Search Algorithms", Forthcoming in *Transportation Science*.

[3] Bräysy O. and M. Gendreau (2003b), "Vehicle Routing Problem with Time Windows, Part II: Metaheuristics", Forthcoming in *Transportation Science*.

[4] Laporte G., M. Gendreau, J.Y. Potvin and F. Semet (2000), "Classical and Modern heuristics for the vehicle routing problem", *International Transactions In Operational Research* 7, 285-300.

[5] Tan, K.C., L.H. Lee and K. Q. Zhu (2000), "Heuristic methods for vehicle routing problem with time windows", *Artificial Intelligence in Engineering* 15, 281-295.

[6] Psaraftis HN (1988). Dynamic vehicle routing problems. In B.L. Golden and A.A. Assad, editors, Vehicle routing: Methods and studies, pages 233-248. North-Holland Amsterdam.

[7] Psaraftis HN (1995). Dynamic vehicle routing: Status and prospects. Annals of Operations Research, 61:143-164.

[8] Savelsbergh MWP and Sol M (1998). DRIVE: dynamic routing of independent vehicles. Operations Research, 46:474-490.

[9] Kimbrough, S.O., Lu, M., Wood, D.H., Wu, D.J.: Exploring a two-market genetic algorithm. In Langdon, W.B., Cantù-Paz, E., et al., eds.: Proceedings of the Genetic and Evolutionary Computation Conference (GECCO 2002), San Francisco, CA, Morgan Kaufmann (2002) 415–21

[10] Kimbrough, S.O., Lu, M., Wood, D.H., Wu, D.J.: Exploring a two-population genetic algorithm. In Cantù-Paz, E., et al., eds.: Genetic and Evolutionary Computation (GECCO 2003). LNCS 2723, Berlin, Germany, Springer (2003) 1148–1159

[11] Kimbrough, S.O., Lu, M., Safavi, S.M.: Exploring a financial product model with atwo-population genetic algorithm. In: Proceedings of the 2004 Congress on Evolutionary Computation, Piscataway, NJ, IEEE Neural Network Society, IEEE Service Center (June 19–23, 2004) 855–862

[12] Goldberg, D.E. (1989), *Genetic Algorithms in Search, Optimization and Machine Learning*, Addison Wesley, Reading, MA.

[13] Holland, J.H. (1975), “Adaptation in Natural and Artificial Systems”, University of Michigan Press, Ann Arbor.

[14] Solomon, M.M. (1987), “Algorithm for the Vehicle Routing and Scheduling Problems with Time Window Constraints”, *Operation Research* 35, 254-265.

[15] Potvin, J.-Y. and S. Bengio (1996), “The Vehicle Routing Problem with Time Windows Part II: Genetic Search”, *INFORMS Journal on Computing* 8, 165−172.

[16] Confessore, G., Galiano, G., Stecca, G., 2006, General Purpose Application For Vehicle Scheduling And Routing In Complex Distribution Environments, 5th CIRP International Seminar on Intelligent Computation in Manufacturing Engineering, eds. Roberto Teti, pp 299-302, ISBN 88-95028-01-5.

Integrated Data Management in Factory Planning and Factory Operation. An Information Model and its Implementation

Tobias Weimer[1], Ralf Kapp[2], Peter Klemm[1] and Engelbert Westkämper[2]

[1] Institute for Control Engineering of Machine Tools and Manufacturing Units, Universität Stuttgart, Germany

[2] Institute of Industrial Manufacturing and Management, Universität Stuttgart, Germany

Abstract

Fast and smooth ramp-up of production and adaption of production capacity to the contemporary market induce new demands on the interaction of the two phases factory planning and factory operation as well as on their integrated IT-supported combination. This paper focuses on the IT requirements which result from the integration of factory planning and operation. Based on functional and on tool-oriented requirements, an information model is deduced which is qualified for the planning and for the operation of a factory. The elaborated information model represents the logical basis for an object-oriented data management as realised at the Universität Stuttgart within a research project in partnership with a medium-sized enterprise.

Keywords:

Information Modelling; Factory Operation; Digital Factory

1 INTRODUCTION

The following article deals with the modelling and maintenance of information. The information examined is that from the areas factory planning and operation.

Information used and generated in the field of factory operation can be discerned from planning information by several characteristics. Obviously planning information is virtual and connected to scenarios whereas operation information is related to elements in the real world, e.g. an order from a real customer, a error message from a real machine tool. Other characteristics are the amount of data underlying the information, the precision of information and the ratio between information regarding the structure of an enterprise and the information based on mass data.

Factory planning and operation are thematically close, since one would not be possible or needed without the other. In planning production, for example, the technical bases are created on which the operating phase will have to rely on for longer periods. The operating phase, on the other hand, continuously provokes re-planning of an existing production. In spite of this interrelationship, there is historically a strong disconnection which has only recently been softened.

The separation of factory operation and planning has its roots in the various tasks that form the basis of the phases and in the different people or contractors who are in their own positions working these tasks off. Up to now, the planning and the following operation were performed in a sequential way. Nowadays rising demands force enterprises to more frequent re-planning of factories and to an intensified parallel work on operating and planning factories.

In order to meet the growing demands of a global market in many areas a stronger integration of the two phases is being worked on. With this work, a faster and more efficient ramp-up of the production after a change in a product generation or technology should be achieved, as well as an adaptation of the capacities to fluctuation in product mixture of demand.

The integration does not only affect aspects of the organization and work flow but also aspects of IT. In regard to these integration efforts, e.g. the term 'Digital Engineering' may be mentioned here. In publications [1, 2] it is described how data from the planning phase of a production plant are used in the engineering phase and thus synergies can be utilized.

In this paper, the use of data – and of information – in various life cycle phases of a factory is examined and it is shown, which requirements are placed on the modelling of information. The requirements are based on several functions and tools of the two phases factory planning and operation.

The elaborated demands finally result in the generation of an information model which is suitable for planning and for operation of a factory. It will be in parts described here. Its implementation within an interdisciplinary research project of the Universität Stuttgart is in excerpts presented.

2 INFORMATION, DATA AND MODELS

In this paper the term ‚information' is used according to the following definition: ‚Information eliminates uncertainty. For the recipient it consists of date, meaning and practical value' [3]. This leads to the following definition: 'Data are information which is being processed in a computer (system-internal) and which is formulated for processing following precise regulations'. [4]

In order to be able to use the data of factory planning and operation in a cross-phased manner, their unified, standardized mapping in a data storing system is necessary. The basis of standardization is a data model. As shown in Figure 1, the data model is the end result of a series of modelling stages.

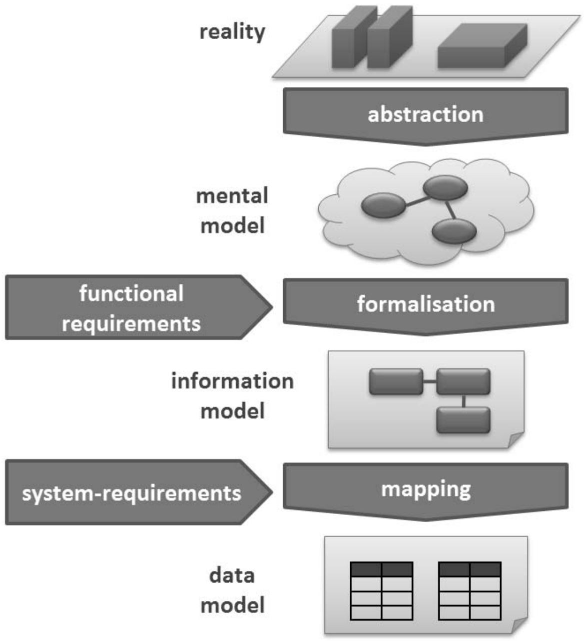


Figure 1: Modelling stages according to [5].

The modelling chain starts out with the Mental Model which is generated in the mind of the modeller by abstracting reality as well as by following conclusions. By formalizing the facts in consideration of the functional requirements of later use, an information model results by textual or graphic recording. By taking into account the requirements of the storage systems on the structure and by mapping of information in data bases, a data model is finally generated.

3 GENERAL REQUIREMENTS ON AN INFORMATION MODEL FOR FACTORY PLANNING AND PRODUCTION

An information model is situated, as shown in Figure 1, at the boundary between the real word and the system-internal world and thus also at the boundary between the object to model and the IT. It has to fulfil a number of requirements that are presented in the following.

3.1 Three types of information

With information modelling, three basic types of information can be distinguished.

Information about the modelled object

The most frequent type of information to be modelled is that from the real world. This can be either the weight of a product or an error message from a machine. The latter has no physical representation but as a part of the information to be modelled it is also assigned to the real world.

Information for software tools

In order to provide future safety, the information models are kept as tool-neutral as possible. But in many cases a complete tool neutrality is not possible. Many information-processing tools – i.e. the software for the direct support of the user in contrast to the data management tools – need application-specific information. Therefore, information as, e.g. a tool-specific well-defined product identifier has to be also includable in a tool-neutral information model.

Information for interfaces and transformations

The third type of information supports transformations. As seen in Figure 2, information can be available in differing forms and be exchanged by transformation.

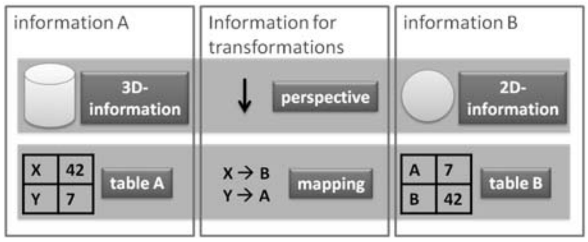


Figure 2: Different forms of corresponding information, connected by transformation information.

If different functions or tools have access to the same information, these data should only be filed once as a master and when needed be transformed.

- ➔ Requirement: Information in the model must be representable in regard to the modelling object, the function/tool and also in regard to transformation.

3.2 No redundancy

Redundancies are a source of inconsistency. To avoid redundancies, two types must be distinguished: on the one hand the evident redundancy when e.g. processing time is modelled as a property of the process as well as a property of the product.

On the other hand, it is more difficult to check for hidden redundancies. They can occur when – as described in chapter 3.1 – one type of information is provided with additional transformation-information and thus redundancy towards a third information form emerges.

- ➔ Requirement: Avoidance of evident and hidden redundancies.

3.3 Scalability in granularity

Strategic decisions, for example, need a more compressed information material whereas daily operative decisions are based on a very detailed knowledge of a few facts. Thus the application of a digital model of an enterprise includes information provision of differing granularity.

Such problems of scalability have to be considered during the design of the information model. They can be solved, depending on the specific case, by an intelligent use of hierarchies. In contrast to one basic hierarchy it is useful to utilize different dimensions of hierarchies. Thus, a resource in the financial hierarchy of cost centres is associated with another element in the geographical hierarchy of the architecture model.

The elements for modelling the hierarchy are assigned information in regard to the scaling and aggregation of information.

- ➔ Requirement: Information can be modelled in differing granularity.

3.4 Significance of existing standards

During the generation of information models, two kinds of standards have to be considered. On one hand those standards which are related to the modelling object, and on the other hand those standards that regard the tools and methods for modelling. Even though the latter have no semantic impact, the tools and methods used for creating an information model have a strong influence on the outcome.

An example for the standards first mentioned is the data exchange between ERP (Enterprise Resource Planning) and MES (Manufacturing Execution System) which is standardized in ISA95.

Important existing standards in the field of information modelling are UML (Unified Modelling Language) and XMI (XML Metadata Interchange). UML is due to its varying diagram units and its graphical representation very widespread. UML models can be mapped in XMI and as an XML data file they can be flexibly used.

- ➔ Requirement: Consideration of standards concerning the modelling object as well as concerning the modelling process.

3.5 Expandability

Already during the generation of an information model its later expansion has to be considered. Expansions here have their cause inside the enterprise as, for example, new IT support of a function or external causes like the passing of a new standard. Naturally, especially the later extension with future standards causes problems in regard to redundancy and uniqueness.

As known from fields like software development, object orientation is an accepted method to use advantages like enclosure of properties of objects and like inheritance. These features simplify a later expansion of a model [4].

- ➔ Requirement: Object-oriented modelling for easy expandability.

Expandability requires the manageability of a model. It involves clear organisation, logic structure and continuous documentation of the model.

The documentation has to be conducted all the way to the level of properties, in some cases even to the level of attributes. A clear structure of the information model can be achieved by the generation of class packages. Simple methods as, for example, graphical modelling tools and the pigmentation of elements in groups provide easy handling of the model.

- ➔ Requirement: Detailed documentation and clear structuring.

3.6 Automatable Model Check

An information model is a formalized set of rules and consequently it must contain any restriction which is imposed on the enterprise model based on it. So there are, for example, only zero to seven workdays per week permitted. The mapping of such a multiplicity between two elements is known from standards like UML [6].

But the requirements of enterprise modelling exceed the possibilities provided by UML. Particularly the exclusion of associations between certain objects is not possible. A restriction of this kind is necessary in the following case: different objects can be attached to the ends of an association due to inheritance – though it might not be desired in certain combinations.

A good example is the relationship ‘contains’ shown in Figure 6 between elements of the class ‘space’. Due to inheritance, this association may occur between all elements: A factory can include buildings, but a building cannot include any factories. Therefore, the association is assigned with a restriction that describes the forbidden combinations of association endings. When the enterprise model is generated, the restrictions and multiplicities of the information model can be checked automatically.

In addition, each property of a class must be assignable with attributes, in order to facilitate the utilization by various software tools of an enterprise and to allow model validations. One example could be the property ‘shipping weight’ which contains further attributes besides its own value. In this case, its dimension and unit are i.e. ‘mass’ and ‘kg’.

- ➔ Requirement: Restrictions, multiplicities and dimensions can be modelled.

3.7 Inheritance between objects in a library and objects in scenarios

Modelling of an enterprise depends on the use of a library containing objects, as described in [7]. Based on these preconfigured prototypes, objects can be instantiated. When using prototypes from a library to model an enterprise, the modeller is dependent on efficient methods to complete all properties of a newly instantiated object. Therefore libraries with efficient inheritance relationships are useful. Some of the properties should be inherited from the prototype others should be new defined.

For example: the property ‘producer’ of a class ‘robot’ is assigned with the attribute ‘inherit_value_from_library’, but the property ‘serial number’ not. Later, in the enterprise modelling, this means that the user while instantiating a robot from the object library only needs to punch in the serial number. The property ‘producer’, though, is inherited from a standard value of the class in the library.

- ➔ Requirement: Consideration of static character of properties for future inheritance from object libraries.

3.8 Scenarios

During the planning of an enterprise variations are generated. These variations will be called scenarios. Besides an enterprise model of the reality, an ideal scenario and differing planning scenarios are developed, assessed and compared.

- ➔ Requirement: Modelling of scenarios.

To catch all planning objects of one variation in the information model we define a so-called scenario object. Scenarios are the main objects of the planning phase and they contain all objects that constitute an enterprise, as, for example, resources, key figures, orders, organisation units, suppliers, relations etc.

With scenario objects, basic requirements have to be considered, as, for example, the properties of a scenario in regard to purpose, installer, date, status etc.

An information model used for factory planning and factory operation has to provide the possibility to support scenarios as well as inheritance relationships between scenarios. In Figure 3 the usage of inheritance is shown. New Scenarios can be created based on an existing scenario, in this example the real scenario, by coping or derivation. Here copying means creating decoupled scenarios without inheritance relationship, in contrast derivation creates coupled scenarios.

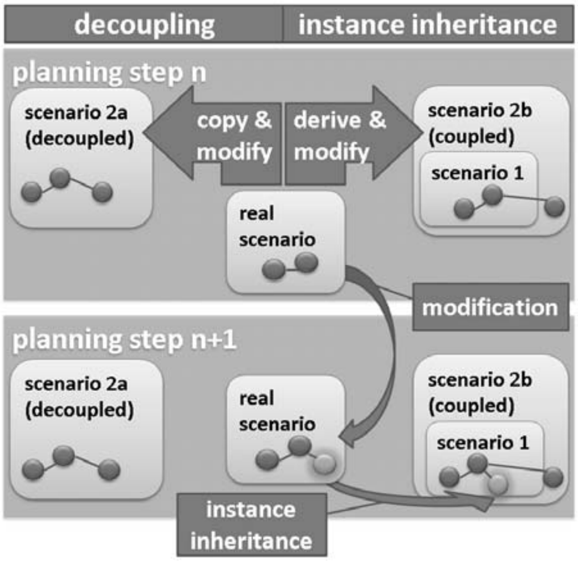


Figure 3: Difference between coupled and decoupled planning scenarios.

In the approach described by this paper, the real scenario is updated automatically by the data of factory operation and can serve as the basis for planning. When the reality changes, the real scenario is adapted. E.g. by new additional products. Through inheritance the changes also affect the coupled planning scenarios, according to Figure 3. Thus, the time-stretched work in planning scenarios always accounts reality.

- ➔ Requirement: Modelling of mass data under consideration of information models of data storage systems of operative units.

4 EXEMPLARY REQUIREMENTS FROM SPECIALIZED TOOLS

In an associated research project, several tools of factory operation and factory planning are to be connected by relying on a shared database. Figure 4 shows planning tools and operational tools linked to the 'Date Engine' which holds the model of an enterprise.

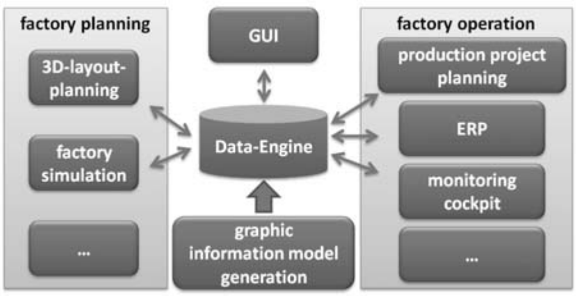


Figure 4: The Data Engine is accessed by a GUI (Graphical User Interface) as well as by tools for factory planning and factory operation.

4.1 Factory simulation

In the factory simulation the goal is to find answers to short and middle-term questions, concerning to the behaviour of the whole factory, which cannot be answered by static calculation. For example, whether for a period of some days sufficient production capacities are event discrete available depending on different (real or virtual) environment scenarios (different system loads (list of customer orders), delivery performances of the component suppliers, …)

In this approach an executable event discrete simulation model is automatically generated based on the data and structure of the digital enterprise model. Therefore the most important input data is the working plan, BOM, PPC (production planning and control) -methods and parameters, resource capacities and the system load. The user can think and model in logistic object and it is not required to have much simulation know-how.

In contrast to static calculations, the input data for a simulation has to be more detailed concerning time information. For example instead of a weekly cumulated amount of orders a list of single orders is needed. Those orders need to be equipped with details like date of order and assured date of delivery. Furthermore the available capacity of a resource has to be described in form of detailed shift models.

- ➔ Requirement: high level of details regarding time-dependent information

Current and historical data on a high level of detail can often be found in the IT-systems which support factory operation, e.g. ERP. To be able to answer short-term questions quickly with factory simulation, it is necessary to load operative data into the Data Engine with little effort.

- ➔ Requirement: automated import of data requires mapping of the information models of the operational IT-systems to the information model of the Data Engine .

The task of simulating a factory requires planning specific information like the start date and end date of a simulation run.

- ➔ Requirement: an object holding general simulation-specific information has to be attached to a scenario

Additional to the information available from the operative Systems information is needed to characterise in which way real objects are represented in the simulation model. To achieve maximum flexibility and usability a generic simulation specification was chosen. By the automated model generation a so-called model generator combines components from a building block library to a complete executable simulation model (cp. figure 5). Which available simulation building block is used for a particular logistic object has to be defined in the enterprise model. Standard building blocks and individual customized building blocks can be chosen. With this approach it's possible to model real objects with the required level of detail.

- ➔ Requirement: Posibility to relate real objects in the enterprise model to different types of simulation building blocks and parameterise them type specific.
- ➔ Requirement: new customized ‚building blocks' with specific parameterisation need to be integrated into the information model with little effort.

Normally a factory does not start operating empty. So at the begin of the simulation run there is the need to set an initial situation. All current stock levels, machine states, released production orders, position and machining progress of every lot have to be set in the simulation model. Thus ramp-up effects in the simulation run can be avoided und the simulation is able to give detailed forecasts of the near future.

- ➔ Requirement: PDC data needs to be represented in information model.

As well PDC data are an output of a simulation run. In the data engine they are assigned to the simulation-scenario and key performance indicators are calculated according to real performance indicators, as well defined in the enterprise model. So these figures are presented in the cockpit analogue to real key performance indicators. This enables operational staff easily to understand the results of simulation experiments in a well-known environment

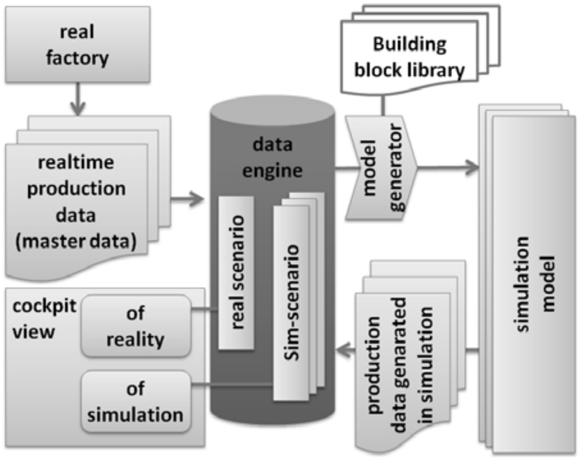


Figure 5: Scheme of the generic model creation.

For new questions, far away from operational exigency, there is the need for a quick and easy definition of new KPIs.

- ➔ Requirement: Handy configuration of new simulation specific KPIs analogue to operative used KPIs.

4.2 Factory layout planning

A central function within the planning of a factory focuses on the layout of the shop floor. It is done after the planning of the production-processes and after rough decisions regarding resources. Layout planning aims towards an optimal layout of the production with objectives like providing utilities to machine tools or placing buffers for products. Input from the factory simulation (e.g. material flow intensity or floor space required for work in process) is used in layout planning to back decisions and it is also possible to visualize it within a layout tool. A good example is the Sankey Diagram which can visualize material flow in a planned production by drawing simulation results onto the corresponding paths in the layout software.

Software tools used for layout planning are divided into 2D and 3D graphic tools. The exemplary tool chosen here is iPlant [8]. It is a 3D tool which provides a participative user interface for easy layout planning.

To connect iPlant to a central enterprise model, the model has to be able to hold 3D data in the form of VRML files. Therefore the information model needs to have the capability of properties as binary large objects (blobs).

Since the model of an enterprise is based on objects like machine tools, each carrying a VRML, the layout tool needs information about the relative position of each VRML. Besides properties describing the position, there have to be properties for orientation of the VRML.. The assignment of a VRML, a relative position and orientation per atomic objects is needed to use groups of machines in different ways of deriving or copying them to several buildings or scenarios. Another interrelated property contains the scale of the VRML-3D-data, since it is not sure whether the used scale while creating it was meters or millimetres.

Additional information to be modelled concerns transfor-mation of 3D data to 2D views. This information should consist of a description of a slice and a perspective. Thus the information defines the transformation end allows 2D views of complex layout data.

- ➔ Requirement: Properties like orientation, position, 3D blobs and transformation information for 2D view of 3D data need to be integrated into the information model.

4.3 Cockpit Portal

In most enterprises daily and strategic decisions are based on information about production, customer orders, etc. This information can be presented as Key Performance Indicators (KPI), arranged on Scorecards (SC). The concept of arranging several SCs, each containing KPIs, to show all relevant information about an enterprise in one software tool is called ‘portal’ in this article.

The KPIs presented to decision makers are part of a bigger picture: a problem in a producing company (for example high rate of rejections) results in a certain pattern of KPIs reaching their thresholds and blinking red for instance. The problem indicated by the pattern leads to tasks which have to be accomplished to solve the problem. Tasks are connected to responsible persons. This chain of problem-pattern-tasks-person should be taken into account when modelling KPIs and SCs. The result in the information model should be an object for each part of the work flow.

A huge benefit of using KPIs in an enterprise model which connects factory operation and factory planning is clearness. If daily KPIs are used to demonstrate planning results, the process of reviewing results and scenarios becomes much more ergonomic and efficient.

- ➔ Requirement: KPIs and SCs need to be modelled in regard of underlying problems, tasks and responsible persons.

5 IMPLEMENTATION

The deduced information model structures classes and relations into the following packages:

products	processes	resources
organisation	partners	information
key figures	disturbances	capacity
active objects	technical integration	

In Figure 6 a simplified excerpt of the information model is shown. Here a cut-out of the packet ‘organisation’ can be seen. The four organisation structures which can be assigned with a resource are easy to see. Next to a cost centre and a production segment a resource can be assigned to a ‘room’ which ultimately covers the aspects financial, organisational and physical organisation

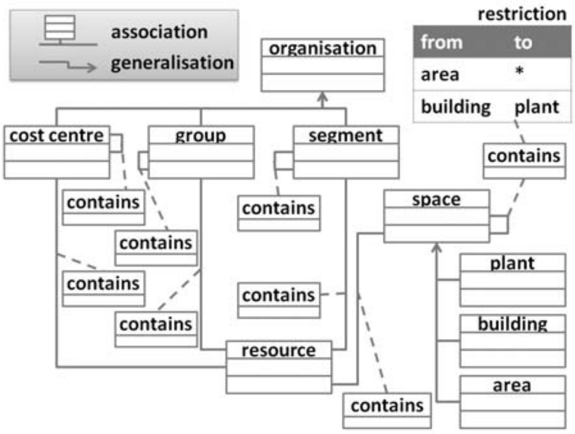


Figure 6: Simplified cut-out of the information model. The element ‚resource' is assigned to four organisational elements.

In order to allow for the disciplinary organisation of the employees, a fourth organisational structure 'group' was added.

The fact that all organisational objects may only contain organisational objects of the same type, makes the hierarchical structure possible. In connection with the classes 'factory', 'building' and 'area' all derived from the class 'space' the problem results that a building can also contain factories. To avoid this, a restriction was assigned to the association 'contains'. It includes all forbidden combinations of objects at the endings of the association.

6 USER INTERFACE – USING THE DATA ENGINE

For the work with the Data Engine adapters for various applications have been generated. The modeling of an enterprise first is done by means of a graphical user interface that has access to the Data Engine via network. A screenshot in Figure [7] demonstrates the partitioning of the interface. In the left window inheritance trees of the modelled scenarios are displayed. Decoupled scenarios have red coloured icons, coupled scenarios green ones. After selecting a scenario the assigned objects are displayed in the middle area and their attributes can be edited in the right one. To handle mass data queries can be defined in the middle area to filter the displayed objects.

7 PROSPECTS

The objective of the research project of the Universität Stuttgart is a practical, prototypical software for the integrated deploy of data in factory planning and operation. As could be seen in the course of the work, the basic information model is a central component in such an interdisciplinary project.

The work also pointed out that the vision of one singular standardized information model for all of the manufacturing industry lies momentarily still in the future. It cannot be concluded ultimately, whether such a complete model can ever be realized in view of the immense differences of the production branches, the speedy developments in the IT sector and the multitude of case-specific requirements.

8 ACKNOWLEDGMENTS

Research was carried out with the support of the Transfer Unit 059 'Transformability in Multi Variant Serial Production' at the Universität Stuttgart. The studies are sponsored by the Deutsche Forschungsgemeinschaft (DFG).

9 REFERENCES

[1] Sackerer, M., 2005, From digital factory to real production. PS/IPC/Drives, Elektrische Automatisierung - Systeme und Komponenten, Fair & Congress, Pages 310-319, VDE-Verlag

[2] Deussler, S., Schloegl, W., 2006, Integrated usage of standard-MES-solutions - configuration of the production control system with digital factory data. Automatisierungstechnische Praxis – atp, Volume 48, Issue 10, Pages74-82, Oldenbourg Industrieverlag

[3] Lyre, H., 2002, Informationstheorie. Eine philosophisch-naturwissenschaftliche Einführung. München, Fink

[4] Schneider, U., Werner, D., 2007, Taschenbuch der Informatik. 6., neu bearb. Aufl., München, Wien, Carl Hanser Verlag

[5] Klemm, P., 2007, Vorlesung: Softwaretechnik für Prozessplanung und Leitsysteme II, Universität Stuttgart

[6] Webpage of Unified Modeling Language, 2008, http://www.uml.org/

[7] Kapp, R., Weimer, T., Westkämper, E., 2006, Factory cockpit: A factory simulator for the entire factory life cycle, 39th CIRP International Seminar on Manufacturing Systems, University of Ljubljana, Ljubljana, SI, June 7-9, 2006

[8] Webpage of CAD-Tool iPlant, 2008, http://www.i-plant.de/

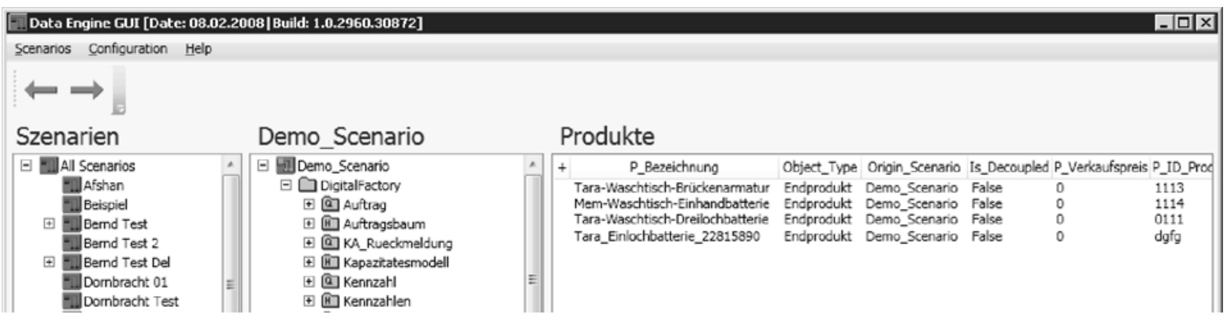


Figure 7: The graphical interface provides many modelling possibilities while designing a scenario.

Future Challenges in Process/Resource Planning Due to Increasing (Product) Variants

Lars Weyand, Helmut Bley
Institute of Production Engineering / CAM, Saarland University

Abstract
Factories that are flexible in production and cost-efficient are needed. But obviously, it is not so easy to plan/realise such factories – the complexity of the task increases with every additional (product) variant to be considered. All fields of the production engineering process are affected by the above-mentioned complexity problem. But, this paper only deals with shortcomings/challenges within the first steps of production planning. Future planning experts need suitable concepts to perform their tasks in an easier, faster and more precise manner. The transparency within production planning has to be improved in order to simplify decision making.

Keywords:
Process/Resource Planning, Variants, Decision Making

1 INITIAL SITUATION

Competition among manufacturing companies becomes harder and harder. Stagnating sales figures as well as price deductions are obvious signs for this.

Especially, the automotive industry is confronted with such a situation at the moment [1], [2]. New competitors are coming into the European market and they offer comparatively cheap products. As a consequence, the pressure on European car manufacturers increases.

In order to defend their market share and to avoid a direct competition on the low-price level, European car manufacturers try to focus more on special customer wishes. This trend has already started (cp. with figure 1), but all in all, the European car manufacturers are still at the beginning of an individualisation process.

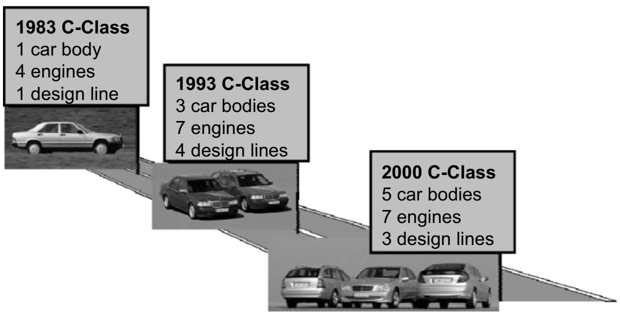


Figure 1: Evolution of the variants of the Daimler C-Class in accordance to [3].

Of course, customers have nowadays the possibility to choose among different car bodies, engines and design lines, but – in order to be honest – the choice is still limited. After a general decision about a car body and an engine, the customer usually has only some liberties in equipment combinations. Perhaps, this will be different in the future.

2 SUITABILITY OF THE EXISTING PRODUCTION ENGINEERING PROCESS

Individualised products consist of special parts and every variant of a part has to be considered within the production engineering process. Usually, the more different parts exist, the more the design complexity, the planning complexity as well as the manufacturing complexity increases.

All fields of the production engineering process are affected by the above-mentioned complexity problem caused by a huge number of (product) variants. But, the preparation of the different fields for the future variant task is different. Figure 2 shows the nowadays unbalanced situation.

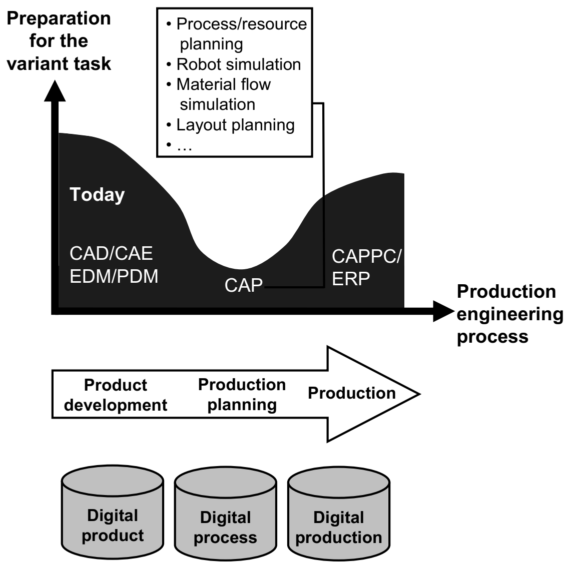


Figure 2: Preparation of the production engineering process for the variant task.

The production engineering process can be subdivided into three different fields: the field of product development, the field of production planning – in this case, we especially discuss process/resource planning – and the field of real production.

From our point of view, the fields of product development and production itself already seem to be well prepared for the future variant task – there are real variant management concepts as well as powerful IT-tools available in those areas. But in the field of production planning, the situation is different. In production planning, there are only isolated IT tools in use [4], [5]. Furthermore, a lot of activities are still done manually (without IT support) by the planning experts. The planning results still strongly rely on the experience of the engineers and integrated variant management concepts are up to now not available.

3 CLASSIFICATION OF VARIANTS

In general, it has to be differentiated between product, process and resource variants:

Product Variants

According to [6], product variants are objects of similar shape and/or function, usually with a high portion of identical parts.

Process Variants

Different products or product variants usually lead to process variants [7]. For example, it is maybe cost-efficient to join a product variant A that consists of a special part α based on a gluing process, but in respect of a product variant B that consists of a special part β it is maybe necessary to use another process – for example a welding process because of stiffness requirements.

Resource Variants

Resource variants often occur because of product/process variants [7]. In connexion with the above-mentioned example, it could be sensible to employ a worker for the gluing process of the part α (product variant A), but for the welding process it could make sense to use a robot for the joining of the part β (product variant B).

Up to now, it is only mentioned that product variants can lead to process/resource variants. But, there can be other reasons for process/resource variants, too. Marketing or personal decisions can also cause process/resource variants even if they are not really necessary. However, the circumstance that product variants can cause process/resource variants is in our case important.

In respect of processes, an additional differentiation between processes which add value and processes which do not add value is possible. In some cases, the ratio of value adding processes and no value adding processes can be seen as an indicator for a good planning. However, the above-mentioned differentiation shows that not only manufacturing processes have to be regarded within process/resource planning, also no value adding processes – logistic processes, verification processes and so on – have to be considered.

4 CURRENT SHORTCOMINGS AND CHALLENGES WITHIN PROCESS/RESOURCE PLANNING

In early phases of process/resource planning, a lot of work is usually done manually by planning experts. First process/ resource concepts are often generated based on manual drafts – digital tools are generally rarely used in early planning phases. Afterwards, the manual results are stored in a digital planning tool in order to be able to do subsequent time and cost analyses based on digital methods. Figure 3 shows how "documenting" looks like in a digital planning tool.

"Documenting" takes place based on relations to be set among elements of the product structure and their corresponding processes and resources. Thus, a non-ambiguous allocation of processes and resources to specific elements of the product structure is theoretically possible – a boundary condition for subsequent time and cost analyses.

The above-mentioned workflow is suboptimal. It is unfavourable to store only the results of first planning activities in a digital planning tool. The way of solving the planning task is

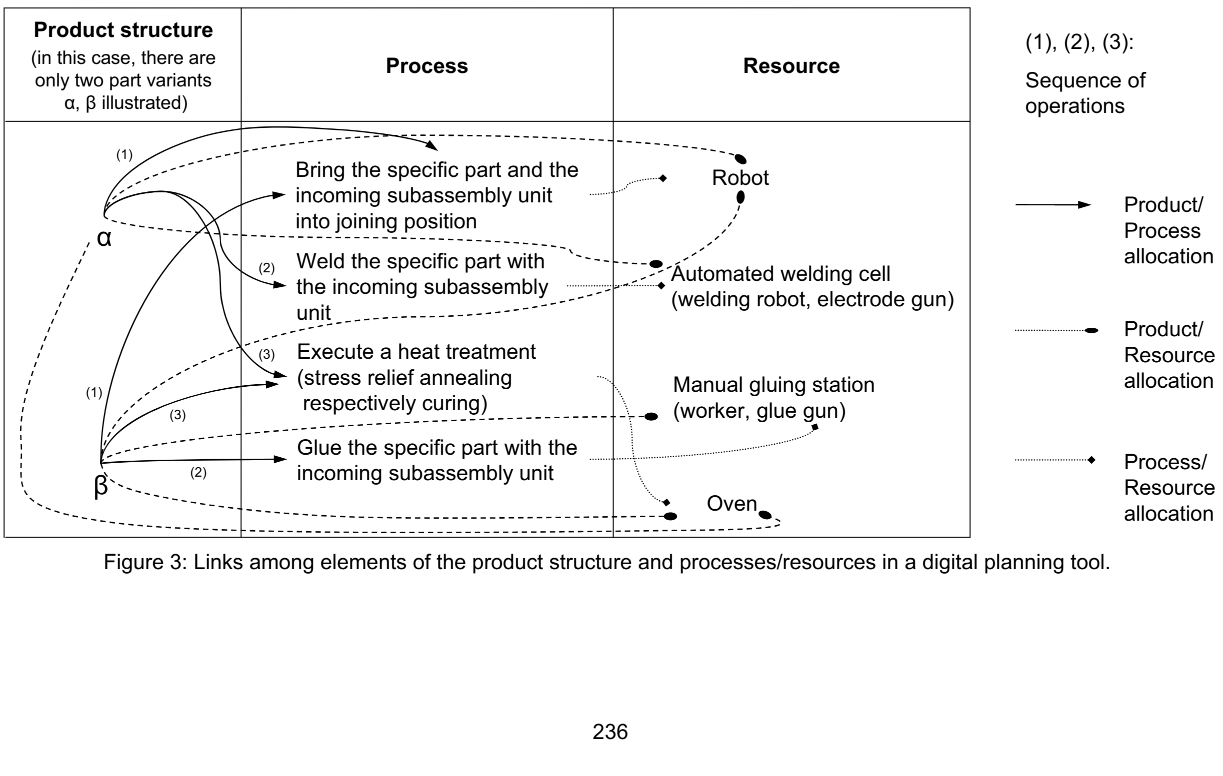


Figure 3: Links among elements of the product structure and processes/resources in a digital planning tool.

also important and has to be documented, too. Furthermore, it should not be forgotten that such an early, manual planning (without digital support) is a time consuming matter and that it is usually more prone to error. From this point of view, it seems to make sense to use digital planning tools from the first step of process/resource planning. In doing so, important ideas of involved engineers can be stored, planning modifications can be done easier and interim results can be distributed among engineers without effort.

At this point, it has to be mentioned that there is already a possibility to use digital planning tools from the beginning of process/resource planning. Nowadays, planning experts have the chance to work with so-called digital process/resource graphs. The graphs allow a graphical step by step solution of the planning task. But up to now, this solution is not accepted in industrial practice. The graphs become too fast too ambiguous. Figure 4 shows the main disadvantages of the available standard solutions (left side) and moreover an improved approach of the Institute of Production Engineering / CAM (right side).

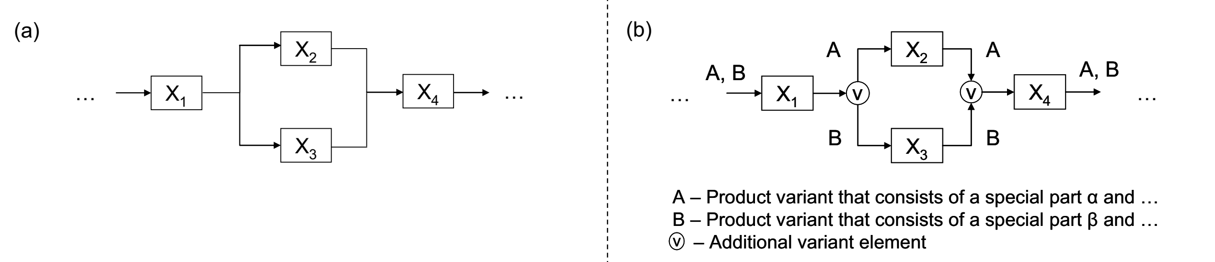
Ambiguous graphs represent a risk in the production engineering process. Mistakes in planning can occur because of the fact that involved engineers misunderstand the graphs. The final effects of occurring planning mistakes depend on their heaviness and on the point of time of their identification. The factor time plays a crucial role in this context because – in general – the later a mistake in planning is detected, the more money is needed for corrective activities. In worst case, it could happen that the mistake in planning is not found until the ramp-up phase of hardware components. In this case, it is possible that primary favoured product variants can not be produced. According to that, the planning quality is directly coupled with the production and product quality. The planning quality influences the launch time of the products as well as their production costs. From this point of view, it makes sense to view the process/resource planning as one of the most important phases in the production engineering process. Moreover, it should not be forgotten that correct decisions about variants to be produced and variants not to be produced can only be made based on a high-quality process/ resource planning with high transparency.

Up to now, the above-mentioned transparency is missing in digital planning tools. In general, the data bases already consist lots of data/information that are necessary for decisions, but unfortunately, it is not so easy to separate important data as well as to understand the stored information. Usually, only one engineer – the engineer who has stored the data/information in the data base – really knows the coherences why special processes and resources are necessary. In respect of a distributed planning procedure with lots of engineers involved, the missing transparency seems to be unacceptable or at least unfavourable.

5 NEW PROCESS/RESOURCE PLANNING APPROACH

Bley et al. [7], [8], [9], [10], [11] have already made a lot of proposals to improve the transparency within process/ resource planning. In order to come to a non-ambiguous process/resource graph, the best solution seems to be to integrate additional elements – so-called variant elements – in the graphs. The additional elements will improve the readability of the graphs and will allow other involved engineers to understand where and why special processes/resources are necessary. Figure 4 (b) shows the advantages of such an approach in comparison to available standard solutions.

On the one hand, there is no doubt that the introduction of



(a) and (b)

Process graph:	Resource graph:
X_1= Bring the specific part and the incoming subassembly unit into joining position	X_1= Robot
X_2= Weld the specific part with the incoming subassembly unit	X_2= Automated welding cell (welding robot, electrode gun)
X_3= Glue the specific part with the incoming subassembly unit	X_3= Manual gluing station (worker, glue gun)
X_4= Execute a heat treatment (stress relief annealing respectively curing)	X_4= Oven

Disadvantages of the current digital process/resource graphs:

- The graphs usually only consist of boxes and edges
- The graphs become very soon very complex and ambiguous
- A differentiation between simultaneous and alternate operations is not possible

Advantages of the enhanced digital process/resource graph:

- The graphs are non-ambiguous
- Adequate visual functionalities to determine which variant is taking which way in the graph are available
- A differentiation between simultaneous and alternate operations is possible

Figure 4: Standard approach in respect of process/resource graphs (a) and improved approach (b).

additional variant elements is a step in the right direction for increasing the transparency in current available process/ resource graphs. But on the other hand, there is also no doubt that additional improvements will be necessary in order to reach the goal of a future user-friendly, high-quality digital planning. Special filter operations as well as digital methods for automated identification of critical variants have to be developed. Furthermore, additional research activities have to be started that try to answer the question how precedence conditions can be stored and visualised in a digital environment. Precedence conditions have to be considered because of their impact on the assembly sequence. Moreover, a detailed, systematic planning workflow has to be worked out in order to ensure that the new approach can be used in practice. First ideas for a general planning workflow are presented in the following.

Planning experts do not only get the product structure as input for their planning task, they also get additional information from the marketing department. Predicted sales figures as well as car configurations to be checked with respect to manufacturing costs are available. At this point, it has to be mentioned that the planning experts usually only check some car variants – they are not able to check all possible configurations. Moreover, it is important to know that the marketing specifications not necessarily deal with variants that can be produced.

However, in connection with the presented approach it seems sensible to do a step by step planning, beginning with the car variant having the highest predicted sales figure. Afterwards, the next car configuration with the second highest predicted sales figure should be regarded and so on. This means, instead of starting a new planning procedure for every new product configuration, the already worked out process/ resource graph has to be checked with respect to its suitability for additional variants. In doing so, the original process/resource graph could be steadily supplemented with special processes and resources required for other additional configurations. Based on such a method, decisions in respect of additional special processes respectively resources, needed for special product variants, can be done systematically in the future. Furthermore, performed changes would be always obvious and clear for other engineers. Those points represent very important advantages that can be reached only in connection with the presented variant elements.

6 CONCLUSION

Manufacturing companies are nowadays confronted with a tricky situation. There are a lot of competitors on the market and the customers want to have products that fulfil their individual demands. In order to survive under such difficult conditions, European car manufacturers try to produce more different products as well as product variants. But, car variants cause trouble in the production engineering process. Especially, the field of production planning is up to now not prepared for the variant task. In first steps of production planning, a lot of work is still done manually (without IT-support) by the planning experts and real variant management concepts are missing.

Of course, modern digital tools already offer possibilities to support planning experts from the beginning of process/ resource planning, but the available digital process/resource graphs still have serious disadvantages. The main problem is that the graphs become too fast too ambiguous. Therefore, they can not be used in practice.

But in this paper, a promising approach for the creation of non-ambiguous graphs is presented. Based on additional variant elements, the readability of the graphs can be improved. Thus, misunderstandings among engineers can be avoided. Moreover, planning experts also have a possibility – at least in connection with the proposed workflow – to decide about special processes and resources, needed for special product variants, in a systematic manner.

7 ACKNOWLEDGMENTS

This research activity was partially supported by the research project “Flexible assembly processes for the Car of the Third Millennium (MyCar)” funded by the Commission of the European Union.

8 REFERENCES

[1] Becker, H., 2007, Auf Crashkurs – Automobilindustrie im globalen Verdrängungswettbewerb, 2. aktualisierte Auflage, ISBN: 978-3-540-37351-3 (in German).

[2] Papakostas N., Makris S., Alexopoulos K., Mavrikious D., Stournaras A., Chryssolouris G., 2006, Modern Automotive Assembly Technologies: Status and Outlook, 1st CIRP-INTERNATIONAL Seminar on Assembly Systems, pp. 39- 44.

[3] Bär, T., Haasis, S., 2003, Steps towards the Digital Factory, Proceedings of the 36th CIRP International Seminar on Manufacturing Systems, pp. 171-175.

[4] Franke, C., 2003, Feature-basierte Prozessketten-planung in der Montage als Basis für die Integration von Simulationswerkzeugen in der Digitalen Fabrik, PhD thesis at Universität des Saarlandes (in German).

[5] Bley H., Zenner C., Bossmann M., 2005, Advances towards an integrated product and production development process, Proceedings of 2. German - Israeli Symposium for Design and Manufacture - Advances in Methods and Systems for Products and Processes, pp. 129-137.

[6] DIN 199, Teil 2, 1977, Begriffe im Zeichnungs- und Stücklistenwesen (in German).

[7] Bley, H., Franke, C., Zenner, C., 2003, Integrated Data Management and Variant Management – Milestones on the Way to the Digital Factory, Production Engineering, Volume X/1, Annals of the German Academic Society for Production Engineering, pp. 105-108.

[8] Bley, H., Franke, C., Zenner, C., 2004, The User as the Main Focus for the Actual Realization of the Digital Factory, International Conference on Competitive Manufacturing, pp. 337-342.

[9] Bley, H., Zenner, C., 2005, Feature-based Planning of Reconfigurable Manufacturing Systems by a Variant Management Approach, 3rd International CIRP Conference on Reconfigurable Manufacturing.

[10] Bley, H., Zenner, C., 2005, Handling of Process and Resource Variants in the Digital Factory CIRP Journal of Manufacturing Systems, Vol. 34, No. 2, pp. 187-194.

[11] Bley, H., Zenner, C., 2006, Variant-oriented Assembly Planning, Annals of the CIRP Vol. 55/1, pp. 23-28, 20.

Human Aspects in Manufacturing Systems

Modernisation of Industrial Engineering. Enhanced Participation of Employees

Detlef Gerst[1]

[1] Faculty of Mechanical Engineering, Institute of Production Systems and Logistics, Leibniz University Hannover, Germany

Abstract

This paper addresses the opportunities and barriers of and prerequisites for the participatory design approach employed in industrial engineering. It discusses the results of a research project carried out at several sites of a German car manufacturer. With the intention of involving the manufacturing workforce in the process of setting their performance targets, the management of this car manufacturer reorganised the basic principles of its time and capacity planning. The paper argues that, on the basis of a so called High-Performance Work System, this concept improves and increases the manufacturing workforce's commitment, motivation and competence.

Keywords:

Industrial engineering; Participation; Production system

1 INTRODUCTION

Traditional corporate industrial engineering (IE) apparently no longer meets the demands of companies today. The reason for this is that increasing demands for more flexibility and higher cost pressure are making it necessary to involve employees in the planning, controlling and optimising of work processes. This stands in direct opposition to the primarily centralised approach of IE which is widely practiced. The possibility of extending worker participation to IE activities therefore raises the following question: How can this be achieved without limiting IE's vital functions in the areas of controlling and process-optimisation? The following is an attempt to answer this question and is based on research done at four sites of a German car manufacturer.

2 INDUSTRIAL ENGINEERING AND WORKER PARTICIPATION

2.1 Tasks of Industrial Engineering

IE concerns the planning, designing and controlling of a company's operational systems. Theoretically, it should unite economic goals with a humane form of work organisation. In this case, humane means taking the values and needs of workers seriously [1]. However, in operational practice, giving equal fundamental value to economic and humane criteria does not mix well with a one-sided orientation toward economic optimisation [2].

IE depends on the collection of data for its main tasks, and time management studies are the basis for defining production schedules, determining standard times and continually improving operating systems. How workers should perform their tasks is thereby determined by a centralised IE Department which is not organised to include worker participation. For their time studies and rationalisation, IE's experts do not consult the workers, but the works council (Betriebsrat). IE communicates the determined standard times to lower-level management, the foremen and forewomen (Meister/Meisterin), who monitor the achievement of goals (figure 1).

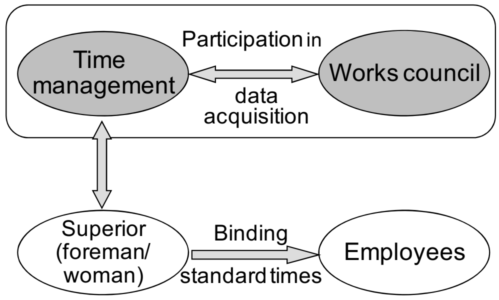


Figure 1: Traditional time management.

Not involving operational workers in setting standard times has its disadvantages. These include limiting the learning opportunities of workers and a loss of motivation and flexibility. Under these conditions, a company can hardly expect its production workers to be committed to applying their skills to kaizen activities or to take active roles in the company beyond the duties outlined in their contracts. Against this backdrop, the management of the car manufacturer in this study has recognised the need for modernisation and has begun to question the efficiency of those work processes which are only managed centrally.

Many articles have been published lately about new directions in IE. However, the authors of these articles are primarily interested in standardising work processes [3] [4], while they fail to consider the involvement of operational workers in IE activities. This is a point that deserves criticism because IE will not be able to successfully free itself from Taylorist forms of work organisation without the direct participation of workers.

2.2 Advantages of Participation

Early studies in the sociology of organisations already pointed out that Taylorist work systems could have unwanted side-effects. One of the gravest side-effects is called 'dysfunctional organizational learning' [5]. This means that standardised practices and their justification are more important to workers than targeting optimal production and meeting the changing demands of customers. Under these circumstances, Taylorist organisation leads to a loss of flexibility. Furthermore, a centralized planning and controlling of working processes requires a central department, which uses up resources without generating added value. Therefore, limiting the division and specialisation of labour and using partially autonomous units of organisation to meet complex requirements is beneficial [6].

In this respect, management has the task of convincing workers to voluntarily cooperate and show willingness [7]. One way of achieving this is to involve workers in those company decisions which do not necessarily need to be reached centrally. This is the core of the participatory approach in work organisation.

Since the 1990s, we have seen three main arguments that stress the economical benefits of participation:

The **organisation** argument reasoning that decentralised decisions and partially autonomous work groups provide relief for management and departments. Participation also enables a speedier reaction to interruptions in the production run.

The **resources and potential** argument that points out how participation is beneficial for building up intelligent and competent personnel. Improvements in human resources can broaden the company's range of possible actions and can become a factor that competitors cannot imitate.

The **corporate culture** argument stating that participation is an intelligent solution for mediating between the conflicting interests of management and workers. It allows trust to develop and encourages the 'extra-role behaviour' of workers.

According to these three arguments, participation is an integral part of a High-Performance Work System, which also serves as a benchmark for the German company in this study.

2.3 High-Performance Work Systems

The term High-Performance Work System (HPWS) refers to a type of production system which empirical studies have proven is extremely economically efficient [8]. The core components of a HPWS are autonomy in work processes, intensive communication, self-managing teams, worker participation in decisions, incentives and qualifications [9].

Several studies have proven that a HPWS encourages employees to become active beyond the duties stated in their contracts [10]. This 'extra-role behaviour' can be seen in the optimisation of operating systems and the flexible combination of knowledge and skills in the work process. Thus, a HPWS creates gains in productivity and flexibility while creating attractive working conditions.

High-performance work systems can be characterised by highly standardised work methods, but because workers participate in setting system standards, these also have a democratic character [11], enhancing employee motivation.

3 INTEGRATION OF TIME MANAGEMENT AND PARTICIPATION

In the following, I will present the results of an empirical study done at four sites belonging to a German car manufacturer. Under the banner '**Reorganisation of Time Management**', the company set itself the goal of involving its employees in the process of setting performance standards. The study is based on 8 on-site inspections, 32 interviews of experts, 25 group discussions and 545 written surveys of group members and lower management personnel (foremen/forewomen).

3.1 Challenges and Goals

The reorganisation process within the company was intended to create a new space for thoughts and actions in order to mediate between the interests of operational personnel and management. This mediation was also the special challenge of this labour policy project. The operating plan targets which companies set always consist of reducing costs and increasing productivity, and these goals of rationalisation conflict with the interests of operational employees who want to maintain their jobs in the long-term, their ability to work and their employability. Therefore, labour policies must cultivate a process that satisfies the point of view of both production workers and management personnel. One possible solution is a **model of dialogue** in which lower management and their employees jointly establish performance standards and optimise accompanying operating systems in order to reach the company's operating targets. The participation of production personnel in the structuring of time management sets the following goals:

- The broader understanding of operating processes on the part of operational employees.
- The greater responsibility of operational employees in the company's operations.
- The establishment of a culture of trust between management and production personnel.
- The utilisation of operational employees' skills in the process of kaizen.
- The repositioning of IE as an in-house service provider.

3.2 Fundamental Approach and Operational Practice

Production employees and their immediate operational superiors play key roles in the reorganisation of time management (figure 2).

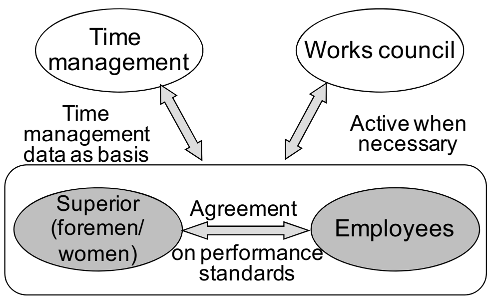


Figure 2: Reorganised time management.

Foremen/forewomen and employees jointly set goals by establishing a dialogue in which they decide how many employees are needed to complete a certain production volume. Company management therefore cannot one-sidedly set the performance standards. In return for a fixed performance standard, employees receive secure, invariable wages. Time management experts from IE collect data and assist in the process implementation. Each group chooses a representative to take part in the decision making process.

This person receives special training and mediates between his or her team colleagues and the foreman/forewoman.

This strategy was applied in the entire company through its company regulations, but it has been implemented in a variety of ways. In some plants, the dialogue for setting targets was teamed with a kaizen process and ambitious goals to save costs. However, this procedure was not typical for all the sites in question. In some locations, IE staff plays very different roles in the goal-setting process: IE can have a leading role, it can function as a service provider, or it does not play any role at all.

3.3 Ambivalent Operational Experiences

For each of the different operational groups, the operational experiences with this approach were ambivalent. A mere 41% of the foremen/forewomen and 22% of the group members surveyed believe the implementation of the new strategy is successful. This contrasts with 64% of foremen/forewomen and 58% of group members who believe the new regulation fundamentally makes sense. This implies that there are deficits in the implementation and organisation of the otherwise generally accepted approach. IE staff and managers also have mixed feelings.

The following problems with implementation were observed:

- Groups tend to argue with managers higher than foreman/forewoman about the performance standards to be agreed upon: Managers want to cut down on personnel in order to pressurise group members to improve, while group members strategise to maintain a certain amount of leeway and autonomy. The majority of the group representatives feel they are put under pressure to sign contracts for which they are later criticised by their colleagues.
- The relations between the groups and the IE Department can also be somewhat strained. Group members accuse IE of often being unwilling to cooperate, and IE believes the groups and foremen/forewomen are unwilling to contribute to goals of rationalisation.
- The result of this is that agreeing on goals takes a great deal of time or does not occur at all. Another point is that reorganisation does not lead to a culture of trust in many of the production sectors. Instead, group members feel reorganisation is purely a form of rationalisation and the promise of participation is not being kept.

These points of criticism represent the experiences of employees in roughly 60% of the reorganised areas of production. According to the surveys, the factors affecting the success or failure of this process do not include the following: the technical level of production, the difference between assembly and manufacture, and the level of qualification of employees.

4 CONDITIONS FOR SUCCESS

After assessing the interviews, it was determined that the conditions for a successful reorganisation of time management depend on the type of work organisation and on the quality of the goal-setting process.

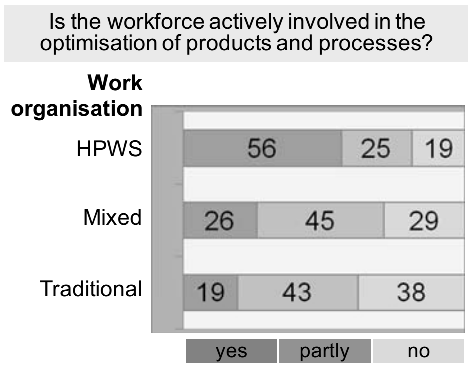
4.1 Type of Work Organisation

Ideally, a HPWS can be distinguished from a **traditional form of work organisation** (figure 3) based on Taylorism, which does not cultivate a dialogue with employees and has a foreman/forewoman who organises the operating business of the group as the lowest level of management. In a **HWPS,** partially autonomous work groups exist, employees and managers communicate according to a model of dialogue, and production sectors are managed by foremen/forewomen who are able to encourage groups to independently manage themselves and who are also more involved in planning and improvement than traditional foremen/forewomen.

criteria	HPWS	Mixed	Traditional
Work form	Partially autonomous group work	Contains elements of HPWS and traditional models	Taylorist work forms
Level of participation	Dialogue model		Antagonistic model
Type of superior	Upgraded foreman/ woman		Traditional foreman/ woman

Figure 3: Types of work organisation.

The type of work organisation has a measurable effect on the experiences of group members and their superiors. An excellent example of this is worker participation in improvement processes. Compared to a traditional form of work organisation, the reorganisation of time management in a HPWS results in a high level of employee participation in improvement measures (figure 4). This means that the type of group work, the level of participation and the type of foreman/forewoman are necessary conditions for achieving success.



n = 107 Foremen/women in 4 sites; data in %

Figure 4: Labour policies and kaizen participation.

From the point of view of IE and labour policy, another condition for success is the quality of goal-setting processes.

4.2 Quality of Goal-Setting Processes

The process of setting goals must be organised in a way which takes the conflicting interests of management and employees into account. However, it is unlikely that the company's operating targets can be smoothly coupled with employees' interests in the open discussions of the goal-setting process. In order to avoid unproductive power struggles between employees and management, a structured approach is vital (figure 5).

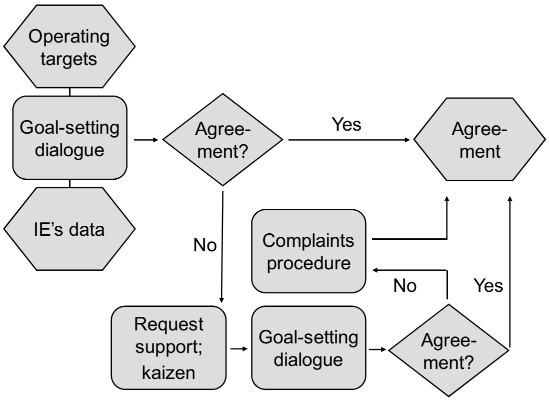


Figure 5: Ideal procedure model for setting goals.

The basis for lower management's and production personnel's ability to set goals together is time management data and the plant's operating plan targets. It is the responsibility of lower management to take these basic requirements into account in its dialogue with the workers and to therefore assume the role of leaders and mediators. The foremen and forewomen must initiate the process of setting goals while taking the workers' point of view into account. If no agreement can be reached, they initiate the next step. They can request assistance, for example from IE personnel, or they can initiate kaizen activities to increase the likelihood of achieving operating plan targets. If no agreement is achieved after a second dialogue has been initiated, they should initiate a **complaints procedure** as soon as possible. A committee with an equal number of representatives from the works council and management is responsible for such matters.

This process model offers ideal conditions for a functional dialogue by allowing conflicts to be expressed without the risk of hindering the agreement process. Having a dialogue also prevents workers from becoming overtaxed through unjustly high performance standards [12]. The company in this study planned to implement this process model in all of its sites, but in reality this was rarely done with consistency. The main reasons for this were overtaxed and underqualified lower management personnel and a lack of acceptance of the dialogue model on the part of middle management.

Where these obstacles were not present, the reorganised time management system proved very practical. It not only allows a high level of worker participation, it also encourages a sharper focus on the company's operational targets.

5 SUMMARY

In this paper, a German car manufacturer serves as an example for how including workers in the process of setting performance standards can modernise IE. However, reorganising time management can only lead to positive results under the following two conditions: The work system must comply with the basic principles of a High-Performance Work System, and the goal-setting process must guarantee that the company's operating targets are taken into account and the performance standards are agreed upon, even in cases of conflicting interests.

6 REFERENCES

[1] REFA-Verband für Arbeitsstudien und Betriebsorganisation (ed.) 1991, Planung und Steuerung. Teil 1. Reihe Methodenlehre der Betriebsorganisation, Carl Hanser, München.

[2] Zülch, G. 1997, Industrial Engineering, Luczak, H.; Volpert, W. (ed.), Handbuch Arbeitswissenschaft, Schäffer-Poeschel, Stuttgart: 183-186.

[3] Deuse, J.; Wischniewski, S.; Fischer, H. 2006, Renaissance des Industrial Engineering. Methoden für die Umsetzung Ganzheitlicher Produktionssysteme, wt Werkstattstechnik online, 96/1-2:57-60.

[4] Haller, E.; Schiller, E.; Abraham, S. 2005, Wandel der Arbeitsorganisation im globalen Wettbewerb am Beispiel der Pkw-Montage, Zülch, G. (ed.), Arbeitsorganisation in der globalen Wirtschaft. Shaker Verlag, Aachen: 48-65.

[5] Merton, R. K., 1940, Bureaucratic Structure and Personality, Social Forces, 18: 560-568.

[6] Baecker, D. 1999, Organisation als System, Suhrkamp, Frankfurt am Main.

[7] Chester Barnard, 1938, The Functions of the Executive, Cambridge, MA.

[8] Bailey, T. 1992, "Discretionary Effort" and the Organization of Work: Employee Participation and Work Reform since Hawthorne, Teachers College Institute for Education and the Economy Working Paper, New York.

[9] Thompson, M.; Heron, P. 2005, Management capability and high performance work organization, International Journal of Human Resource Management, 16/6: 1029-1048.

[10] Appelbaum, E.; Bailey, T; Berg, P.; Kalleberg, A, 2000, Manufacturing Advantage: Why High-Performance Work Systems Pay Off, Cornell University Press Ithaca, NY.

[11] Adler, P. 1995, "Democratic Taylorism": The Toyota Production System at NUMMI, in: Babson, S. (ed.): Lean Work. Empowerment and exploitation in the global auto industry. Wayne State University Press, Detroit, Michigan: 207-219.

[12] Gerst, D. (2006): Von der direkten Kontrolle zur indirekten Steuerung. Eine empirische Untersuchung der Arbeitsfolgen teilautonomer Gruppenarbeit. Rainer Hampp Verlag, München, Mering.

Towards Optimal Worker Assistance: Investigating Cognitive Processes in Manual Assembly

Christian Stoessel[1], Mathey Wiesbeck[2], Sonja Stork[1], Michael F. Zaeh[2] and Anna Schuboe[1]

[1] Experimental Psychology, Ludwig-Maximilians-Universität München, Munich, Germany

[2] Institute for Machine Tools and Industrial Management, Technische Universität München, Munich, Germany

Abstract

The integration of cognitive systems in currently humanly dominated, manual assembly environments is the core issue of the project ACIPE (Adaptive Cognitive Interaction in Production Environments). This paper presents a novel concept of workbench augmentation for adaptive human worker observation and guidance. An increase in worker performance is expected from a cognitive assistance system integrated into the workbench, which can support the worker via context sensitive, adaptively generated instructions at the right time, location and with appropriate content. In order to assist the worker adequately, a deeper understanding of mental workload and related cognitive processes and limitations during manual assembly is needed. The augmented workbench serves both as a research tool for detecting cognitive bottlenecks in manual assembly, as well as an implementation platform for the worker assistance system, leading to a more efficient manual assembly performance.

Keywords:

Manual assembly, worker assistance, human cognition, mental workload, augmented reality

1 INTRODUCTION

The integration of cognitive systems in currently humanly dominated, manual assembly environments is the core issue of the project ACIPE (Adaptive Cognitive Interaction in Production Environments). It is tightly integrated in the excellence cluster 'Cognition for Technical Systems' (CoTeSys), a large scale, long-term research initiative funded by the Deutsche Forschungsgemeinschaft (German Research Foundation, DFG). The cluster CoTeSys brings together Munich-based researchers from engineering, computer science, psychology and neurosciences to strive for the realization of cognitive technical systems. The cluster's research findings will be implemented in different demonstration platforms. Within this context, a crucial goal of ACIPE is to develop concepts and prototypical technical implementations, which will allow the realization of a truly context-sensitive guidance and assistance system for manual workplaces.

2 COGNITION IN PRODUCTION ENVIRONMENTS

Skilled human workers turn mechanical workshops into today's most flexible and widely applicable forms of production. However, this flexibility and generality comes at very high production costs, which restrict the use of mechanical workshops to building prototypes and a limited range of highly specialized and valuable products. The source of flexibility in this case can be easily identified: the cognitive capabilities of the humans that operate it. Humans can perceive their environment with multiple sensory organs, plan actions, learn and adapt behaviors and they can interact in multiple ways with their surroundings. Moreover, humans can do this robustly despite changing contexts and situations. The realization of comparable cognitive capabilities in technical systems bears an immense potential for the creation of industrial automation systems that are able to overcome today's boundaries. Conventional automation systems fail to account for the demand on flexible manufacturing of highly variant and customer individual products [1]. In order to reach a high degree of flexibility and efficiency, automated processes, such as guidance systems, have to be integrated with human workers [2]. Such a tight interaction between humans and machines becomes only possible by introducing intelligent interfaces that enable adaptive support during assembly. Adaptive in this context means the real-time integration of sensory information on the assembly progress, the production environment and the current state of the worker in order to generate a context-sensitive set of instructions and appropriately accommodate the information output to the worker. This knowledge will allow for interactive multimodal guidance and support of the worker in manual and semi-automatic assembly.

2.1 Understanding cognitive processes in manual assembly

When striving for an assistance system that can support the worker adequately, it is of crucial importance that the system delivers the right kind and amount of information at the right time and place so that the worker can receive and process it effortlessly. It is therefore not sufficient for a system to be aware of the environment and the current state of the product in the manufacturing process, but additionally it has to incorporate information on the cognitive processes involved during manual assembly. By cognitive processes we understand the entirety of human mental functions dedicated to information processing, such as perception, attention, memory, problem solving and action. Only on the basis of such knowledge it can be ensured that the assistance system chooses the right kind of support which does not overstrain the human worker. This is particularly important in those situations where high demands on the worker's cognitive capacities threaten an efficient and faultless completion of the workpiece. The research presented in this article focuses on the investigation of cognitive processes and respective cognitive bottlenecks during a manual assembly task. It provides a framework how these insights can be incorporated in the technical realization of an assistive workbench.

The information processing framework

In order to disentangle the cognitive processes involved in manual assembly, it is useful to adopt an information processing framework. Such a framework (e.g., Figure 1, left) is based on the notion that environmental stimuli, which are taken up by the human through the sensory organs, are processed in a roughly serial manner through a couple of defined processing stages, until finally a response is generated to act upon the environment. This basic taxonomy is useful in breaking up the cognitive processes involved in manual assembly into different substeps. Since in manual assembly usually some instruction detail has to be detected, identified, translated into an appropriate motor response and this response has to be executed correctly, this framework seems especially appropriate to structure the involved cognitive processes. An example of how this framework can be applied to a manual assembly task is provided in Figure 1 (right). Such a framework, although oversimplified, is nevertheless useful to describe specific sources of human error: As human processing resources are limited, they have to be distributed or allocated to specific mental processes, and delays or errors in performance can arise if processing resources are insufficient or not allocated appropriately.

Selective Visual Attention

The authority deciding which items (features, objects…) will be processed at all, and which information is passed over to the next processing stage, is referred to as attention. Even though we can deploy attention covertly without moving the eyes to a certain region in the environment, we cannot attend to an unlimited number of objects at different spatial locations simultaneously. During manual assembly for example, we cannot attend to the workpiece and the instruction manual at the same time if they are spatially separated. There are competing theories in psychological literature on the exact role, locus and timing of attention, but they all agree on its selective character, its ultimately limited resources and the measurable time costs it takes to switch attention from one location or object to another. Attention can be directed reflexively, by external, exogenous cues, like a salient object in the environment, or voluntarily (endogenously), by instruction or intention. It is evident that an assistance system should guide the worker to attend to the right locations and objects, and avoid unnecessary shifts of attention. The worker can work more efficiently, if a strong exogenous cue guides her/his attention (for example a red piece among grey ones, or a highlighted box surrounded by nonhighlighted boxes) and if she/he knows where to look next for the next relevant piece of information.

Multiple Task Performance

For most tasks, demands on our attention become first obvious when we have to perform several tasks simultaneously: While driving and conducting a conversation at the same time is usually a simple task, conversation might come to a halt if the driving situation becomes difficult, and all available attention has to be allocated on the driving scene. Similarly, in an assembly context, the worker has to divide his attention between the instruction manual, the work piece, and

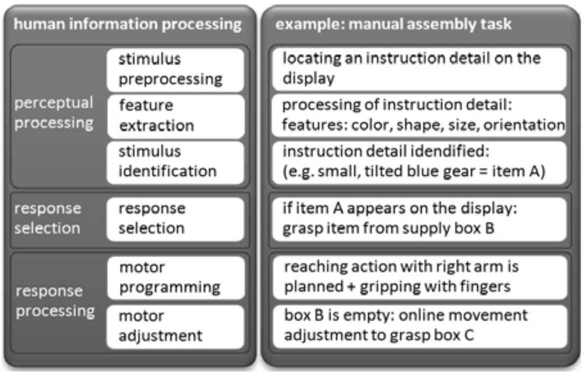


Figure 1: Left: Human information processing stages (adapted from Sanders, 1990, [3], p.37) Right: Mapping to corresponding processing stages within a manual assembly task

the factory environment. Humans have to allocate attention appropriately to aspects of the tasks, and in order to optimize performance, they have to find the right combination of attentional strategies (attending to different sources of information) and perceptual strategies (only extracting the necessary amount of information from a source) [3]. Efficient, coordinated performance depends on the ability to switch from one task (component) to another. Whenever the worker has to switch tasks, (e.g., when the next assembly step requires a new set of parts/tools) some cognitive processing time is needed to adapt. In order to establish an efficient assembly workflow, a switch of tasks should be reduced to the necessary minimum. In order to avoid task switching, information to the worker should be organized coherently, for example the sequence of assembly instructions and the spatial organization of the workplace should be in a logical and comprehensible manner.

Mental Workload

The previous section described the important role of attention in selecting information, and how it needs to be carefully distributed to objects, spatial location and tasks on the background of limited cognitive resources. Equally important from an applied perspective is to measure and evaluate the mental demands imposed by tasks and to describe how people cope with multiple demands on their attention in order to avoid performance break down and errors. The term **mental workload** has been used in applied psychology and human factors literature to describe information processing demands imposed by the performance of cognitive tasks [3]. Wickens suggested that different tasks draw on different mental resources with their specific capacity limitations, and only if two tasks need the same resource, decreased performance is observed [4]. He proposed a four-dimensional resource model described along the dimensions **input modalities** (visual - auditory), **processing codes** (spatial-verbal), **processing stages** (perception – central processing - responding) and **response modality** (manual-verbal). This model has become popular to describe mental workload, as it has been proven useful to represent tasks in a relatively simple framework and allows making predictions on concurrent performance of multiple tasks. In the context of a manual assembly operation, this framework could be applied, for example, to the concurrent tasks of mounting some parts and skipping through assembly instructions. If both operations have to be carried out simultaneously and involve the same response modality (e.g. both require manual intervention) the workload is higher due to increased demands in hand coordination, compared to a situation where the task of flipping pages in the instruction manual can be carried out concurrently through a verbal response. Therefore, task design should avoid cognitive bottlenecks within the same dimension.

3 TOWARDS A COGNITIVE WORKER ASSISTANCE SYSTEM FOR MANUAL ASSEMBLY

In order to reach the goal of providing an adequate support for the worker, progress has to be made on several levels.

On a theoretical level, the cognitive processes involved during a manual assembly task and processing bottlenecks need to be understood. The previous paragraph provided a framework of information processing stages during manual assembly. It stressed the importance of directing the worker's attention appropriately, to present the task coherently and with a minimum of attention switching, and, more generally, to design tasks based on these findings. Only if sufficient progress is made in understanding the cognitive demands during manual assembly, an assistance system can be effectively tailored to increase worker performance.

Secondly, on an implementation level, we investigate by which technology we can enable an assistance system to appropriately adjust to the worker. If the system should be able to adjust the worker support online, immediately, and in a context-sensitive way, then we are in need of a tracking and sensor technology that provides constant information on the environment, the state of the assembly process, and the actions or even intentions of the worker. A further technological challenge is the mode of information presentation to the worker. Especially if we deal with untrained workers or highly customized products, the worker has to rely strongly on the assembly instructions. The work instructions must describe concisely what needs to be done, in which succession, and with which tools and materials. Consequently, the instructions need to be carefully planned in order to minimize operator learning time, while at the same time being economical to prepare, reproduce, distribute and change [5]. While in many manual assembly workplaces today paper instructions manuals are still abundant and have clear advantages on the side of their production and reproduction, their disadvantages and limitations are also obvious: They are static in respect to the content that is displayed (e.g. no animation, context highlighting possible), their content cannot change dynamically with changing products and individual worker needs, physical size is fixed and takes up workspace, and manual assembly workflow has to be interrupted if pages need to be flipped manually. These limitations are contrary to the demands on a flexible assistance system, which should provide intuitively understandable assembly instructions and should be able to issue context-sensitive warnings or direct the worker's attention by exogenous cues, while at the same time minimally interfere with the assembly process. Regular display technology (e.g. mounting a TFT monitor to the workbench) can surmount already some of these limitations: e.g. several product instructions can be provided on the same physical instruction medium, animations or 3D inspections of a virtual workpiece are possible. However, it is still limited to a fixed location in space and needs to be attended by the worker, who consequently has still to perform cognitive mapping and search tasks to match the instructions with the real assembly scenario. **Augmented Reality (AR)** technology seems to be an appropriate candidate to deliver instruction information interactively at the right location in space and at the right time when needed.

3.1 Potential of Augmented Reality for manual assembly

Augmented Reality technology, such as head mounted displays (HMD) and tracking systems, has been employed in an impressive variety of applications (e.g., in architecture, medicine, entertainment, production or maintenance). AR has the potential to be successfully applied within an industrial working context (e.g. the assembly line), and to deliver updated, accurate, and useful information at the place where it is needed (e.g. a representation of the next product to be assembled), which will 'eventually lead to shorter production times, less training efforts, the reduction of errors, and finally to lower production costs' ([6], p. 284). Tang and colleagues have pointed out that AR technology has the potential of reducing head and eye movements, reducing the costs of attention switching and supporting spatial cognition and mental transformation [7]. However, successful working AR scenarios in manual assembly environments are rarely found. This discrepancy between being highly adequate from a theoretic point of view and the lack of successful implementation is explainable by the many technical problems that these systems are still facing. In the manual assembly domain, AR scenarios have been developed for example in car door assembly, cable wiring, mainboard assembly or toy settings. Common to these approaches was the use of a HMD, which was used to overlay the real world with additional, spatially coregistered pieces of information. To summarize the findings, AR technology has shown beneficial mostly only in cases when the tasks were sufficiently complex (e.g. [8]). It was most effective during the selection phase of assembly, where workers have to find the right piece and position where to put it, but not during manual execution of the task. In many cases, the technical limitations – strain during continuous wearing of HMDs, their limited resolution, contrast and field of view, problems with registration, unsatisfying speech based interaction – outnumbered the benefits [8,9].

The approach we took tried to combine the benefits of AR technology – providing additional information on top of the real world environment at the right time and location where it is needed – without having to face the drawbacks that the use of HMDs – co-registering techniques and speech based interaction – usually brings about.

4 WORKPLACE SETUP

In order to develop an adaptive interaction model to support the worker more efficiently, two lines of research were pursued in parallel: The first challenge was to develop a working environment that satisfies the ergonomic constraints for effortless manual assembly, and includes a well-adjusted combination of display and tracking technology. The worker can thereby be presented with individually tailored instructions at the time and place when and where she/he needs it. The second challenge was to develop a realistic and at the same time adequate scenario and measurement

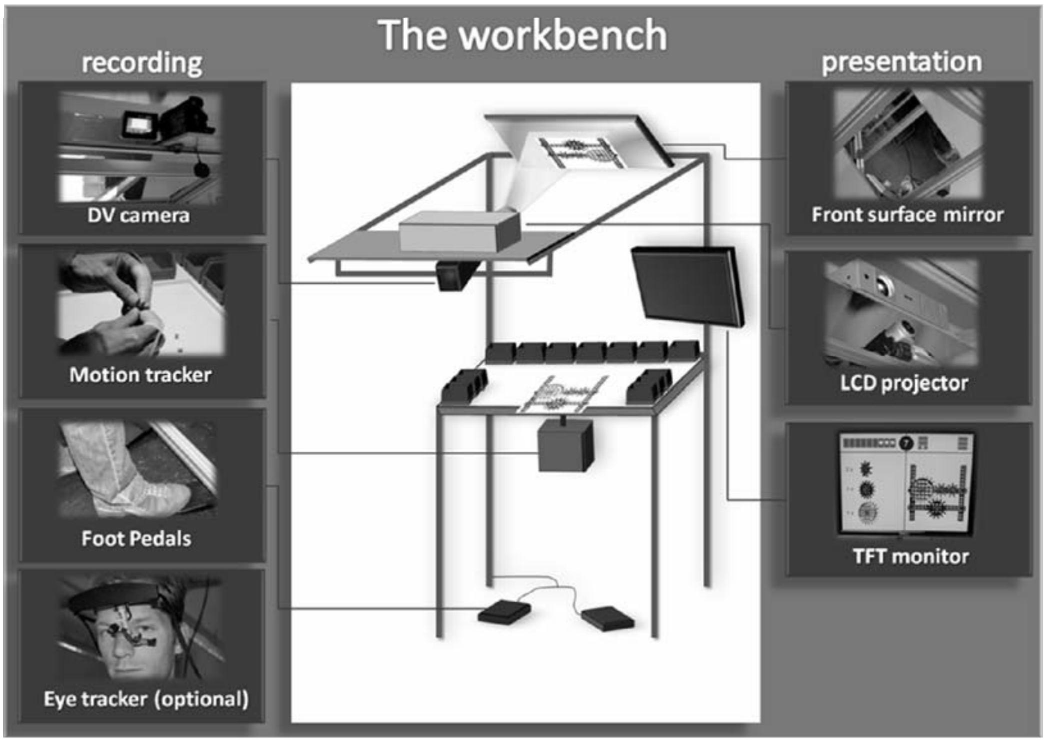


Figure 2: Schematic depiction of the workbench.

methods which allow investigating the cognitive processes involved in assembly tasks in great detail, in order to achieve a better understanding of where workers are facing problems and why they arise. Cognitive bottlenecks need to be identified and specific processing stages need to be differentiated (see section 2). We met these challenges by designing a special workbench, which was equipped with a customized information presentation technique, a foot pedal based interaction concept and a tracker system in order to achieve an online feedback of human behavior to the machine (Figure 2).

4.1 Information Presentation

One of the crucial points in assisting the worker is the manner in which instructions are presented. Following the ideas of worker-adaptive information display [2] and contact analog information presentation in accordance with attentional and visual processing mechanisms in humans, we opted for an augmented reality solution where the instructions become an integrated part of the workbench. As most difficulties in creating a usable AR system were related to HMDs and tracking technologies we decided against such a system. Instead, the relatively static environment of a manual assembly workspace offered ideal conditions to make use of a fixed projection augmented reality setup. By projecting assembly information directly onto the workbench, it is possible to combine the advantages of an LCD projector over a HMD (increased contrast, luminance, resolution and colorspace, unhampered field of view, no strain on nose or head) with the possibilities to provide the information at the spatial location where the worker would need it, and to provide **contact analog** information on static objects within the work space. Furthermore, this approach has the potential of adapting the instruction display in real time according to the worker’s needs. This kind of AR support to a manual assembly task also bears the benefit of being cheaper to acquire, easier to install and operate and more likely to be tolerated by the worker for a daily use than HMD based approaches. In order to project the information onto the working plane, a standard workbench for manual assembly was extended overhead with a horizontally mounted LCD projector. As this setup should be operable not only in a laboratory environment but also within a real factory setting with variable lighting conditions, care was taken that the projector provided enough luminous power (4000 lumen), while at the same time providing a good projection ratio (projection distance : projected image diameter) and low operating noise. A front-surface mirror was installed at an angle of 45° at the front end of the workbench (Figure 2) so that the projected image got reflected to cover the whole working area. With this setup, we were able to provide contact analog information, for example a highlighting of only those boxes where parts for the current assembly step had to be taken from. In addition to the projection directly onto the working surface by means of the LCD projector and the front-surface mirror, this setup includes a 20” TFT monitor for alternative or complimentary information presentation that is located at an ergonomically favorable position, can be adjusted in angle and is easily (de)mountable.

4.2 Interaction Concept

While in many interaction concepts for AR applications in production environments a speech recognition system is employed, these systems were often error-prone due to insufficient speech recognition for commands that were altered by slang or background noise, and consequently annoying to the user (e.g. [9]). Particularly a noisy factory environment seems inappropriate to use speech as input modality for worker commands. In order to provide an interaction mode which is not based on speech or manual

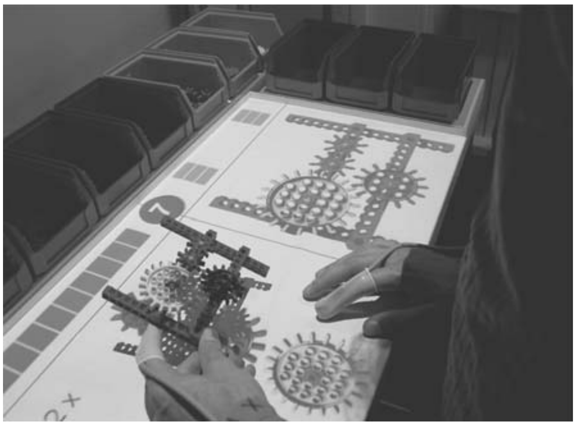


Figure 3: The augmented workbench including projected instruction and contact analog highlighting (boxes).

operation of some input device (which would interrupt the manual assembly workflow), we devised two custom built foot pedals which can be connected over USB to a control computer. The foot pedals are placed at easily reachable and adjustable positions, one operated by the right foot and one by the left foot. By operating the right foot pedal the worker can advance one step in the assembly guide, by operating the left foot pedal the worker can go back one step. Even though the ultimate goal is to develop a system that automatically adjusts to the worker's skills, experience and current mental workload and that way for example automatically detects mistakes and offers instant solutions, it is yet important to provide some means of user interaction so that the worker does not perceive himself as being controlled by the system. A foot-pedal input system realizes such a user control through a simple forward-backward flipping through the assembly steps, while, which is important, the focus of view, attention and bimanual handling can rest on the work piece.

4.3 Performance Measurement

The workbench offers the possibility to mount tracking technology by which the worker's performance can be registered online, interpreted in terms of skill level and current workload, and instructions can be modified accordingly. Currently, the workbench is equipped with a Polhemus Liberty motion tracking system (Figure 3) and a DV camera which can record worker behavior over the mirror. Behavioral variables, such as time-to completion for a certain assembly substeps, error rates and error categorization give a first measure of worker performance. Under controlled experimental conditions, these measures are helpful to compare e.g., different instruction modes or the effect of instruction complexity on assembly performance. As an example, the time it takes from instruction presentation until the worker initiates a movement can be regarded as an indirect measure of how easily the worker could decode the presented information. The scenario also incorporates an eye tracking system, which is especially valuable to identify the worker's fixation at a point in time and is an excellent estimation for attention allocation. Fixation patterns can reflect search patterns on which the workbench and instruction layout can be optimized.

5 EXPERIMENTAL STUDY

An experimental study was designed to test and validate the proposed workbench setup and gather empiric data of how worker performance is modulated by task setup and complexity. In order to quantify the benefits of projected and contact analog information, three presentation conditions were chosen: In condition 1, the instructions were presented on a 20” TFT monitor mounted to the workbench. In condition 2, instructions were projected directly onto the working area, and in condition 3 the worker received additional contact analog information on which parts were needed in the current step (Figure 3). Performance data was gathered with continuous motion tracking, foot pedal registering and video recordings of the assembly and complemented with a short questionnaire focusing on usability issues.

5.1 Study Design

Thirty subjects (20m/10f) participated in this pilot study. Most of the participants were students (age 20-30, average 23.9 years) with dominantly technical engineering background. The task was to assemble three models with LEGO Technic bricks according to a detailed instruction (Figure 3). In order to account for individual differences in past experience with LEGO bricks, a short training phase was conducted in which the participants had to build a simple LEGO model. Each participant had to perform three assembly tasks in counterbalanced order. Participants received the instructions in one of the three presentation modes (monitor, projection or contact analog). The assembly tasks were self-paced. The use of LEGO bricks to simulate a real assembly tasks offers a couple of advantages over similar construction systems: Firstly, they can be combined purely manually without the use of additional tools, second they are available in an unrivaled variety of colors and shapes, third, they are consistent in the way the parts can be mounted, and no particular motor skills or specialized training is necessary. The three models that had to be built were all similar in basic structure – they resembled a windmill – but differed in detail. Each model had to be constructed from a sequence of defined assembly steps that differed systematically in complexity with regard to the amount of pieces, the diversity of pieces and the class of assembly (e.g. ‘gear’ or ‘frame’). Performance was measured by assembly time per construction step, the number and type of errors made and the efficiency of grasping movement trajectories.

5.2 Results

Results showed that complexity of the assembly step had a huge impact on assembly performance: certain assembly steps (e.g. gear assembly) could be performed quite fast (on average 6.8 seconds per part), while others (e.g. gear fitting) took much longer (Figure 4). It could also be shown that the proposed projection method of instruction presentation was in general beneficial to the worker, especially in those situations where assembly instructions were highly complex (Figure 4). In relatively simple assembly steps, no difference of instruction type was observed.

6 CONCLUSIONS

In order to achieve more efficient and less error prone manual assembly, workers can be supported by a context-sensitive guidance and assistance system. Especially in the case of complex and highly variable assembly, productivity could be

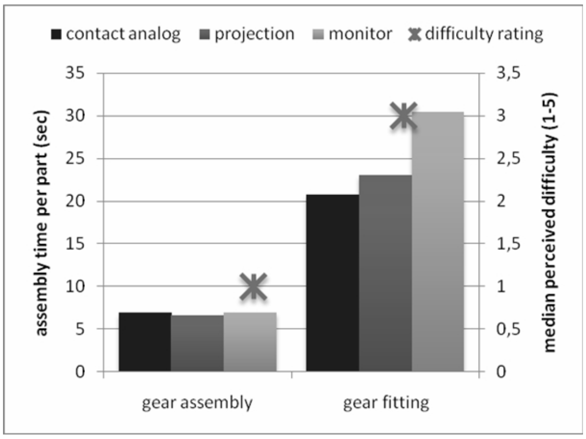


Figure 4: Average mounting times per part for two types of assembly steps. Contact analog and projection presentation mode benefit only for the more complex task (gear fitting). Step complexity reflected by user difficulty rating ('x', ranging from 1: not difficult at all, to 5: very difficult).

raised if the human is supported by an assistance system which provides instruction presentation tailored to the capabilities of the worker, the current situation, and helps to avoid and overcome errors.

The setup as it was proposed presents a first step towards such an assembly assistance system. We have shown that with limited technical effort it is possible to augment a standard workbench with instruction information directly on the workbench, reducing the need to switch attention back and forth between instructions and assembly area. Furthermore, the foot-pedal interaction concept allows continuous bimanual assembly. This scenario can serve as a research platform to investigate cognitive processes during manual assembly in greater detail. The experimental pilot study showed that the proposed approach including instruction projection directly onto the working plane and context-sensitive highlighting of objects led to fewer errors and faster assembly on more complex assembly steps. Contact analog information was especially helpful in drawing the worker's attention to relevant work-pieces and their location in cases where assembly parts were very similar to each other and hard to disambiguate.

7 FUTURE DIRECTIONS

So far, we included mainly behavioral assessments of worker performance. In order to investigate the underlying cognitive processes more detailed, this setup will be complemented with an eye tracking system (Figure 2). The combined movement and vision data will provide valuable insights of how efficiently the instruction information is taken up by the worker. Insights from controlled experiments will feed the design process of adaptive, worker tailored instruction presentation. Wireless tracking technologies need refinement to track the state of the environment, the work piece and the manual movements of the worker. The information gathered by the system may be fed into an environmental and a human model, which in turn dynamically influence the instruction presentation. Online analysis of movement patterns could be used to classify workers as belonging to different skill classes, and adopt instruction display accordingly: a novice on this task would need a different instruction presentation than an expert. Furthermore, assembly errors could be detected and even prevented by immediate feedback to the worker.

Further research on cognitive processes during manual assembly and further development of the assistive workbench will help to support the worker better in challenging assembly situations. By this adaptive accommodation of machines to human needs, human skill can be integrated tightly into the production process to achieve high degrees of efficiency and flexibility, where conventional automation fails.

8 ACKNOWLEDGEMENTS

Special thanks go to Frank Wallhoff, Alexander Bannat, Florian Friesdorf, Rainer Ullman and Cordula Vesper for their supportive collaboration in the ACIPE project. Furthermore, we would like to acknowledge the funding support of the Deutsche Forschungsgemeinschaft (DFG).

9 REFERENCES

[1] Putzer, H., Onken, R., 2003, COSA - A generic cognitive system architecture based on a cognitive model of human behavior, Cognition, Technology & Work , 5:140-151.

[2] Zaeh, M.F., Lau, C., Wiesbeck, M., Ostgathe, M., Vogl, W., 2007, Towards the Cognitive Factory, Proceedings of the 2nd International Conference on Changeable, Agile, Reconfigurable and Virtual Production, Toronto, Ontario, Canada.

[3] Johnson, A., Proctor, R.W., 2004, Attention: Theory and practice, Sage Publications Inc, Thousand Oaks, CA, USA.

[4] Wickens, C.D., Hollands, J.G., 2000, Engineering Psychology and Human Performance (3rd. ed.), Prentice Hall Inc, Upper Saddle River, NJ, USA.

[5] Crowson, R., 2006, Assembly Processes: finishing, packaging, and automation, Taylor & Francis Group, New York, NY, USA.

[6] Regenbrecht, H., 2006, Industrial Augmented Reality Applications, in: Haller, Thomas and Billinghurst, Emerging technologies of Augmented Reality: Interfaces and Design, 283-304, Idea Group Publishers, Hershey, PA, USA.

[7] Tang, A., Owen, C., Biocca, F., Mou, W., 2003, Comparative Effectiveness of Augmented Reality in Object Assembly, Proceedings of the SIGCHI conference on Human Factors in Computing Systems, 73-80, Ft. Lauderdale, FL, USA.

[8] Wiedenmaier, S., Oehme, O., Schmidt, L., Luczak, H., 2003, Augmented Reality (AR) for Assembly Processes: Design and Experimental Evaluation, International Journal of Human-Computer Interaction, 16(3): 497-514.

[9] Schmidt, L., Wiedenmaier, S., Oehme, O., Luczak, H., 2005, Benutzerzentrierte Gestaltung von Augmented Reality in der Produktion, in: Stary, Mensch & Computer 2005: Kunst und Wissenschaft - Grenzüberschreitungen der interaktiven ART, 51-60, Oldenbourg Verlag, München.

Social aspects of Plant monitoring and visualization

Kristian Martinsen[1] and Halvor Holtskog[2]
[1] Raufoss Technology and Industrial Management, Norway
[2] Department of Engineering, Gjøvik University College, Norway

Abstract
Monitoring and visualization of manufacturing plants can be both an aggregation of process monitoring to plant level as well as visualization of performance indicators and material flow. This paper contains a survey on 5 Norwegian manufacturing plants and suggestions on how to utilize of these systems for organizational learning and knowledge creation. The paper contains one in-depth case study on one of the plants.

Keywords:
Process and Plant Monitoring; Socio-Technical Systems; Knowledge Creation

1 INTRODUCTION

Kunzman et al [1] introduce the term “Productive metrology” and describes the mechanism for calculation of benefits from measurements expressed in economic numbers. Although the benefits might not be easily quantified, they contain generation of information, knowledge and know-how from measurements in manufacturing. One group of tools for Productive Metrology is *Plant monitoring and control (PMC) systems* used for measurement visualization, monitoring and control at a plant level. A PMC system is on top of real-time control systems and collects data into a common central database from a large amount of distributed sources in the manufacturing plant, such as materials, work pieces, processes, machines, tools and equipment.

There have been many studies on use of ICT in general and whether ICT systems increase productivity or effectiveness or not [2], [3]. This paper does not address this topic, but take the starting point that PMC systems do exist and are used. The paper indicates a potential for utilization of these systems for increased knowledge creation. The survey among Norwegian manufacturing industry underlines this potential.

1.1 Socio-technical system

All manufacturing plants have the characteristics of socio-technical systems with three main types of relations: relationships between machines, between people and between machines and people. The technical system and the social system follow different laws: the technical system is subject to the laws of natural sciences, while the social system follows the laws of human sciences. Yet, the systems are correlative as the functional task of manufacturing requires the interaction of both. Thus, improving either of these systems independently cannot optimize the system as a whole. Only by jointly optimizing both systems can the best match be achieved [4]. In this respect is the socio-technical perspective not only important while designing a system, but equally important in creating a platform for learning and development. A monitoring and visualization system can be one of the central interfaces in this learning platform and inter-dependency between the technical and the social systems.

1.2 Organizational learning and knowledge creation

In a manufacturing plant, there is usually a mix of novices, competent and expert individuals. While the training of novices typically is a mix of apprenticeship and formal education, the progress of experts is a more informal process. Kjellberg [5] points out the differences between Formalized learning and Entrepreneurial learning. She points out that the first type of learning typically leads to fragmented knowledge creation, and the second typically more holistic knowledge creation. Problem solving, innovations, involvement and learning from others are important factors in the Entrepreneurial learning. Learning as a process that takes place in a participative framework based on interaction and co-participation, becomes a feature of practice which might be present in **all sorts of activities** not just in clear cases of studying, training and apprenticeship [6].

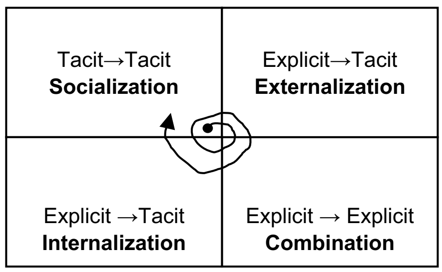


Figure 1. Nonaka’s modes of knowledge creation [7].

Nonaka [7] define knowledge as ‘justified true belief’. Knowledge reflects what a person have experienced and believes is true. He splits knowledge into tacit and explicit knowledge according to Polanyi [8], [9] where the statement 'We know more than we can tell' stand as a general assumption. Codified or explicit knowledge refers to formal coded information which is easily transmitted to others. Tacit knowledge, on the other hand, has a personal quality which makes it difficult to formalize. It contains a cognitive side with mental models, beliefs, paradigms, viewpoints and so on that helps us to understand our world, and a technical side that covers skills, craft, etc. Nonaka [7] defines 4 modes of knowledge creation (SECI model) as shown in Figure 1; Tacit to tacit, tacit to explicit, explicit to explicit and explicit to tacit. The SECI model indicates that knowledge creation follows a spiral along these modes.

2 EVOLUTION MODEL

The evolution of a PMC into a tool for knowledge creation can be described in 5 stages where each step is dependent on the previous steps;

Stage 1: Data Collection: The most basic use of a PMC is simple data collection. The motivation is typically the need for documentation of the process in case of product failure.

Stage 2: Aggregated Process Monitoring and Control: Data and values critical for productivity and/or quality are aggregated and visualized on a plant level. Out-of-limits alarms etc. gives operators information and tools needed to make necessary actions for overall process control.

Stage 3: Automated Performance Measurements and Visualization: Key performance indicators such as Overall Equipment Efficiency (OEE), Productivity, throughput times etc. are automatically calculated and visualized. This can be connected to visualization of production plans etc. K.K.B. Hon [10] gives an extensive list of performance measurements, but points out the importance of clear cause and effect relationships.

Stage 4: Knowledge Sharing and Communication: If the PMC system includes knowledge sharing and communication, we choose to call this level 4. This can be integrated tools for passing information to the next shift, between operators, from operator to maintenance personnel, management, etc. The tools can be interlinked with events such as tool failure or machine breakdowns reported in the PMC.

Stage 5: Integrated in Knowledge Creation Process: The highest level is when the PMC is integrated in a company or a plants knowledge creation process. This is described in detail in the next chapter.

2.1 PMC Survey in Norwegian manufacturing industry

Table 1 shows some results from a survey made on Norwegian manufacturing plants on the level of their PMCs according to the above mentioned stages.

Table 1: Results from PMC Survey in Norway.

	Stage 1	Stage 2	Stage 3	Stage 4	Stage 5
A	Yes	Yes	Yes	Partly	No
B	Yes	Yes	Partly	No	No
C	Yes	Partly	Partly	No	No
D	Yes	Partly	No	No	No
E	Yes	No	No	No	No

Plant A, D and E are automotive Tier 1 suppliers, Plant B and C manufactures consumer goods for an international marked. All plants are highly automated and have PMC system installed. All of the companies have data collection and more or less process monitoring and control, but only one company are using the PMC for communication and knowledge sharing tool, and even they only half-hearted. None of the surveyed company can be seen as using the PMC as an integrated tool for knowledge creation.

3 PMC INTERGRATED IN KNOWLEDGE CREATION

Since a PMC system will collect and display a large amount of data from the manufacturing, the system can be an important source for knowledge creation about the manufacturing system, processes and products. One can argue that Nonaka's SECI model of Knowledge Creation is present throughout every stage in the evolution model. To be perceived as integrated in Knowledge Creation process however, the PMC should preferably reflect the following:

- PMC as a central role in the information infrastructure
- Transparent information flow with information pull
- Decision support on all levels
- Tool for organizational unlearning
- Tools for reflection and improvement work-shops

3.1 Information infrastructure

A PMC system has the potential of playing a central role in a plant information infrastructure. To achieve this, the system should be working according to the following principals:

1. *Supporting or enabling* function, designed to support a wide range of activities, not especially tailored to one
2. *Shared by a larger community* – in a sense that it is the same singe object used by all users
3. *Open* – in the sense that strict borders of the use of the information is impossible to define
4. *More than pure technology*, it should be a part of the socio-technical system. It is the 'container' of explicit knowledge which is a fundamental piece in the SECI process

3.2 Transparent information flow

Information management is often considered to be important control power. With the introduction of transparent information flow the possibility of filtering and manipulate information is lost, and may imply a change from information push towards information pull. In an information push infrastructure, the sender of information decides who is in the need for this information, and send it to the receivers he choose. This can lead to persons that do not need this information are receiving it anyway, and the opposite: persons that need the information are not getting it. Especially among manufacturing process operators, there is often a feeling of lack of enough information on 'what is going on'. In an information pull infrastructure, the receivers need to seek information themselves for example at an intranet page. In most real cases there will be a mixture of push and pull, but the ratio might be different. This seeking of information is the essence of knowledge creation or commitment according to the Knowledge Creation process. Push and pull of information can many times translate to the conversion of knowledge along the epistemological dimension (tacit vs. explicit).

3.3 Decision support

To be able to benefit from the PMC system for process control, the operators must be able to make decision based on the signals from the monitoring as well as all other available information. This requires a high degree of process knowledge, as well as skills:

- Efficient problem solving is a key success factor. A problem has to be solved as quickly as possible
- Effectively avoid the reoccurrence of a problem. Efficient root cause detection and elimination is a key factor to reduce the number of faults
- High degree of abstract knowledge (transforming knowledge to models as a tool for knowledge creation)
- Scientific approach (systematic optimization and knowledge creation)

Efficient problem solving is more related to technical insight and experience from similar situations, whereas root cause detection is related to a systematic model based approach, i.e. scientific working methods. Analysis and reflections based on data collected in the PMC and the experts' analytical skills is the basis of this.

3.4 Organizational Unlearning

Knowledge creation is seen as the foundation for the improvement processes. One hinder for improvement, however, is what we believe to be true but actually is not true. When the PMC system is used, it is important to have an open mind, and to be as little biased as possible. Some of the knowledge gained will be confirmation of beliefs; other might contradict well established beliefs. The use of a PMC system opens a possibility to reflect and rethink.

To change established beliefs has proven to be very hard. One way to achieve this is so called 'organizational unlearning' [11] where learning is subdivided into increasing of knowledge and into changing of attitudes. Attitudes include routines and habits and are generally unwritten rules. Unlearning of attitudes is therefore hard to achieve systematically. McGill [11] points at learning culture and open information flow, similar as discussed earlier in this chapter to achieve unlearning. In addition is the responsibility for managers to promote unlearning addressed.

3.5 Workshops

Improvement (kaizen) and reflection workshops are basic tools for knowledge creation in most modern manufacturing plants. These can be both arranged at pre-determined intervals and/or be trigged by an unwanted event, such as quality loss or equipment failure. To use data from the PMC in such workshops is not uncommon. There might be, however, a lack of good report generation and a lot of work need to be done beforehand to collect and display data. The ideal PMC should be able to automatically generate special reports and initiate the workshops. In addition, there should be integrated tool for fast analysis and visualization of data when special needs emerge during the workshop.

4 CASE STUDY AT RAUFOSS TECHNOLOGY AS

Raufoss Technology AS (RT) is developing and manufacturing aluminum wheel suspension parts and systems. The manufacturing system has continuous automated lines and includes processes such as hot and cold forming, heat treatment, machining and assembly. They implemented a fully integrated PMC system at Raufoss when this plant was rebuilt in 2000/2001 to meet the needs of a new contract on control arms for the GM epsilon platform (among others Opel Vectra and Saab 9-3).

4.1 Initial technical functionality of RT PMC system

The PMC system at RT is based on Siemens WinCC with self-developed modules for shift reports, performance metrics and communication. The PMC has the following functionality:

Aggregated Process monitoring and Control

Based on Siemens WinCC, a plant monitoring and control system was build. The interface has a hierarchical structure with the complete plant on top, the main processes on the next level. This corresponds with the organization with one team on each main process (in total 4 teams on the shop floor). When control values are out-of-limits, an alarm signal is shown on the screen and the process is halted. The operator can look at historical values on all process measurements and control values.

Material flow and manufacturing plans

Current manufacturing plans, expected output, as well as current actual output of each shift are shown. Material flow indicators such as number of parts in buffers etc. are continuously displayed.

Performance metrics

Overall Equipment Efficiency (OEE = Quality rate X Availability X Speed rate) and productivity is automatically calculated and reported for each shift, each week and year for each main process and the complete plant. The performance of each shift is shown in a graph with a time-bar dividing the total time on a shift into four classes;

1. Planned stops
2. Waiting to receive raw materials or deliver parts to next process
3. Unplanned stops
4. Running manufacturing

Connected to the graph showing the division of time consumption is a curve shows the current productivity. The class 2 (waiting) applies only to a single process and not on the complete plant. Furthermore is class 2 not included in time losses while calculation OEE for a single process. The idea is that the team on one process should not be given a bad OEE result if the fault is the preceding process not delivering, or the succeeding process not being able to receive and buffers are respectively empty or full. The OEE for the total manufacturing line will, of course be affected.

Pareto reports on stops and quality losses

Reports on the number of unplanned stops grouped in different time spans, as well as types of quality losses are reported on Pareto charts.

Reports and communication between shifts

The start and end of unplanned stops are automatically reported as well as the station the stop was initiated. If the stop was caused by an out-of-limits process control value, this value is given. More important, however is the possibility for operators to give written comments and messages to the next shift, maintenance personnel etc. about problems and stop causes.

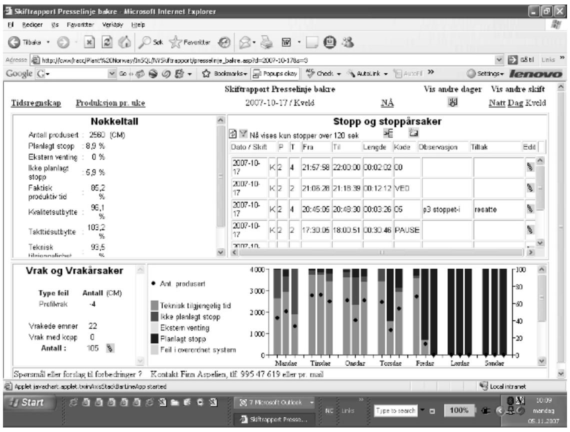


Figure 2: Example of actual shift report page.

4.2 The knowledge creation process

In the survey the workplace was viewed as an arena for knowledge creation and learning. Interviews with several operators from different part of the factory and different employment time where made. As a novice the apprenticeship model is dominant, where the novice acquires tacit knowledge through shared experience with the PMC system. It was found that the knowledge of experience (commitment, reflection and logical thinking of the mentor) was a key factor for knowledge creation. This is the case for the novice as well as for the mentor, for both the socialization and externalization mode and is supported by the notion of mutual engagement [12]. Another key factor for the mentor is the variety of experience. Together, knowledge of experience and variety of experience determine the quality of the mentor's tacit knowledge. A system of this kind will bring in a shared repertoire and a shared language. It was found evidence of this in the interviews. This in turn leads to the knowledge of rationality, 'which describes a rational ability to reflect on experience' [7]. The operators do this in the briefing / debriefing between shifts. Furthermore, there is some reflection in action [13] at the production floor by visiting other teams and sharing information. But it is clear that they did not do this in a systematic combination mode according to Nonaka. Even so, the conceptualization of these shared experiences was clear to the group of operators. When asked about the benefits of the PMC system, they gave more or less equal answers and have the same idea of what changes could be made in order to make the system better. These answers tell us that there is a certain level of common reflection among the operators. The transitions between shifts and teams are the core interaction rhythm that Nonaka means is essential for accelerating the knowledge creation process. The written comments and messages are the equivalent to explicit knowledge. In the effort to minimize unwanted breakdowns and productivity loss, the operators use the system to look at situations in the past and what happened on the previous shift. From a managerial side the system of PMC is connected with the Total Productive Maintenance work, and the reflection actions are connected to operator self maintenance etc.

5 CONCLUSIONS AND FURTER WORK

This paper describes an evolution of PMC system implementation used as a tool for learning process where knowledge is created. The cultural aspect, however, which is essential in the community of practice theory have not been addressed in this paper. Future work will look into it and will be implemented in the evolution model.

6 ACKNOWLEDGMENTS

This is a NORMAN CRI research work founded by the Norwegian Research Council and partners from Norwegian manufacturing industry. The authors wish to thank NFR as well as industrial partners taking part in the survey, and in particular RT for opening up for the in-depth study.

7 REFERENCES

[1] Kunzman, H. et al. 2005, Productive Metrology, Adding Value to Manufacture, Annals of the CIRP, Vol 54/2 pp. 691 - 704

[2] Davenport, T. H., 2007, Information Technologies for Knowledge Management, chap. 7 in Knowledge Creation and Management, Ichijo, K. and Nonaka, I. (eds), Oxford University Press.

[3] Carr, N. G., 2004, Does IT Matter?, Harvard School Publishing Corporation.

[4] Frei et. al, 1993, Work Design for the Competent Organization, Westport: Quorum.

[5] Kjellberg, A., 1999, Teams - What's next? From fragmentation and Consciousness to Responsiveness by Competence Management for Modular Manufacturing Learning, Annals of the CIRP 48/2, pp. 599 - 609

[6] Lave, J., Wenger, E., 1991, Situated learning: Legitimate peripheral participation, Cambridge: Cambridge University Press.

[7] Nonaka, I., 1994, A Dynamic Theory of Organizational Knowledge Creation, Organization Science, Vol. 5, No. 1, pp.14-37

[8] Polanyi, M., 1958, Personal Knowledge, Chicago, IL: The University of Chicago Press.

[9] Polanyi, M., 1966, The Tacit Dimension, London: Routledge & Kegan Paul.

[10] Hon, K.K.B., 2005, Performance and Evaluations of Manufacturing Systems, Annals of the CIRP, Vol 54/2 pp. 675 - 690

[11] McGill, M. E., J. W. Slocum, 1993, Unlearning the organization, Organizational Dynamics, Vol. 22(Autumn 1993), pp. 67-79.

[12] Wenger, E., 1998, Communities of Practice, Learning, Meaning, and Identity, Cambridge University Press.

[13] Schön, D. A., 1983, The Reflective Practitioner, New York: Basic Books.

Cluster Manufacturing Management to Improve Equipment Efficiency and Productivity

Suthep Butdee[1] and Serge Tichkiewitch[2]

[1] Thai-French Innovation Center (TFIC) &

Integrated Manufacturing System Research Center, Department of Production Engineering

Faculty of Engineering, King Mongkut's University of Technology North Bangkok, Thailand

[2] G-SCOP – Grenoble Science Conception Optimization and Production, Grenoble Institute of Technology

Abstract

This paper purposes concept of cluster manufacturing management in order to improve equipment efficiency and productivity for SMEs. The cluster group performs manufacturing tasks as collaborative engineering in the design phase. They share the process of designing information of product using CAD/CAPP/CAE and PDM. ERP (Enterprise Resource Planning) system management is implemented to manage the cluster company group with the same platform which would be established inside companies. Instead of working separately, those companies would share facilities to produce standard parts. Consequently, the lead time and cost of production are reduced; the quality standard is more consistent. Each company's ERP is linked to SCM in order to control the delivery of components to assembly parts of final production stage of the selected company which is considered the most capable and appropriate. The transportation cost is optimized at this stage. The productivity of equipment efficiency is increased clearly in this model.

Keywords:

Cluster Management, Overall Efficiency Equipment, Collaborative System

1 INTRODUCTION

Presently, business management requires new concept and model which are driven by technology in order to survive in today's highly competitive environment. In response to increasing customer demand and dynamic competition, companies are under high pressure to shorten time to market by providing tailored products to the customer for the economy of scope, to reduce time via mass production for the economy of scale, and to decrease time to profit by increasing efficiency of the entire lifecycle for the economy of service. The technology needs are included; the speed up product development, the enhancement of manufacturing and the capacity of supply; the improvement revenue from lifecycle efficiency [1]. The manufacturing productivity indicators are determined as the performance of enterprise competitiveness. They are divided into two stages: the undesirable stage and desirable stage. The undesirable stage consists of the inventory, quality cost, time to market according to schedule and unscheduled changes, whereas the desirable stage contains maximized overall productivity, quality, profitability and customer requirement. Overall productivity is the cumulative gain or loss. A higher level of productivity in one specific department or discipline is not a good measurement. Productivity means creating concepts that positively impact the whole system – both the upstream and the downstream operations. The overall productivity is defined as the ratio of the throughput to the operating expense. The throughput is defined as useful outputs so that the productivity entails the effective measure of how input (people, materials, means, etc.) are utilized in a certain period (measured in terms of operating expenses), in order to realize useful outputs [2]. The throughput is defined as follows:

$$T = \sum_{i=1}^{No} [P_i * N_i * P_{vi}] \qquad (1)$$

The equation 1 represents the *Pi* as the proportion of acceptance outputs which are non-defective of variant *i*; the *Ni* as the total number of outputs produced of variant *i*; *Pvi* as the production (or throughput*)* value per acceptable output *i*; the *No* as the number of outputs or assembly variants.

In this study, the concept of company cluster, which means several companies are grouped together in order to share the process of designing information by using the digital tools, is grouped together. The ultimate target goal is to improve equipment efficiency and productivity. In addition, the cluster manufacturing model enhances the competitiveness. The cluster can serve big quantities of customer orders. They can control the product process to make the quality of product consistently. The digital tool consists of CAD/CAPP/CAE and PDM for the design stage where the cluster can share the product model, drawings, product structure and process planning. ERP is established in the same platform and used in such cluster. The supply chain management concept is employed to serve the effective transportation among the manufacturers and supplier inside the cluster.

2 RELATED PREVIOUS WORKS

Nylund et al [3] presented the concept of an adaptive and autonomous manufacturing system which is based on the principles of Holonic Manufacturing System (HMSs), fraction factory and service oriented architecture (SOA). The authors presented the FMS 2010 project framework which was based on manufacturing knowledge and skills that are used in every design and development case. The adaptive manufacturing system was based on the HMS which consists of six components. They are knowledge base, digital presentation, autonomous units, co-operative units, service oriented and self similarity. Jana et al [3] stated that OEE – overall equipment effectiveness is one of the critical measurement indicators of the TPM strategy that the main principle lies in maximizing the saleable output from equipments by enhancing productivity in the three areas – availability, performance and quality. OEE strives to focus in the areas of machine efficiency, operator performance and quality of products. The shop floor of a government tool organization was determined by the OEE criteria. After six months later, the productivity improvement increased from a meager 23%-56%. Silva and Houten [5] expressed that the advanced manufacturing business of concurrent engineering system is often complemented by a parallel team organization, which brings various functions together from the very beginning of product development. At the same time, there is a shift from a producer's market to a consumer's market. This means the increasing in the demand for a higher product variety. The introduction of such product variety increases two basic categories of costs: production costs and market mediation costs. Consequently, companies must carry out the paradoxical task of providing exclusive products at low cost, high quality, and at short delivery times. An advanced logistics management strategy must be determined together with a product's functional specifications to achieve an optimal life cycle performance. Jacobsen [6] stated that design of production system is a challenging activity. This paper presented MERIP model that means Human Resources in Production translating from the Danish language. Humans have an inborn curiosity, creativity and propensity to create a meaning in what they are doing, which are the basis for creating motivation. Humans have a dignity, which also should be considered. The human resources are involved in new ways in order to establish new competitive production system. The MERIP showed the production system consists of four equal valued building blocks, technology, human resources, information and organization. Gilad et al [7] presented a model partitioning and clustering algorithm (MPCA) which determines similarity between two models and assign to parallel assembly lines. According to their similarity, Lehman's algorithm deals with optimization for balancing mixed model assembly lines. Spath et al [8] presented that the globalization of markets has caused companies to become part of spatially distributed dynamic supply networks, in which, next to the fundamental production process, system and business overlapping cooperation of the processes has increased in importance. This paper created the organization and resource model with organization units and allocated resources. The model is divided into four levels: enterprise, plant, manufacturing and resource group. Therefore, the available resources are taken into account involving through customers and suppliers using the concept of supply chain management. In additions, the Enterprise Resources Planning – ERP systems are introduced to associate the company. However, they have not designed to simulate the company overlapping supply chains. They perform on fixed capacity quantities and production times as basis. The supply network simulation requires event orientation that considers production processes and stock keeping processes as well as transportation. The major objective is to optimize the task execution with reference to the time to market and target costs. The available resources and their organizational structures are mapped into an organization model in order to check whether the necessary resources for the appropriate time are available or engaged to other parallel activities. Consequently, the structure of business processes as well as the allocation of the resources can be clearly presented through the internal and intermediate organizational structure across collaborative companies. Zancul [9] stated that the ERPs are enterprise management systems whose main features include a wide scope of functionalities, capacity of adjustment to several kinds of enterprises and data integration. These systems are basically made up of modules and a central database. The modules have all the functionalities to support the enterprise, such as marketing, sales, purchase, production, human resources management, as well as management of physical and financial resources. Data employed by each module is stored in the central database in order to be handled by other modules, ensuring the integration among the business processes. The broadened scope of ERP, with the inclusion of new functionalities is product life cycle management (PLM). The main functionalities of ERP consist of project management, product data management (PDM), and computer-aided process planning (CAPP). Nyhuis [10] presented the process model for factory planning which contains a description of factory planning processes with the focus on the process and facility planning. It serves the extended demands in factory planning by defining interfaces to other planning departments. Murgu [11] expressed efficient management of the information flows that is critical for the efficiency of the whole system. High productivity in all operations conducted within manufacturing companies is needed in order to be a core player on the competitive market today and tomorrow. Basic tools and methods associated with lean production on the shop floor have been around since the beginning of the 80's and they are now mature, available and widely used. Ming et al [12] presented that in the modern global economy, companies are facing ever-increasing challenges for short time to market to enter into the market early, for reduced time to volume to occupy the market quickly, and for decreased time to profit to get return from market shortly. Product life cycle management –PLM is recognized as one of the key leading technologies to facilitate companies to overcome the challenges, which will offer companies a new way to rapidly plan, organize, manage, measure, and deliver new products and services much faster, better and cheaper in an integrated way. PLM has recently been recognized as a new strategic business approach in support of collaborative creation, management, dissemination, and use of product assets, including data, information, knowledge, etc., across extended enterprise from concept to end of life – integrating people, processes, and technology. Stefanovic [13] purposed the possibility of establishing relationships between processes in supply networks and functioning of the entire system. In this model, the integrated system, all relevant factors for supply network management, both at the global level and at the single process level are observed. The concept is to form up process library of supply network, which would contain process description, input, output, and the way the process is realized. Supply chain management in today's fast changing business environment is facing many challenges. Hilmes [15] studies the presentation of a specification supporting the production development under concurrent development to improve competitiveness. Lu [16] stated that the collaborative is characterized by more durable and pervasive relationships. The group works together toward a common goal which the team attempts to find solutions satisfying to all involvements. Figure 1 shows the revolution and comparison of the design engineering: (a) sequential engineering, (b) concurrent engineering and (c) collaborative engineering. The sequential one takes more time to finish the production, whereas the concurrent one take more advantages by using the overlapping process. However, the concurrent engineering is still lack of the cooperation. Therefore, the collaborative engineering takes the most advantages because it is overlapping method and cooperation. It can mean the integrated design.

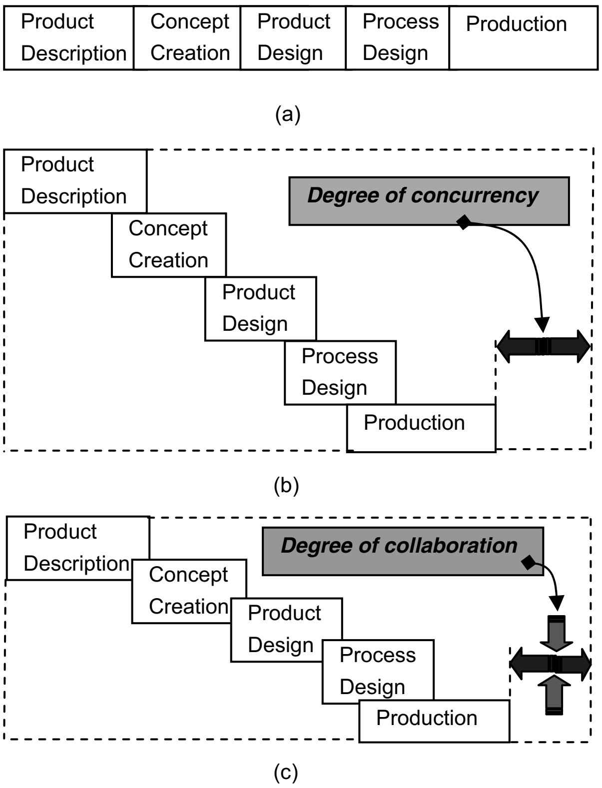


Figure 1 Revolution of Design Engineering Process [15]

3 CLUSTER MANUFACTURING MODEL

In this study, the five companies where they work together in the same area are collaborative as a manufacturing cluster. Four of them are bus body makers and the fifth is a supplier. This research objective is to integrate them in order to optimize and utilize the facilities. Figure 2 shows cycle of the company cluster that is managed as collaborative manufacturing system in order to improve productivity of the whole group cluster. It shows the link of FAC_1, FAC_2, FAC_3 until FAC_N. The system components consist of CAD – Computer Aided Design, CAPP – Computer Aided Process Planning, PDM – Product Data Management, ERP – Enterprise Resource Planning, SCM – Supply Chain Management and LCM – Life Cycle Management. The concept is linked as PLM – Product Life Cycle Management. The cluster shares the design of product via the product data management. The product data are shared and the job design works are distributed to each company of the group cluster. ERP system is used to manage the manufacturing activities. However, the ERP is applied for individual company. The SCM is managed by a company in the cluster. Raw materials, spare parts and other accessories for producing products are economically and effectively supplied. In cluster group of companies, it can be several factories.

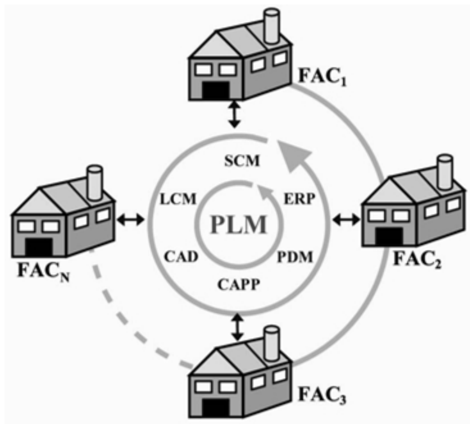


Figure 2. The Cycle of Collaborative Factory Cluster System Management

The relation of collaborative design occurs in the stage of design. The new product model is created on CAD. Then the process is designed based on the standard production route. CAPP is used in this step. The model is exported to the CAPP in STEP – Standard for the Exchange of Product Model Data. BOMs – Bill Of Materials are generated. Machine, equipments and tools are selected based on each process. The PDM linking the product model on CAD to ERP is used to manage the factory shop floor. The purchase orders are received from customers, which will be planned for production process both master plan and daily plan. Production time and cost are estimated. Finally, the finished products deliver to customers. The next section presents the case study of cluster manufacturing management.

4 CASE STUDY

The case study presents the process of the group of bus body manufacturing in Thailand where they join together under the concept of collaborative cluster manufacturing model. This project is supported by the Thai Government through the IMSRC consulting organization. There are more than 30 companies establishing since the last 20 years in the same areas. Five companies are selected and controlled as the pilot cluster. One of them is a supplier enterprise and the other four are bus body manufacturing companies. Currently, the pilot cluster is creating the sharing a large quantity order by sharing facilities and time. Collaborative system has been established based on working group. The six wheel bus is set as the pilot project. The quantity order is 500 buses that are over capacity for one company. Each designer of each company is shared in the design phase. The BOM data is then shared for each individual company in order to autonomous control their factory. However, the production plans are controlled by the cluster collaborators. Figure 4 shows the sample digital mockup bus body design. Collaborative engineering group cluster are linked and shared CAD design.

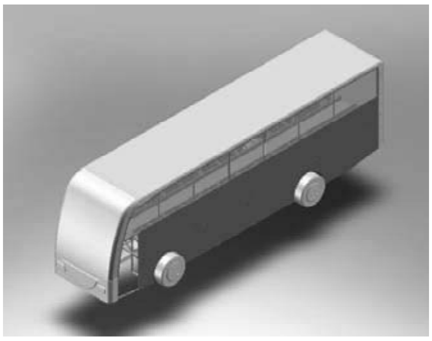


Figure 3. The Pilot Digital Mockup Bus Design

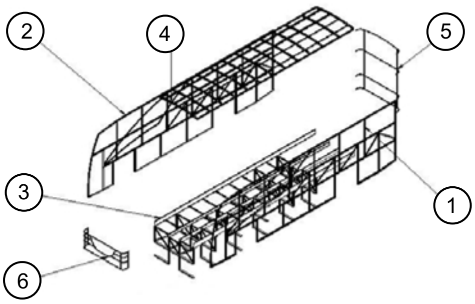


Figure 4. The Bus Body Frame Parts

The bus structure consists of eight main components. They are chassis, bus body frame including left side, right side, front side and back side, mirror, seat, interior parts, electrical control system, sound system and outside coating. All components are assembled as shown in Figure 4. They contain six parts. Each part is explored as the sub-components. In this step the planning process is created as the standard process that is then distributed to the cluster member and managed by ERP. It is composed of eleven modules: customer, supplier, design, marketing and sales, accounting, purchasing, production planning, inventory, manufacturing, quality control and maintenance.

5 SUMMARY

The concept of cluster manufacturing management has been presented. The related previous works are reviewed that some of the idea are adapted. The concept of collaborative design has been successfully implemented together with ERP program that is used to manage the factory which is controlled by the cluster collaborator. SCM, LCM and PLM concept are applied to be the effective cycle method of the manufacturing system management. This concept is clearly shown for the equipment efficiency and productivity improvement.

6 REFERENCES

[1] Ming, X.G., Yan, J.Q., Lu, F.W. and Ma, D.Z., 2005, Technology Solutions for Collaborative Product Lifecycle Management – Status Review and Future Trend, The International Journal of Concurrent Engineering: Researches and Applications, 13/4: 311-319.

[2] Prasad, B. (1996) Concurrent Engineering Fundamentals, Prentice Hall Inc.

[3] Nylund, H., Hokkanen, H., Saminen, M. and Addersson, P. (2007) Life Cycle of Digital Manufacturing Supporting Design and Development of Adaptive Manufacturing Systems, The Proceedings of the 12th Annual International Conference on Industrial Engineering Theory, Applications and Practice, Mexico, pp. 544-549.

[4] Jana, D.R., Banerjee, A., Sahay, Chittaranjan and Jaquez, L. (2007) Improving Overall Equipment Effectiveness (OEE): A Case Study, The Proceedings of the 12th Annual International Conference on Industrial Engineering Theory, Applications and Practice, Mexico, pp.674-679.

[5] Silva, R.J. and Van Houten, F.J.A.M., 2006, Integrated Design Decision Support and Logistics Strategy, The CIRP Journal of Manufacturing System, 35/1: 23-30.

[6] Jacobsen, P., 2003, Design of Production Systems, The CIRP Journal of Manufacturing System, 32/1: 75-83.

[7] Gilad, I., Hazbany, S. and Shpitalni, M., 2006, Parallel Mixed-Model Assembly Lines for Efficient Operations and Cost Reduction, The CIRP Journal of Manufacturing System, 35/6: 513-522.

[8] Spath, D., Sternemann, K.-H and Lanza, G., 2003, Supply Networks Simulation, The CIRP Journal of Manufacturing System, 32/2: 143-150.

[9] Zancul, E.S. and Rozenfeld, H., 2006, Application of ERP Systems in the Product Development Process, The CIRP Journal of Manufacturing System, 35/5: 463-472.

[10] Nyhuis, P. and Elscher, A., 2006, Process Model for Factory Planning, The CIRP Journal of Manufacturing System, 35/5: 429-434.

[11] Murgu, A., Johansson, B. and Pejryd, L., 2006, A Study on the Interaction between Physical and Information Flows in Manufacturing Systems, The CIRP Journal of Manufacturing System, 35/5: 63-68.

[12] Ming, X.G., Yan, J.Q., Lu, W.F. and Ma, D.Z., 2005, Technology Solutions for Collaborative Product Lifecycle Management – Status Review and Future Trend, The International Journal of Concurrent Engineering Research and Practice, 13/4: 311-319.

[13] Stefanovic, D., Majstorivic, V. and Stefanovic, N., 2006, Mehtodology for Process Integration in Supply Networks, The CIRP Journal of Manufacturing System, 35/2: 143-150.

[14] Hilmes, R. and Lien, T.K., 2006, Production System Design Specification – A Tool to Improve Competitiveness, The CIRP Journal of Manufacturing System, 35/2: 119-127.

[15] Lu, S.C.Y. (2006) Complexity on Collaborative Engineering , International Consortium for Complexity Research (ICCR) Forum, Alberta, Canada

Scheduling

Penalty Distribution Method for Scheduling Based Supply Chain Management

Jia Yee Chai[1], Tatsuhiko Sakaguchi[2] and Keiichi Shirase[1]

[1] Graduate School of Engineering, Kobe University, Japan

[2] Organization of Advance Science and Technology, Kobe University, Japan

Abstract

Trends of globalization and advances in Information Technology (IT) have created opportunity in collaborative manufacturing across national borders. A dynamic supply chain configuration is more adaptable by project based manufacturing environment characterize by the involvement of Small and Medium-sized Enterprises (SMEs). This research proposes a three echelons dynamic supply chain model where a job shop style manufacturer represents the center node being focused that will autonomously negotiates contracts with downstream client and upstream suppliers based on parameters derived from a reactive scheduling engine and an inventory management system. Furthermore a method is proposed to distribute delay penalties of previously contracted jobs to multiple new orders being negotiated simultaneously.

Keywords:

Genetic Algorithm, Scheduling, Supply Chain Management, Contract Negotiation

1 INTRODUCTION

Recent trends in manufacturing organization going across national borders to source for customers and suppliers in order to improve competitiveness have attracted interest in issues of dynamic supply chain configuration. A dynamic supply chain configuration acknowledges the possibility of better supply chain configuration depending on the requirement of each demand and the competency of participating organizations. The term competence cells refer to individual organizations that hold particular manufacturing know-how, and are independent in organizational objectives and business strategies. Included in these organizations are many Small and Medium-sized Enterprise (SMEs) that are highly specialized in their field, and engaging in various contract or project based jobs.

Transaction of information, a critical success factor in building an effective supply chain, can be done with ease through internet. A customer can post an order online to be auctioned by interested suppliers. As this process migrates up the supply chain, the phenomenon of dynamic supply chain configuration is observed. The use of automated negotiation software between the customers and suppliers may results in better operation cost, negotiation duration, and also may avoid emotional human acts such as when the partners try to establish social interaction instead of focusing on the objectives of the negotiation.

Thus the objectives of the proposed research are:

- To create a deployable model that enable automatic dynamic supply chain configuration.
- To simulate agent based negotiation process considering production scheduling and inventory management.

2 PREVIOUS RESEARCHES ON SUPPLY CHAIN MODEL

This research is closely related with the work of Tanimizu Y. et al [1] that utilize reactive scheduling method to determine bid due date and bid price of negotiations in a 2 echelons dynamic supply chain. Pre-planned jobs that have their due date delayed due to the insertion of new order into the schedule will be compensated by an amount of penalties. When there are more than one new order in consideration, the manufacturer will create separate rescheduling process for each new order that run in ignorance of one and other. While this method is shown to perform better than previous model that can only negotiate new order sequentially, it has the disadvantage of consuming high computing power as each rescheduling process requires separate genetic algorithm (GA) calculation, and also disregard the interrelationship between the new orders.

Piramuthu S. [2] proposed a knowledge-based framework for automated dynamic supply chain configuration. The negotiation that constructs the supply chain was guided by knowledge database that capture the characteristic the suppliers. Relationship between the objectives being negotiated was not defined, and the prospective supply chain partners were strictly restricted by the knowledge database.

Neubert R. et al [3] proposed a method of automated negotiations of supply contracts for flexible production networks that transform all objectives of negotiations into a common comparable aspect. As a result, the agents being utilized are capable to negotiate as effectively without being limited to certain product or object of negotiation. While it suggested that the parameters of negotiation can be derived from production planning and control function of an organization, it offered no insight as of how to do that.

Kaihara [4] formulate the supply chain model as a discrete resource allocation problem under dynamic environment. It demonstrated the applicability of the virtual market concept, which hinted that the complicated total supply chain optimization problem can be simulated in a multi agent programming environment by properly defining the decision process of each agent according to economic principles. Differ from this model that focus on the optimization of supply and demand management through economic principles, our proposed model will be examine the mechanism of supply chain from the manufacturing operation management viewpoint while using multi agent simulation.

3 PROPOSED SUPPLY CHAIN MODEL

We propose to minimize the complexity of supply chain modeling by using a 3 echelon model that consists of a job shop style manufacturer at the middle of the node, the raw material suppliers at the upstream of the manufacturer, and the client at the downstream of the manufacturer. Figure 1 below shows the concept of the 3 echelon model proposed.

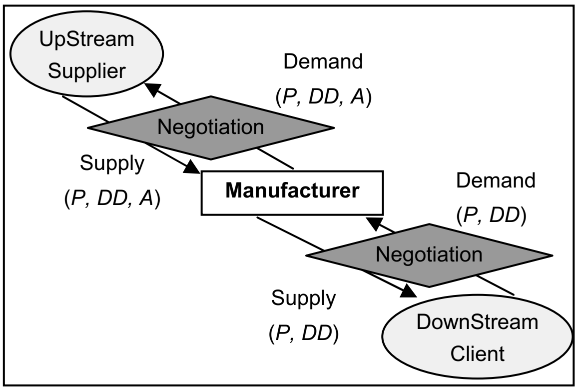


Figure 1: Proposed supply chain model.

The supply chain is dynamically formed whenever the downstream client places an order to the manufacturer. The mechanism to form the supply chain is the negotiation between these organizations. As it is a job shop style manufacturer, the aspects being negotiated with the downstream client are the price (*P*) and due date (*DD*) of the job but not the amount (*A*). The aspects being negotiated between the manufacturer and the upstream raw material supplier are the price, the due date and the amount.

The negotiations proceed sequentially, following the hill-climbing negotiation strategy, the negotiation agents will modify its next offer by relaxing the requirement of certain parameters being negotiated if the opponent also gives way.

For the upstream negotiation, the manufacturer determines the amount by monitoring the difference of the inventory level and the requirement of the order in the schedule. Negotiation of amount will influence the inventory holding cost of the manufacturer. The price being offered for the raw material is set at 55% of the profit the corresponding order as shown by Leenders and Fearon [5]. Negotiation of price will influence the bid price of the downstream negotiation. Negotiation of due date will influence the available start time of the order being negotiated downstream.

In this research we will be elaborating on the downstream negotiation where the proposed penalty distribution method will be influential as the negotiation parameters are derived from the result of reactive scheduling.

4 CALCULATING DOWNSTREAM BID PRICE FOR MANUFACTURER

In a market environment that allows for back ordering, the manufacturer can still fulfill the contract even if the due date of the job cannot be kept. In this case, the manufacturer will have to pay a delay penalty to the job owner of the said contract, calculated with the Equation (1) below:

$$D_i = p_i \times \left(dd_t - dd_{con}\right) \tag{1}$$

Where

D_i = Delay penalty of order *i*.

p_i = Penalty rate of job *i*

dd_t = due date in production schedule at time *t*.

dd_{con} = due date promised to the job owner.

When the manufacturer receives a new order from the client, the GA reactive scheduling engine will try to insert the new order into its current production schedule, and modify the schedule based on the objective function of total weighted tardiness. If the insertion of the new order causes any pre-planned job to delay, the penalty need to be paid to the owner of that is calculated by the Equation (1). In order to prevent excessive lost of revenue due to delay penalties, the bid price of the new order that causes these delays will be attributed with the delay penalties, as shown by Equation (2) below:

$$manufacturerBid = C_P \times PF - D_P + \sum_{i=1}^{P-1} D_i \tag{2}$$

Where

P = New order.

C_P = Cost of producing new order P.

PF = Profit mark up rate of the manufacturer.

D_i = Delay penalty of order *i*.

Logically, the timing of new order being inserted into the schedule will results in different amount of delay subjected to the existing orders. The earliest timing that the new order can be inserted into the schedule is,

$$StartTime_P = t_0 + Tneg_P + Tneg_{RM} + Dly_{RM} \tag{3}$$

Where

t_0 = start time of negotiation for new order P.

$Tneg_P$ = estimated negotiation duration for new order P.

$Tneg_{RM}$ = estimated negotiation duration for raw material
= 0 if raw material is immediately available.

Dly_{RM} = delivery time for raw material
= 0 if raw material is immediately available.

5 PROPOSED PENALTY DISTRIBUTION METHOD

By using Equation (2) to calculate the bid price for a new order, we assumed that this new order is the only reason that some of the pre-planned jobs are delayed after the GA engine's rescheduling process. Now, imagine when two new orders are to be negotiated simultaneously. Once the rescheduling process finishes, we will be able to determine the magnitude of the delay penalties, if any, with Equation (1). Yet we have two new orders in the schedule and we will need to calculate the bid price for each of the order. Obviously Equation (2) cannot be used directly as the distribution of the delay penalties among the new orders is necessary. If distributed illogically the manufacturer will risk not only losing out on the biddings, but also losing out on profit in the case where one of the negotiation failed after the other negotiation has been successfully secured, in which case the incorrectly distributed penalty at the failed negotiation will have to be burdened by the manufacturer as lost profits.

To solve the problem, we propose a method that is loosely based on critical path analysis. Imagine a job consists of several operations that are going to be processed sequentially by different resource (machine). Also assume that S_o is the schedule before the new orders are inserted and all the pre-planned jobs on it are on time. Then we have S_t which is the schedule at time *t* after the insertion of multiple new orders with some pre-planned jobs being delayed. Time *t* is also the time where the manufacturer has to submit the bids to the client on the new orders.

An operation *Opt* is delayed if its start time at S_t (*nEST*) is later than its start time at S_o (*oEST*). The magnitude and the reasons of delay of *Opt* are calculated with Equation (4) below by comparing its position and those of the operations around it between the schedule S_0 and S_t.

$$delayOfOpt = SC + \sum_{t=pEST}^{t=nEST} \left(ePT_t + gap_t \right) \tag{4}$$

Where:

SC = duration of previous operation of the same job in S_t that has a finishing time later than *oEST*.

pEST = potential earliest start time of *Opt* at S_t limited by the finishing time of the previous operation of the same job.
= *oEST* if the previous operation of the same job finish before *oEST*.

ePT_t = duration of processing time of operations of other jobs that exists between the limit of *pEST* and *nEST* at S_t.

gap_t = duration of gap that exists between the limit of *pEST* and *nEST* at S_t.

The logic is, only 3 reasons may cause *Opt* to be delayed:

- The previous operation of the same job is delayed (SC).
- *Opt* cannot start in S_t the same time as S_o due to delay or insertion of operations of other jobs in front of Opt at the same resource (ePT_t).
- Additional gap appeared in front of *Opt* in S_t. (gap_t)

For example we have a manufacturing system with 2 resources and the production schedule is loaded with 3 pre-planned jobs: W01, W02 and W03. The schedule S_0 before the insertion of a new order WX1 is shown in Figure 2. The schedule S_t after the insertion of the new order WX1 is shown in Figure 3.

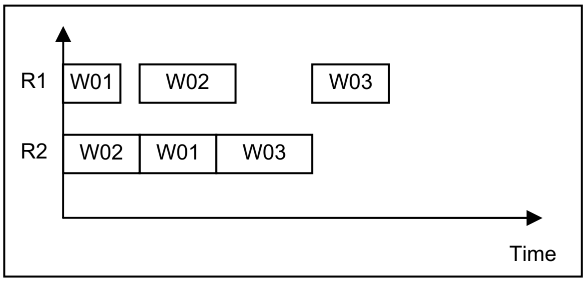


Figure 2: Schedule S_0 before insertion of new order.

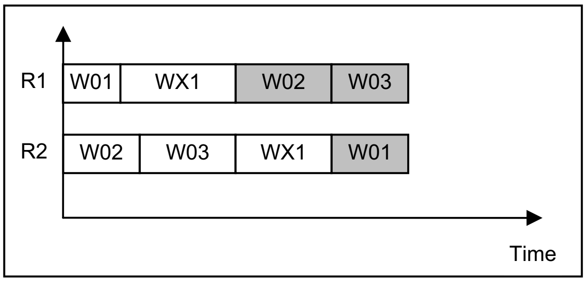


Figure 3: Schedule S_t after insertion of new order.

The operations that are delayed due to the insertion of new order WX1 are: operation of W01 at R2, operation of W02 at R1, and operation of W03 at R1. Table 1 below shows the start time and end time of each operation in both schedule.

Table 1: Start time and end time of each operations.

Operation	S_0		S_t	
	Start (s)	End (s)	Start (s)	End (s)
W01R01	0	30	0	30
W01R02	40	80	140	180
W02R01	40	90	90	140
W02R02	0	40	0	40
W03R01	130	170	140	180
W03R02	80	130	40	90
WX1R01	-	-	30	90
WX1R02	-	-	90	140

Using Equation (4), for operation W01R02:

SC = 0(previous operation, W01R01, is not delayed).

ePT_t = (50 + 50) = 100 (due to jump ahead of W03R02, and insertion of WX1R02).

gap_t = 0 (no additional gap between *pEST* and *nEST*).

Thus 50% of delay (50s) is attributed to the new order WX1, and another 50% of delay (50s) is attributed to GA rescheduling as it brings the operation W03R02 that is originally behind W01R02 to the front.

Using Equation (4), for operation W02R01:

SC = 0 (previous operation, W02R02, is not delayed).

ePT_t = 50 (due to insertion of WX1R01 in front of it).

gap_t = 0 (no additional gap between *pEST* and *nEST*).

Thus 100% of delay (50s) is attributed to the new order WX1.

Using Equation (4), for operation W03R01:

SC = 0 (previous operation, W03R02, is not delayed).

ePT_t = 10 (due to pushing back of W02R01 in front).

gap_t = 0 (no additional gap between *pEST* and *nEST*).

Thus 100% of delay (10s) is attributed to the cause of W02R01 being delayed, that is the insertion of WX1R01.

The delay that is caused by GA rescheduling will be equally divided between all the new orders being considered. After each delayed operation at S_t is accounted for, we use a recursive search algorithm to determine the root cause of the delays. The magnitude of delay at last operation of each delayed job determines the magnitude of delay of the job. The search will begin from those operations and branch upward by following each cause of delay. The search branch will stop when it reaches an operation that belongs to a new order, or an operation has no delay magnitude (the operation start at the same time, or even earlier in S_t than in S_o).

6 CASE STUDY AND RESULTS

Case studies are conducted on networked computers to simulate real life organizations communicating with each other through internet. Two manufacturers, one on each computer, are set to bid against each other for the contracts offered by one client. Both manufacturers start with the same initial production schedule, with 20 pre planned jobs of 10 operations each to be processed on a total of 10 resources in the manufacturing system. GA rescheduling engine for both are configured with same parameter, but with different penalty distribution methods. Manufacturer A utilizes the proposed method, while manufacturer B utilizes an equal distribution of penalty on the new orders being considered. The client will post a new offer every 200 seconds and will collect bids from the manufacturers at every 100 seconds interval. The total duration of simulation is 1800 seconds, in which 10 new orders are released by the client.

Table 2 below shows the profits for the manufacturers at the end of the simulation. Manufacturer A managed to secure more new orders from the client than manufacturer B by utilizing the proposed method. Manufacturer B still managed to capture some new orders as the schedule of manufacturer A was getting loaded by new orders, but the ability of the proposed method to distribute delay penalties intelligently allows the manufacturer A to bid more effectively. 3 orders failed to be secured by both the manufacturers and 1 order was still in the negotiation process when the simulation terminated.

Table 2: Manufacturers Profits.

	Manufacturer A	Manufacturer B
Gross Profit ($)	235005.0	215343.0
Delay Penalties ($)	8010.0	1320.0
Net Profit ($)	226995.0	214023.0
Contract Secured	4 orders	2 orders

For an example of the negotiation, bids at time 1675s are shown in Table 3. Bid of Work28 by manufacturer B had been rejected by the client in the previous iteration. The penalty distribution of manufacturer A is obviously different with manufacturer B. The difference in total penalty is cause by different results of GA rescheduling. At this iteration the bid of Work28 by manufacturer A was rejected by the client.

Table 3: Bids at 1675.

Manufacturer	Order	Price	Due Date	Penalty
A	Work28	2463.9	20369.0	1245.8
A	Work29	3309.5	20479.0	816.9
A	Work30	14689.7	18120.0	2289.7
B	Work29	2587.0	21780.0	1395.0
B	Work30	14995.0	16980.0	1395.0

Then at the next iteration at time 1776s, the bids are shown in Table 4. At this iteration, we can see that the penalty distribution of manufacturer A was clearly in proportion with the due date of the bids, while manufacturer B simply divided the penalty equally between the bids. The result is, the client awarded the contract of Work29 to manufacturer A at this iteration, and reject the bid of Work29 by manufacturer B.

Table 4: Bids at 1776.

Manufacturer	Order	Price	Due Date	Penalty
A	Work29	4995.0	19880.0	1564.2
A	Work30	13431.2	18230.0	2316.2
B	Work29	3568.3	21150.0	1407.5
B	Work30	11412.5	18970.0	1407.5

7 CONCLUSION

As this is still an ongoing research, we recognize there are areas to be improved, such as the distribution of penalty due to GA rescheduling that cannot be directly attributed to the new orders. However, the case study and results clearly shows the effectiveness of the method being proposed to distribute the delay penalties logically between new orders in negotiations.

8 ACKNOWLEDGEMENTS

A part of this work was supported by the Grant-in-Aid for Scientific Research (B) No. 17360063 of the Japan Society for the Promotion of Science.

9 REFERENCES

[1] Tanimizu Y., Maeda Y., Ozawa C., Sugimura N., 2007. "A study on dynamic supply chain considering production schedules – parallel scheduling for orders (In Japanese)," Proceedings of 51st ISCIE(2007) pp 83-84.

[2] Piramuthu S., 2004. "Knowledge-based framework for automated dynamic supply chain configuration," European Journal of Operational Research 165 (2005) 219 – 230.

[3] Neubert R., Gorlitz O., Teich T., 2003. "Automated negotiations of supply contracts for flexible production networks," International Journal of Production Economics 89 (2004) 175 – 187.

[4] Kaihara, T., 2003. "Multi-agent based supply chain modeling with dynamic environment," International Journal of Production Economics 85 (2003) 263 – 269.

[5] Leender, M.R., Fearon, H.E., 1998. "Purchasing and Supply Management 11th Edition," Irwin, Chicago.

Social contract based manufacturing scheduling with combinatorial auction mechanism

Toshiya Kaihara[1], Kentaro Sashio[1] and Katsuhito Miura [2]

[1] Department of Computer Science and Systems Engineering, Graduate School of Engineering, Kobe University, Japan

[2] Department of Computer and Systems Engineering, Graduate School of Engineering, Kobe University, Japan

Abstract

Efficiency and robustness are required simultaneously in manufacturing scheduling. Social contract mechanism takes a metaphor for social interactions amongst decision makers, and its aim is to execute a rational allocation of multiple resources in their society. Then it must have great potential to realise sophisticated manufacturing scheduling system in its nature. As one of social contract mechanisms, we focus on combinatorial auction approach, which is able to calculate optimal solutions in auction manner, and try to integrate it with manufacturing scheduling. First of all, we explain the general concept about combinatorial auction mechanism, and describe the formulation of our targeted scheduling problem with the mechanism in this paper. Then, the performance of the proposed methodology is evaluated with several simulation scenarios in terms of its optimality and robustness. Finally we try to clarify its effectiveness with industrial case study as practical situations.

Keywords:

Manufacturing scheduling; Combinatorial auction; Social contract; Multi-agent; Optimisation

1 INTRODUCTION

It is a time of changes for manufacturing industry today. Market trends are rapidly changing every minute, and deciding what and how much to sell at the right moment is a very crucial and difficult matter. In order to adapt to the dynamic changes in the market, the need for manufacturing systems with high agility and flexibility, that is, manufacturing systems that are capable of producing the desired products at the desired amount timely, has been long pointed out by both business and academia [1]. Several approaches have been proposed to realise the agility and flexibility as well as optimality, and we focus on social oriented approach in this paper.

Social system, which consists of huge number of decision makers, is quite robust against any dynamic environmental change as its nature. We are always trying to seek better situations via micro social interactions in our life. Each micro interaction, i.e. social contract, could emerge adaptive behaviour, which affects macro movement in the entire society. The rational movement should acquire robustness as well as efficiency [2][3].

One of the important aims in social science, such as economics, business management, law, etc., is to clarify the maximum conditions of total welfare in its society. For example, a concept of general equilibrium, where it is assumed that the relevant supply and demand conditions are known and all markets are cleared at equilibrium prices, has been proposed in micro economics. In a market economy, any competitive equilibrium is Pareto optimal: an allocation is a division of all the goods and services among all the people in the economy [4].

As mentioned above, interdisciplinary approaches with engineering and social science must have large potential to realise rational system management with optimality and robustness simultaneously [5][6]. It is obvious that a new concept of "social-oriented manufacturing systems" could take an advantage of such characteristics.

In this paper, we focus on auction mechanism [7], one of the basic social interactions, and try to apply it into one of the important resource allocation problems in manufacturing system, manufacturing scheduling. We propose a new scheduling methodology with combinatorial auction mechanism so as to find an optimal solution with efficient computational performances against NP-hard problems.

2 COMBINATORIAL AUCTION

2.1 General concept

When an auction of multiple items is performed, it is often desirable to allow bids on combinations of items, as opposed to only on single items. Such an auction is often called combinatorial, and the exponential number of possible combinations results in computational intractability of many aspects regarding such an auction. In combinatorial auction, a multiunit auction in which each bidder offers a price for a collection of goods (of the bidder's choosing) rather than placing a bid on each item separately. The auctioneer selects a set of these combinatorial bids which raises the most revenue without assigning any object to more than one bidder [8].

Combinatorial markets where bids can be submitted on bundles of items can be economically desirable coordination mechanisms in multi-agent systems where the items exhibit complementarity and substitutability.

We focus the combinatorial auction mechanism as resource allocation algorithm for manufacturing scheduling in this paper. The relationship between work and jig is defined as the complementarity, and parallel machines are regarded as substitutability in scheduling problems, so the concept of combinatorial markets should be quite affinitive to the scheduling problem.

We divided the mechanism into two modules, such as combinatorial bid creation problem (CBCP) and winner determination problem (WDP). WDP is formulated as general combinatorial optimisation problem in which the total social welfare based on the collected bids is maximised under several constraints shown in Figure 1 (where *k*: process ID).

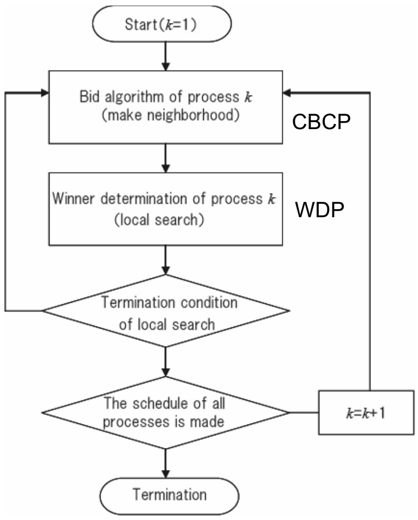


Figure 1: Combinatorial auction algorithm

The combinatorial problem in manufacturing scheduling is classified in NP-hard, so it is required to reduce its searching space for better performance in calculation time. CBCP is used to squeeze the search space rationally based on local utility of bidders. In other words, CBCP is a kind of local optimisation module, and global optimisation is acquired via WDP within the search space created as the aggregation of CBCP. So the sophisticated scheduling algorithm which calculates optimal solutions efficiently is attainable after the social interactions between CBCP and WDP. The algorithm is described in detail at 2.3.

2.2 Formulation

When we formulate manufacturing scheduling problem with the social contract based combinatorial auction mechanism, we classified scheduling problems into 2 types, such as Pull and Push production cases as follows:

Scheduling model 1: Pull production

bidder: process machine

goods: job

Scheduling model 2: Push production

bidder: job

goods: time slot of process machines

Process machines compete the auction to get their goods (jobs) in model 1, while jobs try to seek appropriate process machine according to their utility in model 2. Model 1 and model 2 could naturally emerge the workflow in Pull and Push production systems, respectively.

WDP definition

WDP in each model is defined as follows:

Model 1:

Minimise $$\sum_{m=1}^{M_k}\sum_{j=1}^{B_{k,m}} p_{k,m}^{j} x_{k,m}^{j} \quad (1)$$

sub. to

$$\sum_{m|J_{i,k}\in S_{k,m}^{j}}^{M_k}\ \sum_{j|J_{i,k}\in S_{k,m}^{j}}^{B_{k,m}} x_{k,m}^{j} \le 1 \quad (i=1,\cdots,I) \quad (2)$$

$$\sum_{j=1}^{B_{k,m}} x_{k,m}^{j} \le 1 \quad (m=1,\cdots,M_k) \quad (3)$$

$$x_{k,m}^{j} \in \{0,1\} \quad (m=1,\cdots,M_k),(j=1,\cdots,B_{k,m}) \quad (4)$$

Model 2:

Minimise $$\sum_{i=1}^{I}\sum_{j=1}^{B_i} p_i^j x_i^j \quad (5)$$

sub. to

$$\sum_{i=1}^{I}\ \sum_{j|T_{m,t}\in S_i^j}^{B_i} x_i^j \le 1 \quad (m=1,\cdots,M_k),(t=1,\cdots,T) \quad (6)$$

$$\sum_{j=1}^{B_i} x_i^j \le 1 \quad (i=1,\cdots,I) \quad (7)$$

$$x_i^j + \sum_{s|CT_i^j < BT_r^s \le CT_i^j + st_{w_i,w_r}}^{B_i} x_r^s \le 1 \quad (i=1,\cdots I),(j=1,\cdots,B_i),(r=1,\cdots,I \mid r\ne i) \quad (8)$$

$$x_i^j \in \{0,1\} \quad (i=1,\cdots,I),(j=1,\cdots,B_i) \quad (9)$$

where

M_k : the number of machine in process *k*

I : the number of products

$B_{k,m}$: the number of bids from machine *m* in process *k*

$p_{k,m}^{j}$: bid value of bid *i* from machine *m* in process *k*

$S_{k,m}^{j}$: job set in bid *i* from machine *m* in process *k*

T : the number of time slots

$T_{m,t}$: time slot *t* in machine *m*

$st_{w,v}$: setup time from *w* to *v*

w_i : product type of job *i*

S_i^j : time slot set in bid *j* of job *i*

BT_i^j : process start time slot in bid *j* of job *i*

CT_i^j : process finish time slot in bid *j* of job *i*

In those formulations, Eqs.(2) and (6) describe constraints about treating goods. Winner bid constraint is defined in Eqs. (3) and (7), and sequence constraint is formulated in Eq. (8). Setup time is treated only model 2, because setup time is included and optimised at CBCP in model 1.

CBCP definition

Bidders' behaviour in CBCP is formulated as follows:

Model 1:

Machine utility $$u_{k,m}^{S} = \alpha\left(\frac{P^S}{pt_{k,m}^{S} + st_{k,m}^{S}}\right) + \beta\left(1 - \frac{td_{k,m}^{S}}{T}\right) \quad (10)$$

Bid value $$p_{k,m}^{j} = td_{k,m}^{S} \quad (11)$$

Lower limit of utility $$u_{k,m}^{S} \ge U \quad (12)$$

Model 2:

Job utility $$u_{i,k,m}^{t,t'} = \alpha\left(\frac{P_i}{pt_{i,k,m}^{t,t'}}\right) + \beta\left(1 - \frac{td_{i,k,m}^{t,t'}}{T}\right) \quad (13)$$

Bid value $p_i^j = td_{i,k,m}^{t,t'}$ (14)

Lower limit of utility $u_{i,k,m}^{t,t'} \geq U_R$ (15)

where

$u_{k,m}^S$: utility of machine m for job set S

$td_{k,m}^S$: delayed time against due date

$pt_{k,m}^S$: process time

α , β: weight value (constant)

Eqs. (11) and (14) mean the minimisation of delayed time against due date is the objective function at WDP in both models. We set lower limit about bidders' utility to expect better calculation performance shown in Eqs. (12) and (15).

2.3 Auction algorithm

Combinatorial auction algorithms in two kinds of model are described as follows:

Model 1:

Step1 : Set k=1

Step2 : Auctionner sends new call for auction into all the machines in process k to send the job attributes.

Step3 : Machines in process k create their bids based on their utility (CBCP) and send them to auctioneer.

Step4 : WDP is resolved and optimal schedule is attained in process k.

Step5 : Finish if k=K

Step6 : k=k+1, then go to Step2.

Model 2:

Step1 : Set k=1

Step2 : Auctionner sends new call for auction into all the jobs in process k to send the time slot information of process machines.

Step3 : Jobs in process k create their bids based on their utility (CBCP) and send them to auctioneer.

Step4 : WDP is resolved and optimal schedule is attained in process k.

Step5 : Finish if k=K

Step6 : k=k+1, then go to Step2.

WDP is resolved by commercial solver (ILOG CPLEX 9.1) [9] in this paper.

2.4 Neighbourhood search

WDP is carried out in the search space aggregated by CBCP, and that means single search calculation is a kind of local optimisation approach. Hence several iterations are required to improve the optimality in combinatorial auctions in process k so as to attain optimal schedule.

It is a kind of local search algorithm, and initial value set in iteration i is selected as local optimal solution in the neighbourhood in previous iteration (i-1). The number of iteration is defined as L in this paper.

3 SIMULATION EXPERIMENTS

3.1 Scheduling model

General flexible flow shop is selected as target manufacturing system to validate our social contract based scheduling methodology, shown in Figure 2. Each process consists of several parallel machines with different specification in process conditions.

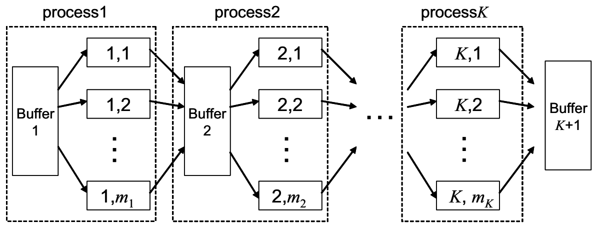


Figure 2: Flexible flow shop model

Where

K: the number of process P_i : production volume of product i

m_k: the number of machines in k

3.2 Experimental parameters

Experimental parameters are listed in Table 1. As the first research to evaluate our proposed algorithm, we assumed a simple but principle flexible flow shop. We have done many kids of experimental scenarios, but we describe some of basic results in this paper. The number of iteration is 10, and 15 trials are carried out at each simulation scenario in all the experiments in this paper.

Table 1: Experimental parameters

K(process)	1	w_i(product type ID)	U(1,5)
M_1(machine)	7	P_i (production volume)	U(1,10)
T(time slot)	100	d_i (due date)	U(30,60)
I(job)	20	$pt_{1,m,w}$ (process time)	U(3,6)
W(product type)	5	$st_{w,v}$ (setup)	U(1,3)

3.3 Experiment 1 : Optimality evaluation

In this experiment, we allow each bidder to create bids with all the combination in goods that means the search space is maximal. As a result, a globally optimised solution must be acquired by WDP in this experiment. The optimality and calculation performance about proposed algorithm are evaluated by this experiment.

Table 2: Optimality evaluation

	TD	MS	Nb	Tp
Model 1	15.53	63.00	2084.9	8.21
Model 2	15.53	63.33	3596.0	53.01

Experimental results are shown in Table 2.

TD: Total delayed time against due date

MS: make span

Nb: Total number of bids

Tp: Calculation time (seconds)

TD is 15.53 at globally optimal solution in this condition, and we have confirmed the proposed algorithm successfully acquires global optimisation in both models with solving combinatorial problems appropriately by this experiment.

As described 2.2, the objective function of WDP is minimisation of delayed time against due date in this paper. Then the MS isn't always converged into optimal solution under the formulation. The number of jobs and timeslots are relatively small in this experiments, then the search space becomes smaller in Model 1 than Model 2. That causes longer calculation time in Model 2.

3.4 Bidder's utility analysis

Bidder's utility is calculated as the weighted value in two parameters described in Eqs. (10) (13). The search space is squeezed with the utility, so it is obvious that it affects a lot on the proposed scheduling algorithm performance. Then we try to evaluate the influence of utilities parameters, α and β, on scheduling performance in this section.

We set U=1.5, $B_{1,m}$=100 in Model 1, R=150, B_i=50. And that means it is not guaranteed to get globally optimal solution because the search space is squeezed according to bidders utility so as to attain better performances in calculation time in this section.

Table 3 and 4 show the results in Model1 and Model 2, respectively. TD_n means the solution at the n-th iteration in neighbourhood search, and TD_{10} is the final solution acquired by the proposed method.

Table 3: Model 1

α	β	TD_1	TD_{10}	MS	Tp
1.5	1.5	24.00	16.53	63.00	15.25
3.0	1.0	24.40	17.13	62.13	13.68
4.5	0.5	25.87	19.80	61.27	14.29

Table 4: Model 2

α	β	TD_1	TD_{10}	MS	Tp
1.5	1.5	27.73	19.80	62.47	32.65
3.0	1.0	30.60	19.93	62.07	34.98
4.5	0.5	32.00	22.33	62.80	36.65

It has been obviously observed that the globally optimal solution about delayed time against due date is not attained in both models as considered.

We have confirmed several general tendencies as flows:

- As the value of β increases, TD decreases in any cases. That is because bidder's utility about due date is getting higher as the value of β is going to larger. That means final TD_{10} is going to closer to globally optimal solution as the value of β increases.

- As the value of α increases, MS decreases in most cases. Parameter α increases the utility about throughput, so it improves make span as it has bigger value.

- Model 1 performs better than Model2. As mentioned before, Model 2 is a kind of Pull productions, and time slots about the machine with higher performance attract bidders (Jobs) greatly. As the consequence, highly performed machine tends to accumulate Jobs, and that causes to increase make span with larger amount of WIP (Work in process).

- Acquired solution is not optimal, but comprehensive and efficient in terms of 2 metrics, such as due date and make span.

- Total calculation times in Table 3 are always longer than Table 2, although search space must be smaller. Calculation process consists of 2 parts: bid generation and optimisation. After the detailed investigation, it has been confirmed that the bid generation requires more calculation than optimisation in this small model. But it is clear that the calculation time becomes shorter in the proposed method as the target system becomes large in size.

Basic analysis has been carried out in those computational experiments. The target model is simple but principal, so general aspects of the proposed method have been clarified in this paper.

4 CONCLUSIONS

First of all, we proposed a new concept about "Social-oriented manufacturing systems", which has interdisciplinary concept between engineering and social science. Then we formulated a scheduling method based on combinatorial auction, which is based on social contract protocol. Several simulation experiments have clarified that our combinatorial auction based methodology conducts comprehensive and efficient solution within practical time.

Our obvious extension is to apply our methodology into large sized flexible flow shop model to validate the applicability into practical case study.

5 REFERENCES

[1] Groover, M. P., 2007, Automation, Production Sytems, and Computer-integrated Manufacturing, Prentice Hall, Upper Saddle River, New Jersey.

[2] Weiss, G., 2000, Multiagent Systems: A Modern Approach to Distributed Artificial Intelligence, MIT Press, Cambridge MA.

[3] Kaihara, T., 2006, A multiagent-based complex systems approach for efficient partnering in virtual enterprise, Information technology for balanced manufacturing systems, Edited by Weiming Shen, Springer, New York, 175-182.

[4] Shoven J. B. and Whalley J., 1992, Applying General Equilibrium, Cambridge University Press, Cambridge. UK.

[5] Wellman, M. P., 1996, A Market-Oriented Programming Environment and its Application to Distributed Multi-commodity Flow Problems, ICMAS-96, 385-392.

[6] Kaihara, T., 2003, Multi-Agent based Supply Chain Management with Dynamic Environment, International Journal of Production Economics, Elsevier Science, Vol. 85, Issue 2, 263-269.

[7] Fujii, S., Kaihara, T., Sahio, K., Yokose, H., Kurahashi, M. And Hayashi N., 2006, A Basic Study of Auction-based Planning and Scheduling for Cell Manufacturing, Lean business systems and beyond, Tomasz Koch Ed., 55-60.

[8] Rassenti V. S. and Buflin, R., 1982, Combinatorial Auction Mechanism for Airport Time Slot Allocation, Bell Journal of Economics, Vol.13, No.2, 402-417.

[9] ILOG, http://www.ilog.com/

Agent-based Dynamic Process Planning and Scheduling in Flexible Manufacturing System

Hossein Tehrani Nik Nejad[1], Nobuhiro Sugimura[1], Koji Iwamura[1] and Yoshitaka Tanimizu[1]
[1] Department of Mechanical Engineering, Graduate School of Engineering
Osaka Prefecture University, Japan

Abstract
This paper deals with an agent-based architecture of an integrated system for process planning and scheduling for multi jobs in flexible manufacturing systems. The process plans and the schedules of the manufacturing resources and the individual jobs are generated dynamically and incrementally. The architecture does not employ a set of predefined process plans, and the process plans of the individual jobs are incrementally and real timely generated from the alternative production routes in the process plan networks. Coordination agents are proposed to generate a suitable the assignment of the job agents to the machine tool agents at each step of the negotiation.

Keywords:
Multi Agent System; Process Planning; Scheduling; Flexible Manufacturing Systems; Mathematical Programming

1 INTRODUCTION

Process planning and scheduling are important manufacturing planning activities which deal with resource utilization and time span of the manufacturing operations. In order to cope with competitiveness and globalization of today's business environment, supply chains become more complex, and manufacturing processes have become more advanced, however products have to be manufactured in higher varieties and smaller batches. It is essential to establish effective and efficient process plans and production schedules to cope with the highly dynamic manufacturing requirements. Some automobile manufactures are gradually adopting their production ways to support the diversity of the customer needs and increase the changing speed for the developing the new products [1].

In recent years, Multi-Agent Systems (MAS) have been widely applied in manufacturing applications because of its flexibility, re-configurability, and scalability [2], [3], [4]. Lim and Zhang [5] have developed an agent-based integrated dynamic process planning and scheduling system to increase the responsiveness of the manufacturing systems. This system does not only effectively integrate the dynamic process planning and scheduling, but also optimizes the machine utilization and provides a platform to assess the reconfiguration of the manufacturing systems. Sugimura et al. [6] proposed a basic architecture for integrated process planning and scheduling from the view points of the distributed decision making. A systematic approach was proposed to select suitable combination of the process plans and to generate suitable production schedules for all the jobs by applying the genetic algorithm and the dispatching rules. Wong et al. [7] proposed an agent-based approach for the dynamic integration of the process planning and scheduling functions. In consideration of the alternative processes and alternative machines for the production of the parts, the

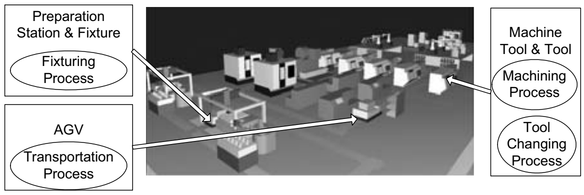


Figure 1. Configuration of the target manufacturing system.

actual selection of the schedule and the allocation of the manufacturing resources is achieved through the negotiation among the part and machine agents which represent the parts and the manufacturing resources, respectively.

The multi-agent architecture has been widely applied to the process planning and the scheduling, as discussed in a review paper [8]. The methods proposed in the literatures deal mainly with the process planning tasks in the static environment in which the product specifications and the manufacturing system status are stable. However, it is now required to develop an integrated process planning and scheduling systems applicable to the dynamic environment in which some unforeseen disturbances may occur.

The objective of the research is to propose a multi-agent based integrated system for process planning and scheduling in the dynamic environment, in order to cope with the product specification changes and the unforeseen disruptions, such as the malfunction of the machine tools. Coordination agents are also proposed to generate a suitable the assignment of the job agents to the machine tool agents at each step of the negotiation. The following issues are discussed in the paper:

1. Target flexible manufacturing systems.
2. Multi-agent system and coordination mechanism for dynamic process planning and scheduling.
3. Flexibility and robustness of the proposed architecture.

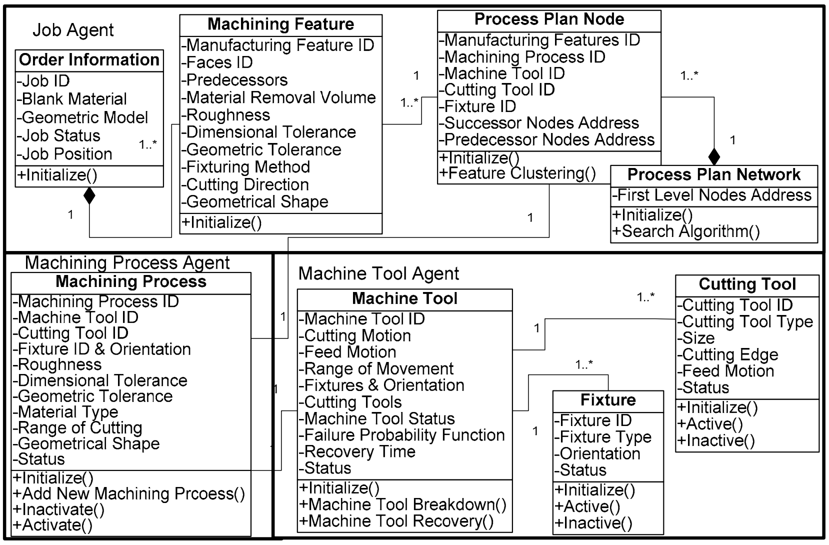


Figure 2. UML Class diagram of physical agents and their relations.

4. Simulation software and experimental results.

2 TARGET FLEXIBLE MANUFACTURING SYSTEM

The basic configuration of the target FMSs are shown in Fig. 1. The manufacturing systems consist of a set of machine tools, preparation stations, input and output buffers, AGVs (Automated Guided Vehicles), fixtures and cutting tools. The input jobs are transformed by a process plan to the finished products through the manufacturing processes. The inputted jobs are firstly fixed on the fixtures at the preparation stations and transmitted to the input buffers by the AGVs. The jobs are then transported to the machine tools in order to carry out the machining processes required. If the refixturing is required to change the positions and orientations of the jobs against the fixtures, the jobs are transported to the preparation stations and the re-fixturing process is carried out at the preparation stations. When all the required machining processes of the jobs are finished, the finished jobs are transported to the preparation stations to separate the jobs from the fixtures, and the jobs are transmitted to the output buffers.

3 MULTI AGENT ARCHITECTURE

A multi-agent architecture is proposed to carry out the dynamic process planning and scheduling. The agent system incrementally generates suitable process plans for jobs, which include the machining sequences of the machining features and the sequences of the machine tools based on the dynamic status of the FMSs. In the integrated manner, it also generates the schedule of the manufacturing resources including machine tools at the same time. In the following sections, the system architectures proposed here are discussed from the viewpoints of the agent definitions and the negotiation protocols and coordination mechanism among the agents.

3.1 Agent Definitions

Two types of agents are considered in this research to develop the process planning systems. They are, physical agents and information agents. The physical agents represent jobs, the manufacturing resources and the machining processes in the FMSs. The manufacturing resources considered here are the machine tools, the preparation station, the AGV (Automated Guided Vehicles), the fixture and the cutting tools. The information agents are virtual agents which are responsible for governing the negotiation protocol and decision making.

Physical agents

The UML class diagram of the physical agents is summarized in Fig.2. The contents of the physical agents are summarized in the following.

Job agents

The job agents represent the jobs to be manufactured in the FMSs. The role of the job agents in the process planning is to certify the correct machining processes of the jobs. The job agents include the following information to describe the orders and the machining features:

1. Job information

The job information section describes the order information, the locations and the progresses of the machining processes of the jobs.

2. Machining features

The machining feature section gives the machining features of the jobs and their technical data such as the types, the tolerances and the roughness. These technical data are required to select appropriate machining processes.

3. Process plan networks

The process plan networks represent the generated process plans in non-linear and hierarchical ways. It includes all the alternative process plans that satisfy the technological requirements of the jobs.

4. Job Status

We consider the following status for the job agents:

- Idle: The job agent is idle and waiting for the next machining operations.

- Machining operation: The job agent is under machining processes on the machine tools.
- Transportation and re-fixturing: The job agent is transported and/or re-fixtured for its next machining operations.

Machine tool agents

The machine tool agents represent the machine tools. The agents representing such resources as the preparation stations and the AGV are not considered, at present, since only the machining processes are discussed in the present research. The machine tool agents are responsible for generating proposals to the machining processes required from the job agents. The proposals include the machining time, the transportation time and the re-fixturing time needed to carry out the required machining processes of the job agents. The machine tool agents include the following information to represent the machine tools in the FMSs:

1. Machine tool information

The machine tool section specifies the shape generation functions, which are represented by the cutting motions, the spindle directions, the feed motions and the maximum product size.

2. Machine tool status

We consider the following status for the machine tool agents in the simulation:

- Idle: The machine tool is idle and negotiating with job agents for next machining operation.
- Machining operation: The machine tool is machining the job agent.
- Breakdown: The machine tool has been broken and is under recovery process.

3. Cutting tool

The characteristics of the cutting tools are described in the cutting tool section, which includes the information about the cutting tool types, the tool sizes and the cutting edge types.

4. Fixture

The fixture section describes the fixture types, and the positions of the fixtures against the spindle axis.

Machining process agents

The machining process agents represent the machining processes of machining features of the jobs, which are carried out by the machine tools. The agents include the following information:

1. Machining process ID which is the combination of the ID of the machine tools, the ID of the fixtures and the ID of the cutting tools.
2. Machining process types and machining features types, which can be generated by the machining processes
3. Surface roughness, tolerances and material removal rate of the machining processes.
4. Machining process status

The status represents the dynamic status of the machining processes in the FMSs. The machining process agents have two statuses.

- inactive: if one of the machine tool, the cutting tool and the fixture related to the machining process are broken-down
- active: otherwise

Information agents

The information agents are virtual agents for governing the negotiation protocol and decision-making.

Production engineering agents

The production engineering agents generate the job agents, the machine tool agents, and the machining process agents to specify the geometric and technological information of the jobs, the machine tools and the machining processes of the FMSs. The agents play a key role for initializing the information of the physical agents.

Job order agents

The job order agents represent the manufacturing tasks. They are information agents, which carry out the negotiation processes between the job agents and the machine tool agents to generate suitable process plans. The agents have crucial influence on the system performance by deploying efficient decision making mechanism to select the appropriate machine tools for the individual machining features of the jobs.

Coordination agent

The architecture of designed multi agent system is distributed, and the benefits inherited from distributed architecture include flexibility and robustness. However, absence of higher authority may result in a lower performance compared with hierarchical systems that are able to achieve global optimization. We introduce a coordination agent to improve the performance. The coordination agents are proposed to a suitable assignment of the job agents to the machine tool agents at each step of the negotiation. The coordination agents make mathematical models according the information sent from the machine tool agents and solve them to find a suitable assignment of the job agents to the machine tool agents.

3.2 Negotiation protocol

A negotiation protocol among the agents is required to coordinate the distributed decisions of the individual agents for solving the integrated process planning and scheduling problems. The problems to be solved here are as follows:

1. Selection of candidate machining processes for individual machining features.
2. Selection of machining resources including machine tools, cutting tools and fixtures.
3. Selection of machining sequence of machining features.

We apply the contract net protocols [9] which are widely used in the agent based scheduling problems. The multi jobs can negotiate with the machine tools simultaneously. In the negotiation protocol proposed here, the individual agents have three types of boards named “request boards”, “proposal boards” and “status boards” for the communication among the agents. Figure 3 shows the role of the boards. The requests from the other agents are firstly sent to the request boards, and the agents scan and receive the requests from the boards every fixed time intervals named RTIP (reading time interval period). The individual agents secondly generate the proposals to the requests, and store them in the proposal boards. The status of the agents are changed and stored in the status board, if necessary. The communication between the boards and the other agents are controlled by a communication system, which is common to all the agents.

Figure 4 summarizes the negotiation protocol proposed here to generate suitable process plans and schedules by the distributed decision-makings of the individual agents and the negotiations among the agents. The negotiation protocol is carried out through the following steps:

Step 1: Initialization

The production engineering agents firstly generate all the job agents and the machining process agents to initialize the status of the FMSs. They integrate the faces of the jobs to define the machining features, which could be generated simultaneously by the same combinations of the machine tools, cutting tools and fixtures, and assign them to the job agents

Step 2: Requests for available machining processes

The job agents select a set of the machining features, which can be machined in the next process, based on the precedence constraints among the machining features. The selected machining features are sent to the machining process agents, and they select a set of available alternative machining processes for the individual machining features, based on the specifications of the machining features, such as the geometries, the sizes, the surface roughness, the tolerances, and the cutting directions. The selected machining processes include the information about the machine tool type, the cutting tools and the fixture orientations. The selected machining processes are sent back to the job agents.

The job agents generate a set of groups of the machining features, which can be machined by the same combinations of the machine tool types, the cutting tools and the fixture orientations. This means that the grouped machining features are machined by one machining process concurrently, therefore, the job agents generate the node representing all the grouped features in the process plan networks. The contents of the process plan networks are described in the previous paper [10]. The individual nodes in the process plan networks represent a set of the machining features machined by one machining process. An algorithm was also presented to generate the nodes representing the machining sequences of the machining features based on the precedence constraints. The information of the generated nodes is sent to the job order agents for the negotiations.

Step 3: Requests generation by the job order agents

The job order agents create requests for the individual nodes of the process plan networks, which are the groups of the machining features machined by same machine tools, and send all the requests to the request board of the corresponding machine tool agents.

Step 4: Proposal preparation by machine tool agents

The machine tool agents read all the requests from the request boards every RTIP (Reading Time Interval Period). The machine tool agents analyze the request messages, and generate appropriate proposals to all the requests.The proposals include the estimated minimal completion time of the remaining machining features of the job agents. A procedure is developed and given to the job agents to estimate the minimal completion time of the remaining machining features. The job agents start the procedures from the start node which is specified by the machine tool agent, and repeat to generate and to select suitable successive nodes with the minimum machining time. When all the machining features are included in the process plan networks, the job agent find both the machining sequences of the machining features and the estimated minimal completion time. The generated proposals are sent to the coordination agents.

Step 5: Selection of appropriate proposals by coordination agent

The coordination agents scan all received proposals from the machine tool agents every RTIP, and assign the appropriate machine tool agents to the job agents. At present, we consider only the flow time of the job agents, and our goal is to minimize the average flow time of all the job agents. The flow time considered here includes the machining time, the transportation time, the refixturing time and the tool changing time. The followings summarize the formulas representing the optimization problems considered here.

Parameters:

$$MP = \{mp_r = (mt_j, fx_f, ct_t) | r = 1,......R\},\ R = |MP| \quad (1)$$

$$MT = \{mt_j | j = 1,2,....n\}\ ,\ n = |MT| \quad (2)$$

$$FI = \{fi_f | f = 1,2,....F\}\ ,\ F = |FI| \quad (3)$$

$$CT = \{ct_t | t = 1,2,....T\}\ ,\ T = |CT| \quad (4)$$

where,

mp_r : ID of machining process.

mt_j : ID of machine tools.

fi_f : ID of fixtures.

ct_t : ID of cutting tools.

$FT^j_{(i,r)}$: Estimation of completion time of job agent *i* (i = 1,2,..m) according to the machining process mp_r (r = 1,2,..R) with machine tool agent mt_j (j = 1,2,..n).

Design variables:

$x^j_{i,r}$ =
- 1: if the machine tool agent mt_j (j = 1,2,..n) is selected for job agent *i* (i = 1,2,..m) according to the machining process mp_r (r = 1,2,..R).
- 0: otherwise.

Mathematical Model:

$$\text{Minimize } Z = \sum_{j=1}^{n}\sum_{i=1}^{m}\sum_{r=1}^{R} x^j_{(i,r)} \times FT^j_{(i,r)} \quad (5)$$

$$\sum_{j=1}^{n}\sum_{r=1}^{R} x^j_{(i,r)} + Dummy_i = 1 \qquad i = 1,2,...m \quad (6)$$

$$\sum_{i=1}^{m}\sum_{r=1}^{R} x^j_{(i,r)} + Dummy_j = 1 \quad j = 1,2,...n \quad (7)$$

$$x^j_{i,r} = 0 \text{ or } 1 \quad (8)$$

We add dummy variables to equations (5), and (6) to change the constraints to sets of equations. Equation (5) is the objective function that is total of the estimated flow time of all

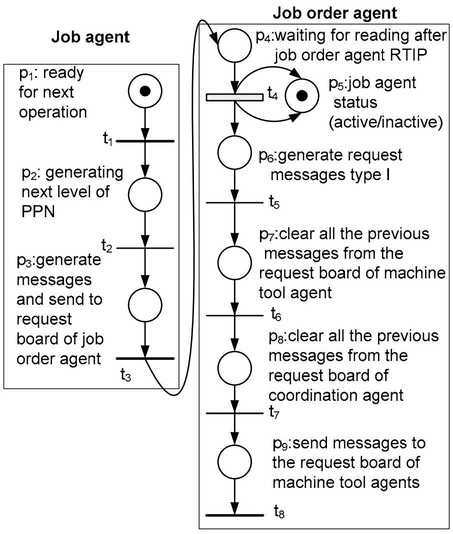


Figure 3. Synchronizing the job agent and job order agent for generating request for machine tool agents.

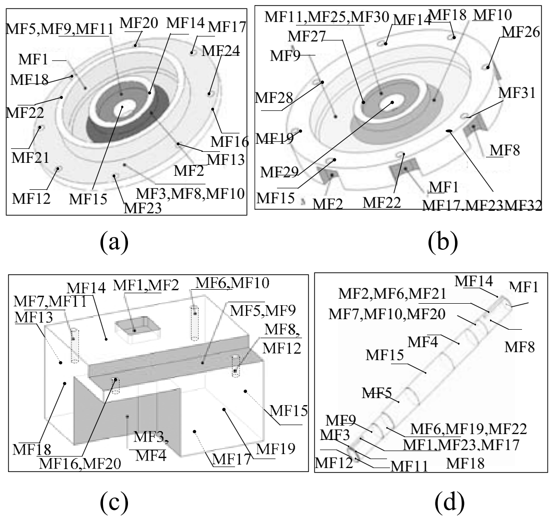


Figure 4. Jobs considered in case studies.

the job agents. Equation (6) is a constraint that only one machine tool agent is selected for each job agent. Equation (7) is a constraint that only one job agent has been assigned to each machine tool agent. The model described in equations (5), and (6) is an integer programming model and can be solved using a commercial optimization packages.

After solving the above coordination problem, the coordination agents inform both the job agents and the machine tool agents that the machining features sent from the job agents shall be machined by the selected machine tools. This means that the coordination agents dynamically generate the process plans and the production schedules of the job agents and the machine tool agents.

The job agents and the machine tool agents selected here carry out the requested machining processes in the next step. Therefore, the statuses of these agents are changed, and the status data are stored in the status boards. All the agents monitor the status data if necessary.

Step 6: Preparation for next operation

When the machine tool agents complete the machining operations of the job agents, the job agents modify their process plan networks. That is, the job agents delete the corresponding nodes representing the group of the machining features which was completed by the machine tool agents. New nodes of the process plan networks are generated to specify the groups of the machining features to be machined in the next step.

The procedures presented in Step 2 to 6 are repeated until all the nodes of the process plan networks are machined and the job agents do not have any remaining machining features to be machined.

4 SYNCHRONIZATION

The Petri nets [11] are used, in the research, for synchronizing the messages and the negotiation protocols between the different agents. This Petri nets control both the sequence and timing of the interaction and the messages between the agents. Each Petri net represents one agent or two interacting agents. Figure 3 shows an example of the interaction between the job agents and the job order agents for generating the requests and send them to the request board of the machine tool agents. These Petri nets are linked with each other with global transition (See transitions, t_1, t_3, t_8 in Figure 3).

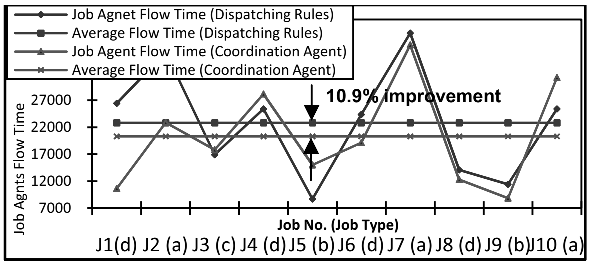
5 SIMULATION AND CASE STUDY

A prototype of the agent based integrated process planning and scheduling system and the graphic system have been developed for the case studies. The system developed here is able to simulate the distributed decision makings of the agents, the negotiation processes among the agents, and also the manufacturing processes in the FMS. The coordination agent use ILOG CPLEX optimization engine [12] for solving the integer programming model of the coordination and for assigning the job agents to machine tool agents.

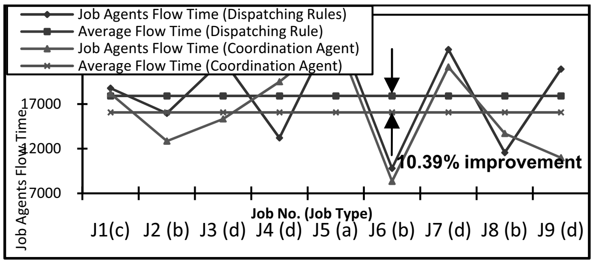
Some case studies have been carried out to verify the applicability and the effectiveness of the proposed system to the integrated process planning and scheduling problems in the FMSs. The FMS considered here includes 7 machine tools and 4 job types. Figure 4 shows the geometries of the job agents for the case studies. The detailed information of the machining features and the machining resources of the case studies are brought in the previous paper [10]. The RTIP in the simulation is set to be 2 sec. for the machine tool agents, 3 sec for coordination agents and 4 sec. for the job agents.

Two case studies have been done to evaluate the impact of introducing the coordination agents in multi agent systems. We compare the results with the previous research [10] which does not consider the coordination agents. In the previous research, the jab agents applying dispatching rules for selecting the machine tools for their manufacturing operations without assisting from coordination agents.

Figure 5 summarizes the comparison of the proposed architecture and the previous method from the view points of the average flow time of all the job agents. In the Figure 5,



(a) Case study with 10 job agents.



(b) Case study with 9 job agents.

Figure 5. Case study and comparison with previous result.

the vertical axis gives the flow time of the individual job agents and the horizontal axis shows the individual job agents and their types.

It is understood, from Figures 5(a) and 5(b), that the multi agent systems with coordination agents generates more suitable process plans and schedules from the viewpoint of the average flow time of the all the job agents. As you can see, the average flow time has been improved 10.9% and 10.39% for cases (a) and (b) of Figure 5, respectively. It is because that the mathematical programming methods applied here are suitable to reduce the average flow time of the job agents of the job shop process planning and scheduling problems.

6 CONCLUSION

A multi-agent based system was proposed for the integrated process planning and scheduling systems for the FMSs. The following remarks are concluded.

(1) A multi-agent system consisting of six basic agents and a negotiation protocol among the agents were proposed to carry out the various tasks in the process planning and the scheduling. The individual agents have the capability for the distributed decision-making and the communications with the other agents.

(2) A systematic procedure was proposed to generate suitable process plans of the jobs and suitable schedules of the machine tools. The proposed method is able to solve the process planning and scheduling problems concurrently and dynamically, with use of the mathematical optimization methods and search algorithms of the process plan networks.

(3) Some case studies have been carried out to verify the applicability of the proposed method to the integrated process planning and scheduling problems in the FMSs including 7 machine tools and 10 jobs. It was shown, through the case studies, that the proposed multi-agent architecture is capable to generate appropriate process plans and schedules.

7 REFERENCES

[1] Nakashima, T.: Innovation of Preparation for production process of new car by digital manufacturing. JSME Conference No 06-19, (2006) 1-3

[2] Go, P., Balasubramanian S. and Norrie D.H.: Bidding-based process planning and scheduling in a multi agent. Compu. Ind. Eng., 32, 2 (1997) 477-496

[3] Sugimura, N., Tanimizu, Y., Iwamura, K.: A Study on Real-time Scheduling for Holonic Manufacturing System. CIRP Journal of Manufacturing Systems, 33, 5, (2004) 467-475.

[4] Iwamura, K., Okubo, Y., Tanimizu, Y., and Sugimura, N.: Real-time Scheduling for Holonic Manufacturing Systems Based on Estimation of Future Status. International Journal of Production research, Taylor ^ Francis, 44 [18-19] (2006) 3657-3675.

[5] Lim, M.K. and Zhang, D.Z.: An integrated agent based approach for responsive control of manufacturing system. Computers and Industrial Engineering 46, (2004) 221-232.

[6] Sugimura, N., Shrestha, R. and Takemoto T.: A Study on Integration of Process Planning and Scheduling for Holonic Manufacturing. CIRP-Journal of Manufacturing Systems, 6 , 34 (2005).

[7] Wong, T.N., Leung, C.W., Mak, K.L. and Fung, R.Y.K.: Dynamic shopfloor scheduling in multi-agent manufacturing systems. Expert Systems with Applications, 31, 3 (2006) 486-494

[8] Shen, W., Wang, L. and Hao, Q.: Agent-Based Distributed Manufacturing Process Planning and Scheduling: A State-of-the-Art Survey. IEEE Transaction on System, Man, and Cybernetics-Part C: Application and Reviews, 36, 4 (2006) 563-577.

[9] Smith, R. G..: The contract net protocol: High-level communication and control in a distributed problem solver," IEEE Trans. Comput., C-29, 12 (1980) 1104-1113.

[10] Tehrani, H., Sugimura, N., Tanimizu Y. and Iwamura, K., "A Search Algorithm for Generating Alternative Process Plans in Flexible Manufacturing System," Journal of Advanced Mechanical Design, System, and Manufacturing. Vol. 1, No. 5, pp. 706-716 (2007).

[11] Proth, M. and J., Xie, X. 1996, Petri Net a Tool for Designing and Management of Manufacturing System, John Willey and Sons.

[12] http://www.ilog.com/products/cplex/

GA Based Reactive Scheduling for Aggregate Production Scheduling

Tatsuhiko Sakaguchi[1], Toshihide Kamimura[2], Keiichi Shirase[2] and Yoshitaka Tanimizu[3]
[1] Organization of Advanced Science and Technology, Kobe University, Japan
[2] Department of Mechanical Engineering, Graduate School of Engineering, Kobe University, Japan
[3] Department of Mechanical Engineering, Graduate School of Engineering, Osaka Prefecture University, Japan

Abstract
A reactive scheduling method is one of the scheduling methods that improve an initial production schedule disturbed by environmental changes in manufacturing systems. In the previous research, Genetic Algorithm based reactive scheduling method have been proposed. In case of dealing with the aggregate production schedule, it is difficult to modify the initial schedule because of the unknown factors which occur in the manufacturing system. In this research, only part of the initial schedule is modified by setting the suitable scheduling range. Consequently, the reactive scheduling method can be applied for the aggregate production scheduling.

Keywords:
Reactive Scheduling; Genetic Algorithm; Aggregate Production Scheduling

1 INTRODUCTION

A production schedule is predetermined before the execution of the manufacturing processes in the manufacturing systems. However, the unscheduled disruptions such as the delays of manufacturing processes or the addition of jobs often occur in the manufacturing systems, and an initial production schedule may not satisfy the constraints due to the disruptions. A dynamic scheduling method is therefore required to cope with the unscheduled disruptions. The reactive scheduling method is one of the dynamic scheduling methods that satisfy this requirement [1].

A reactive scheduling method based on the Genetic Algorithm (GA) [2] was proposed, in the previous research, to modify the initial production schedule when the manufacturing processes delay in the manufacturing systems. In this method, all operations that have not started will be subjected to the rescheduling and thus their positions on the schedule may be changed. However, in the actual manufacturing systems, it is not so easy to change the schedule because it takes long time to change the tools, the NC programs or the transportation route.

Furthermore, if the target for the scheduling is the aggregate production schedule, the more future jobs that based on forecast or loose contracts the schedule includes, the more difficult it will be to solve the scheduling problem with all those unknown parameters.

In this research, only part of the initial schedule is modified by setting the suitable scheduling period. Consequently, the reactive scheduling method can be applied for the aggregate production scheduling efficiently.

2 DETERMINATION OF SCHEDULING RANGE

In the previous research, the reactive scheduling method modifies the initial production schedule by considering all operations which has not started as shown in Figure 1. The

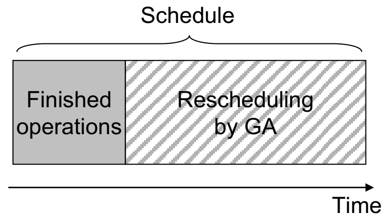


Figure 1: Scheduling range of previous research

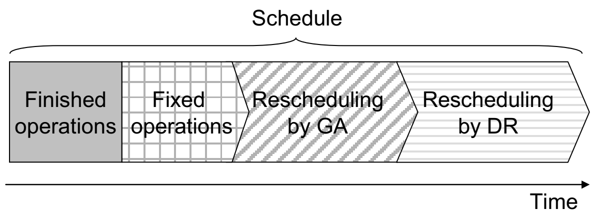


Figure 2: Proposed scheduling range

schedules of the operations which have already started in the manufacturing system are fixed. And the subject for rescheduling is all operations which have not started. However, the larger the number of the operations, the more difficult it will be to modify the schedule. Moreover, when dealing with the aggregate production scheduling, the more future jobs that based on forecast or loose contracts exist in the schedule, the more difficult it will be to solve the scheduling problem with all those unknown parameters. Meanwhile, the nearest operations can not have their schedule easily changed because of the constraint of the set-ups. Therefore, in this research, the suitable scheduling ranges are determined in order to classify the schedule into three blocks as follows:

- The schedules of operations that cannot be changed.

- The schedules of operations that are changed by using GA.
- The schedules of operations that are changed by dispatching rule.

In the first range, the schedules of all operations are fixed because of the set-ups constraints. In the second range, the schedules are changed by using GA as the same way as the previous method. In the third range, the schedules are changed by using dispatching rules periodically. The proposed scheduling ranges are shown in Figure 2. The length of each scheduling range is changed dynamically.

2.1 Definition of processing time

In this research, the operation delays are considered as the disruptions which occurred in the manufacturing systems. It is assumed that the processing time in the manufacturing system can be described with the normal distribution. Therefore, the finishing time of the current executed operation vary as shown in Figure 3. The finishing time is calculated by the following equation.

$$FT = ST + PT \times NRN_{(\mu,\sigma^2)} \tag{1}$$

where,

FT: Finishing time

ST: Starting time

PT: Standard processing time

$NRN_{(\mu, \sigma)}$: Normal random number

μ: Mean

σ_2: Variance

2.2 Rescheduling by GA

Estimation of processing time

In the previous research, a reactive scheduling method modifies the pre-determined initial schedules when operation delay occurs. However, it was assumed that the delay is detected before the delayed operations start in the manufacturing systems. The scheduling system needs to be informed when the operation delay occurs and how long the delay is, before the operations start. In the actual manufacturing processes, it is difficult to predict the length of the delay. The progress of manufacturing operations has to be monitored in order to modify the initial schedule in the context of the actual situation on the manufacturing systems. Moreover, in order to modify the initial schedule when delays occur, the estimated processing time has to be determined.

In this research, the reactive scheduling system detects the operation delays and early finishes, and determines the estimated processing time by the following steps. Figure 4 shows the process of these steps:

1. The reactive scheduling system receives the signal of the starting and finishing time which is sent from the components of the manufacturing system.
2. Three cases are considered in the detection of delay of manufacturing processes. They are:
   (1) The manufacturing operations are executed on schedule in the manufacturing system.

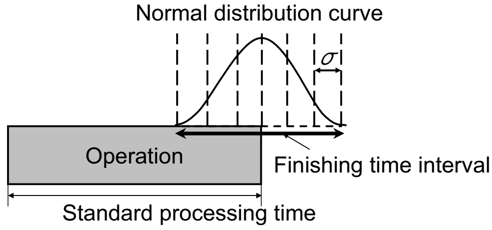


Figure 3: Processing time based on normal distribution

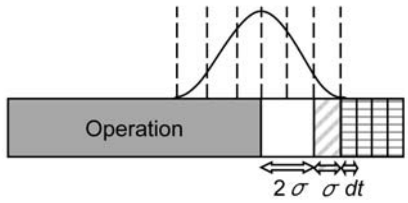


Figure 4: Estimation of processing time

   (2) The manufacturing operations are finished before the standard processing time. In this case, the processing time is fixed by Step 5.
   (3) Delay occurs in the manufacturing system.

   For case (1) and (2), the reactive scheduling process is not executed. On the other hand, for the case (3), the reactive scheduling process is activated in order to modify the initial production schedule.
3. If the reactive scheduling system detects the delay of manufacturing processes, the reactive scheduling system sets on an estimated processing time in order to modify the initial production schedule, because it does not know when the delayed operation will finish. The estimated processing time of that operation is calculated by using the following equation.

$$EPT = PT + 2\sigma \tag{2}$$

   Where,

   EPT: estimated processing time

   PT: initial processing time on the schedule

   By the normal distribution, the probability of the random variable exists in 2σ is 95.44%. Therefore, in this research, the estimated processing time is calculated based on 2σ.
4. The reactive scheduling process is repeated until the modified production schedule satisfies the constraint on due date or all the manufacturing operations finish.
5. If the finishing signal is received from the manufacturing system, the processing time will be fixed and applied to the current schedule. On the other hand, if the finishing signal is not received within ($PT + 2\sigma$), the estimated processing time will be updated with the following equation, and the reactive scheduling process is iterated from Step 4.

$$EPT = PT + 3\sigma \tag{3}$$

   Furthermore, if the finishing signal is not received within the ($PT + 3\sigma$), the estimated processing time will be updated with the following equation.

$$EPT = PT + 3\sigma + dt \tag{4}$$

where dt is the calculation time for one reactive scheduling cycle.

GA based reactive scheduling method

In the scheduling range which schedules are changed by using GA, the variation of the processing time is considered for the criterion of rescheduling. When the operation delays are detected and the production schedule cannot satisfy the constraint of due date, the reactive scheduling process is executed. The reactive scheduling processes consist of the following five steps:

1. The present time T_i (i=1, 2, ...) is set up.
2. The computation time dt is the time in which GA creates a new generation of the populations representing the modified production schedules. The time dt is estimated based on the time needed to generate a new population in GA based current production scheduling process, and it is modified based on the time for creating the modified production schedules through Step 3 to Step 5.
3. Two cases are considered in the creation of the initial population constituted of the individuals which represent the production schedules. They are:
   (1) First activation of the reactive scheduling process at time T_1.
   (2) Second or later activations of the reactive scheduling process at time T_2 or later.

   For the cases of (1) and (2), the reactive scheduling process creates the initial population through the following two steps, respectively:

   - In the case (1), the reactive scheduling system has only the initial production schedule. Therefore, the reactive scheduling process generates the initial population based on the initial production schedule as shown in Figure 5.
   - In the case (2), the reactive scheduling process can inherit the population created in the previous reactive scheduling process. In other words, the last population of the previous reactive scheduling process can be the initial population. Two cases are considered for the inheritance process of the population as shown in the followings:

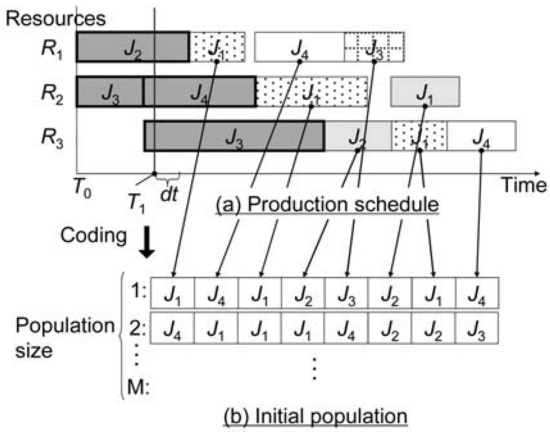


Figure 5: Creation of initial population

   (1) If no operations start between T_i and (T_i + dt), all the individuals of the last population of the previous reactive scheduling process are inherited to a new reactive scheduling process between T_i and (T_i + dt).
   (2) If some operations start between T_i and (T_i + dt), the production schedules of these operations should be fixed. Therefore, a new reactive scheduling process can inherit only the individuals, which are consistent with the schedules of the fixed operations, from the last population created in the previous reactive scheduling process. The other individuals are deleted, and new individuals are randomly created from the inherited ones.

4. The fitness value of each individual is evaluated. The objective function value of the production schedule, which has to be minimized, is selected as the fitness value. In this research, the objective of the scheduling is to minimize the total tardiness. Based on the fitness value, the genetic operators, such as selection, crossover and mutation, are applied to the individuals of the population created in Step 3, in order to create new individuals of the next population.
5. If the shortest objective function value of all the new individuals created in Step 4 is shorter than the objective function value of the current production schedule, the new modified production schedule is substituted for the current production schedule. If the tardiness of the new production schedule is shorter than the constraint on the due date, the reactive scheduling process is terminated.

All the steps from 1 to 5 are repeated, until the created production schedule satisfies the given constraint on the due date or all the manufacturing operations have started in the manufacturing system.

2.3 Rescheduling by dispatching rule

In the scheduling range of which schedules are changed by using dispatching rule, the addition of the new jobs is considered for the criterion of rescheduling. When the new jobs are inserted, the schedules of the operations, including the new operations of the new job, are changed by dispatching rule. W(S/RPT + SPT) rule, which is effectiveness for tardiness minimization problems, is applied in this research.

3 COMPUTATIONAL EXPERIMENTS

3.1 Scheduling conditions

A prototype of reactive scheduling system has been implemented by using an object-oriented programming language, Smalltalk and applied to the computational experiments for the job-shop type scheduling problems. Numbers of jobs, manufacturing equipment and operations considered in the experiments were 70 (10 jobs x 7 units), 10 and 700, respectively as shown in Figure 6. One unit of job consists of ten jobs and each unit has the same set of jobs. Parameters of GA, such as population size, crossover rate and mutation rate, were 50, 0.5 and 0.1, respectively. These parameter values were estimated based on preliminary case

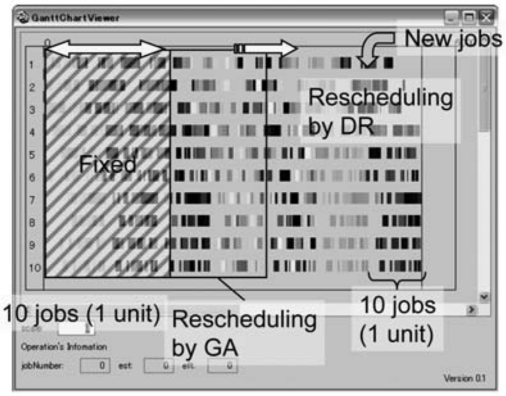


Figure 6: Gantt chart of initial schedule

studies of the job-shop type production scheduling problems. Parameters of normal distribution, such as μ and σ were 0 and 10 % of standard processing time respectively.

The manufacturing system emulator implemented by using Smalltalk sends the signal to the scheduling system when each operation starts and finishes. The scheduling system recognizes the operational delay by comparing the initial schedule to the emulator signal. If the starting time and finishing time on the schedule are equal to the time when the signals were sent, the manufacturing processes was executed in accordance with the schedule. On the other hand, if those are not equal, the operational delay occurred. Therefore, scheduling system execute the reactive scheduling process based on the proposed method. Moreover, three units of new jobs were inserted. The proposed scheduling range was changed dynamically. For example, as shown in Figure 6, the first range which the schedules are fixed will be spreading over time. Consequently, the second range which the schedules are change by GA will move forward.

3.2 Results and discussions

Seven sets of experiments were carried out. In the first three experiments, the suitable length of the GA based scheduling range was verified. The length was set as one-seventh, two-seventh and three-seventh of the make-span of the schedule, respectively. In the second experiment, the initial schedule was modified by using the previous method, which modifies the schedule only by using GA. In last three experiments, three types of dispatching rules were applied for modifying the initial schedule. Figure 7 shows the results of the experiments. X axis represents the method, and Y axis represents the total tardiness (Simulation Time). Through the experiments, the proposed method can modify the initial schedule more effectively than the other methods. However, when the length of the GA based scheduling range was set as three-seventh, the proposed method was less effective than the two dispatching rules. This shows that it is important to determine the suitable scheduling range.

4 CONCLUSIONS

The reactive scheduling method for aggregate production schedule has been proposed in this research. The schedule is divided into three ranges, which is the fixed range, the GA based rescheduling range and the dispatching rule based rescheduling range. The scheduling objective is to minimize the total tardiness. The prototype of reactive scheduling system was implemented, and some computational experiments were carried out. By analyzing the results of the experiments, the suitable scheduling range is verified for aggregate production schedule.

5 ACKNOWLEDGMENTS

A part of this work was supported by the Grant-in-Aid for Scientific Research (B) No. 17360063 of the Japan Society for the Promotion of Science.

6 REFERENCES

[1] Smith, S.F., 1994, OPIS: A methodology and architecture for reactive scheduling, intelligent scheduling, Morgan kaufmann publishers, 29-66.

[2] Sakaguchi, T., Kamimura, T., Shirase, K., Tanimizu, Y., 2007, Reactive Scheduling Based on Status Monitoring for Dynamic Manufacturing Environment, Proc. of 40th CIRP-ISMS, CD-ROM.

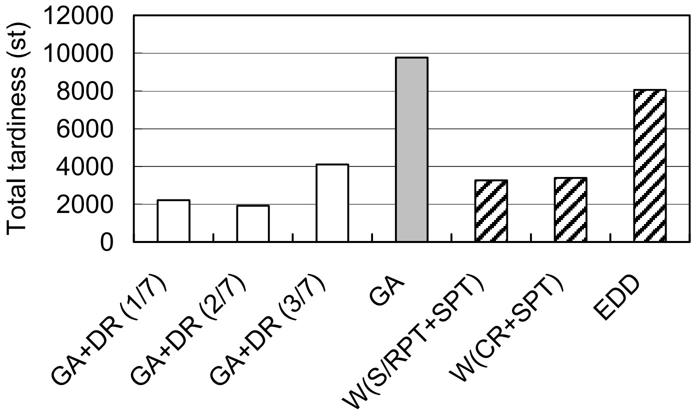


Figure 7: Results of computational experiments

Development of a Multi-Agent Model for Production Scheduling in Innovative Flexible Manufacturing System

Ilaria Baffo[1,2], Giuseppe Confessore[1] and Giuseppe Stecca[1]

[1] Istituto di Tecnologie Industriali e Automazione, Consiglio Nazionale delle Ricerche, Rome, Italy

[2] Dipartimento di Ingegneria dell'Impresa, Università degli Studi di Roma Tor Vergata, Rome, Italy

Abstract

The reduction of a Manufacturing System Complexity is possible adopting a decentralized approach focused on optimization at operation level, and on coordination at higher level of production plant. This point of view, inherited by multi agent models, can increase the flexibility and reduce the complexity of the system management. We apply this principle to a shoe manufacturing system, characterized by an automated transportation who, governed by scheduling rules, can exploit the overtake capability. The operational objectives of the plant are to augment the capacity of the system, to rapidly react to market changes and to obtain a better balance machine work load.

Keywords:

Multi-Agent Model; Production Scheduling; Flexibility; Complexity.

1 INTRODUCTION

The concept of Flexible Manufacturing System (FMS) was introduced in response to the need for greater responsiveness to changes in products, production technologies and markets and has be discussed deeply in literature [1, 2, 3, 4]. FMS have high degree of complexity and often are underused mostly due to lack in software systems and communication technologies able to overcome this hurdle. For this reason it is common to analyze the FMS along two different dimensions: the flexibility end complexity. The first can be analyzed as *internal flexibility*, as the ability to manage in efficient way the plant, and *external flexibility*, as the ability to quickly respond to the market requests. The latter (the complexity) is instead measured in terms of *(i)* plant complexity and, (ii) information domain complexity. The first is a indicator of the number of machines, of products, and of product models [5]. The second one is a function of total quantity of information, information diversity, and of information content, corresponding to the effort to capture and to traduce in useful format the information [6]. The aim of this work is to present an application of a Multi-Agent System (MAS) able to solve in a short computing time a scheduling production problem with a high level of flexibility and complexity. In particular we present and confront a centralized and decentralized approach to solve the previous problem with growing level of information domain. The paper represents a breakthrough step further with respect to the work [7], and it is arranged as following. In the first section we present a discussion about the several approaches of modeling a complex system as a FMS. In the second section we introduce the case study and demonstrate the advantages of use a multi-agent model in order to not increase the algorithmic complexity of the model. In the Section 3 an application to a shoe manufacturing plant is given. Finally, in the last section concluding remarks are discussed.

2 A MAS MODEL FOR COMPLEX SYSTEM

2.1 FMS as a complex system

In our work we refer to FMS as a complex system in which the interaction among the elements of the system, and the type of interaction between the system and the environment, do not enable the nature of the system to be understood simply by analyzing its components [8]. For these reasons a discussion about the best efficient approach to model a behavior of components and of system with a high level of complexity it is open. Often the Petri nets, offering a rigorous and analytic method, are preferable for modeling the control structure of distributed systems. However, the use and application of this model is limited by two problems: (i) the Petri net model become too large and complex even for a modest size problem [9]; (ii) the difficulty to make even a little change to a previously built model. The MAS approach aim to overcome these limits of rigidity and computing complexity in the FMS modeling [10]. The advantages offered by the MAS are: *(i)* the possibility to solve complex problem solving a set of easier local problems; *(ii)* the opportunity to change some problem parameters or to substitute any system element without discard entirely the original model. The difficulty of apply this model approach reside into the activity of architecture design used by the system elements to speak and act. The number and the kind of relations among the system entities became the most important indicator of the complexity model. In this work we want demonstrate that, applying a MAS approach to the same architecture, an increase of information quantity:

1. not increase the computing complexity of the problem;
2. increase the internal and external flexibility of the plant;
3. increase the proximity of the model respect to real system.

3 A MAS MODEL FOR SCHEDULING PROBLEM IN FMS

3.1 Plant Description

We consider an agile shoe manufacturing plant with innovative transportation line. The innovative molecular structure of the transportation, displayed in Figure 1, allows the products to overtake along production lines, increasing the overall flexibility of the plant [11]. The basic element of the molecular structure is the "Tern" (see Figure 2), which is constituted by two rotating tables, called "Table" and "Island", and by a rotating three arms manipulator. The Terns move the parts to be worked while the workmanships are performed by the machines collocated around the Islands. An initial warehouse inserts the forms in the system, and a final storehouse picks up the produced shoes. The Table is used to direct the semi-finished shoes either to the next Tern or to the Island of the same Tern. Each table houses *twelve slots* on which the semi-finished shoes can be placed. Moreover, it moves backward the lasts flowing back towards the warehouse (the last is the object around which the semi-finished shoe is built upon). The Island directs the semi-finished shoes towards the different machining stations, laid around the Island itself. Each island instead houses *twenty-four slots*.

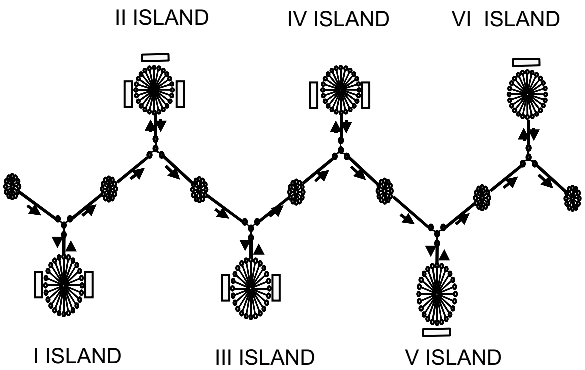


Figure 1: Transportation Line of the plant

The processing time of the jobs on several island can be calculate as following:

- In a deterministic way with a T_{det} formula;
- In a not deterministic way using a average between the case with the isle full of jobs and the case with the isle empty.

In a determinist way we use a T_{det} for calculate the processing time of the jobs on the isles:

$$T_{det} = [N(j\text{-}k) + N(i\text{-}k)] * \max\{t(1), t(2)\} + [N(i) + N(j) - (N(j\text{-}k) + N(i\text{-}k))] * t(1) + [N(i) + N(j) + N(k) - (N(j\text{-}k) + N(i\text{-}k))] * t(2) + [(N(r) + (N(i) - 1)) - [N(j\text{-}k) + N(i\text{-}k)] - [N(i) + N(j) - (N(j\text{-}k) + N(i\text{-}k))] - [N(i) + N(j) + N(k) - (N(j\text{-}k) + N(i\text{-}k))] * t(s) \quad (1)$$

Where:

- $N(r) = 24$ is the total number of island's slots;
- $N(i)$ is the set of jobs to sequence;
- $N(j)$ is the set of jobs that are on the island before that the first machine;
- $N(k)$ is the set of jobs that are on the island between the machine 1 and 2;
- $N(i\text{-}j)$ is the set of couples belonging to set $N(i)$ and $N(k)$ and distant among them 8 slots.
- $N(j\text{-}k)$ is the set of couples belonging to set $N(j)$ and N(k) and distant among them 8 slots.
- $T(1)$ is the job processing time on the machine 1. It is the same for every job.
- $T(2)$ is the job processing time on the machine 2. It is the same for every job.

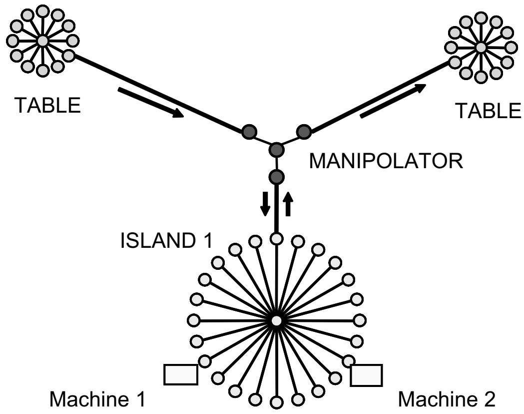


Figure 2: Island – Cell Manufacturing

The formula (1) is applied to every island of the plant and the total processing time is calculated both in approximated or deterministic way.

4 THE MULTI AGENT MODEL

4.1 The Centralized Approach

In the centralized model the enterprise information system acts as the Coordinator Agent (CA) having total decisional and communication power. The CA receives from the market the orders consisting of the specification of products to be produced and the quantity, its due dates, and the release dates of the jobs. The CA has a complete information about the plant characteristics, such as the all parts processing time, computed in approximated or deterministic way. Using this information it can solve the scheduling problem and communicate the starting times at the first island. The best sequence of jobs to be processed is chosen using the EDD (earliest due date) rule. The rule sequences the jobs in not - decreasing order of their due dates and gives an optimal sequence for the single machine maximum lateness problem (1||Lmax). This scenario do not exploit the overtake capability of the system: the flexibility offered by the transportation line is nullified by the decision system rigidity.

4.2 The Decentralized Approach

In this model there are different intelligent agents, each one with an incomplete information about the plant and environment configuration. The agents are connected through a local network and they can communicate and cooperate in order to reach a shared solution. Therefore the negotiation is not necessary because the agents are not competitors. They are cooperative actors that share local information to estimate the goodness of different solutions. There are two types of agents: The Coordinator Agent (CA) who represents the

Information System and the Island Agents (IA) who represent the manufacturing islands. CA has coordination role and it decides if to make changes to the already released production plan based on changes in market request in terms of due dates. For taking this decision it minimize the following decisional function:

$$f = \alpha * \Delta T + \beta * \Delta Cmax \quad (2)$$

where α is a parameter who depends on order priority and β is a parameter depending on how many jobs must be processed, *ΔT* is the change in tardiness of the jobs and *ΔCmax* is the change in maximum completion time after schedule change. Each IA can send or receive messages only to coordinator agent or to neighbor islands. It decide the sequence of jobs to process solving the 1|r|Lmax problem, who consider release dates and it is solved by the agent using Branch & Bound. Agents send and receive messages to share the information and to obtain one shared solution using the overtake coordination protocol (OCP).

The OCP consist of the following steps:

1. The CA receives orders from the market and communicate at the first IA the release and due date of the jobs.
2. The IA solve a local scheduling problem and communicate to next Island Agent when it can start to process the jobs.
3. If due date of one job change during production, then the CA communicate the new due date to the Island Agent following to Island Agent who is working the changed job.
4. The CA find the best schedule and send to CA this schedule with ending times for every job.
5. The CA calculates the convenience of changing schedule using the function (2) and, if overtake is convenient, sends overtake order to the IA.
6. The last IA process the jobs and sends the products to finished product warehouse.

5 TEST RESULTS

We implemented a simulation model using java programming language. We present two exemplificative scenarios of job due date change after the production plan is released to the plant, showing the different decision strategy of the Coordinator Agent (CA) using the deterministic and approximate approach. In the first scenario the due date of job 3 change after the first island and the processing time is calculate as average of these two case: island full and empty of jobs. In the second scenario the change of the due date job is the same, but the processing time is deterministic and obtain applying to every island the formula (1).

Table 1: Due date of three jobs at t=0 and t=1.

Job	**dd (t=0)**	**dd' (t=1)**
1	300	300
2	350	350
3	400	300

The table 1 displays the job information in the two scenarios. In particular, job number, due date before and after the change are reported.

Table 2: Model simulation.

No Det Approach	**t**	Seq	Cmax	Delay	f	**S**
Centralized before dd change	0	1,2,3	594	194		
Centralized with new dd	1	1,2,3	594	294	888	
Decentralized with new dd	1	3,1,2	583	283	866	X
Det Approach	**t**	**Seq**	Cmax	Delay	f	**S**
Centralized before dd change	0	1,2,3	587	187		
Centralized with new dd	1	1,2,3	587	287	874	X
Decentralized with new dd	1	1,3,2	651	351	1002	

The table 2 displays the solution found by the centralized and decentralized models for the two scenarios, and before and after the due date change (denoted with time t = 0 and time t = 1 respectively). In particular the sequence, the maximum makespan (Cmax) and the delay of the job with due date changes are reported. In the last column of the table a *X* mark the best solution of the scenario. In the first scenario, with deterministic time, if we assume that α is the same of β then the Coordinator Agent is willing to change the schedule because the worsening of the *Cmax* is smaller then advantages of reducing the tardiness of job 3. In this case the *external flexibility* meant as the ability to respond quickly to market change is more important then completion of all jobs early. In the second scenario, with approximated processing time, if we assume α equal to β, the Coordinator Agent is willing to do not change the schedule, because the advantage of reducing the tardiness of job 3 do not compensate the worsening of the *Cmax*. In this case, the *internal flexibility*, meant as the ability to manage in efficient way the plant is more important that to respond quickly to market requests. In both cases the value of α and β determine the decision of CA. This theoretic example show how the introduction of an higher level of information in the system not increase the computing complexity and it can determine different decision about the plant. In particular, in the first scenario, the CA chose a decentralized approach but with approximated information about the job's processing time. With the deterministic time the decision change and the advantages of the model are major using a centralized approach. In general, the use of determinist information allow to avoid mistakes about the value of job's processing time estimate. As reported in Figures 3 and 4, the errors of this approximation are: *(i)* 8% respect to deterministic time, using a centralized approach; *(ii)* 5% respect to deterministic time adopting a decentralized approach.

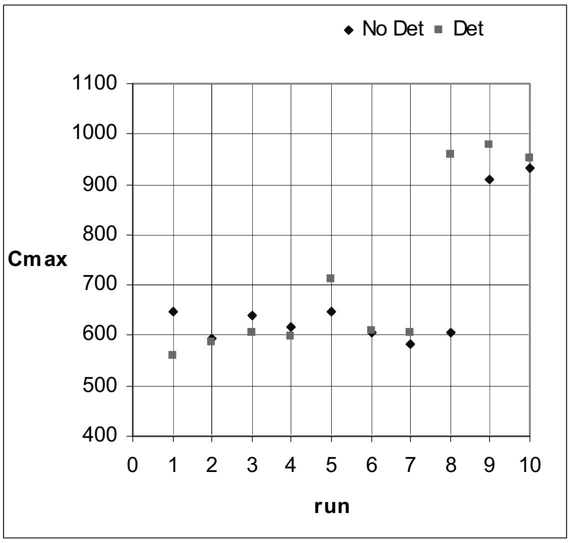


Figure 3. Comparison between deterministic and not deterministic job's processing times using a Centralized Approach.

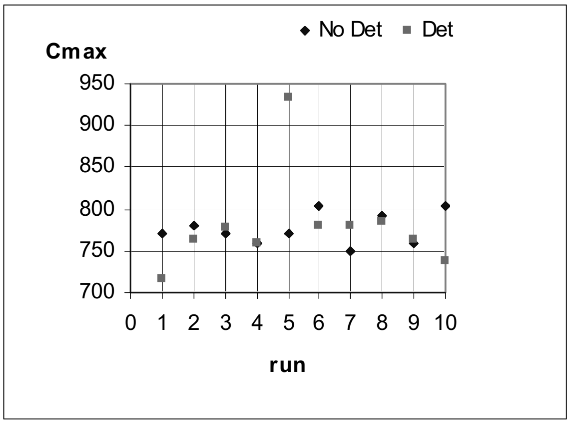


Figure 4. Comparison between deterministic and not deterministic job's processing times using a Decentralized Approach.

6 CONCLUSION

In this paper a multi-agent model able to model complex systems like FMS are presented. The advantages offered by this approach are demonstrated by an application to a shoe manufacturing plant. The results of the system simulation implemented with java language demonstrated that:

- a decentralized approach can increase the flexibility of the plant allowing the job overtakes when this solution is considered optimal by the Coordination Agent.
- An increase of plant's quantity information in the model not require an increase of complexity in the system and in particular in the island agents.
- The introduction on the model of deterministic information and of a decisional function, allow at the same time to have an instrument more flexible and closer to the real system.

7 REFERENCES

[1] El Maraghy H.A., 2006, Flexible and reconfigurable manufacturing systems paradigms, International Journal of Flexible Manufacturing Systems, 17:261–276.

[2] Browne J., Dubois D., Rathmill K., Sethi S. and Stecke K.E., 1984, Classification of flexible manufacturing systems, The FMS Magazine, 2/2:114-117.

[3] Felix T., Chan S., Chan H.K., 2004, A comprehensive survey and future trends of simulation study on FMS scheduling, Journal of Intelligent Manufacturing, 15/1: 87-102.

[4] De Toni A., Tonchia S., 1998, Manufacturing Flexibility: a literature review, International Journal of Production Research, 36/6:1587-617.

[5] Sarkis J., 1997, An empirical analysis of productivity and complexity for flexible manufacturing systems, International Journal of Production Economics, 48:39-40.

[6] ElMaraghy W.H., Urbanic R.J., 2003, Modelling of Manufacturing Systems Complexity, CIRP Annals – Manufacturing Technology, 52/1: 363-366.

[7] I.Baffo, G.Confessore, G.Stecca, Development Of A Multi-Agent Model For Production Scheduling In Innovative Flexible Manufacturing Systems. 40th CIRP International Seminar on Manufacturing Systems. Liverpool, 30 May – 1 June 2007. CDROM.

[8] Cilliers P., 1998, Complexity and postmodernism: understanding complex system, Routledge, London.

[9] Wang J., Deng Y., 1999, Incremental modelling and verification of flexible manufacturing systems, Journal of Intellingert Manufacturing, 10: 485-502.

[10] S. Fujii, T. Kaihara and H. Morita, 2000, A Distributed Virtual Factory in Agile Manufacturing Environment, International Journal of Production Research, Taylor & Francis Ltd., Vol.38, No. 17:4113-4128.

[11] Martinez Lastra J.L., Colombo A.W., 2006, Engineering framework for agent-based manufacturing control, Engineering Applications of Artificial Intelligence, Vol19, 625–640.

A Real-time Scheduling Method Considering Human Operators in Autonomous Distributed Manufacturing Systems

Koji Iwamura[1], Shinya Kuwahara[1],
Yoshitaka Tanimizu[1] and Nobuhiro Sugimura[1]

[1] Department of Mechanical Engineering, Graduate School of Engineering,
Osaka Prefecture University, Japan

Abstract

A real-time scheduling method considering human operators is proposed, in this paper, to select suitable combinations of the human operators, the manufacturing equipment and the jobs for the manufacturing processes. The proposed scheduling method consists of three steps. In the first step, the human operators select their favourite processes which they will carry out in the next time period, based on their preferences. In the second step, the machine tools and the jobs select suitable combinations for the next machining processes. In the third step, the automated guided vehicles and the jobs select suitable combinations for the next transportation processes.

Keywords:
Autonomous distributed manufacturing systems; Real-time scheduling; Human operators

1 INTRODUCTION

Recently, automation of manufacturing systems in batch productions has been much developed aimed at realizing flexible small volume batch productions. The control structures of the manufacturing systems developed, such as FMS (Flexible Manufacturing System) and FMC (Flexible Manufacturing Cell), are generally hierarchical. The hierarchical control structure is suitable for economical and efficient batch productions in steady state, but not adaptable to very small batch productions with dynamic changes in the volumes and the varieties of the products.

Computer systems and manufacturing cell controllers have recently made much progress, and individual computers and controllers are now able to share the decision making capabilities in the manufacturing systems. The network architectures are widely utilized for the information exchange in the design and the manufacturing, and some standardized models, such as STEP [1] and CNC data model [2], have been developed for the information exchange through the information networks for the design and the manufacturing.

New distributed architectures of manufacturing systems are therefore proposed to realize more flexible control structures of the manufacturing systems, in order to cope with the dynamic changes in the volume and the variety of the products and also the unforeseen disruptions, such as malfunction of manufacturing equipment and interruption by high priority jobs. They are so called as ADMS (Autonomous Distributed Manufacturing Systems) [3], BMS (Biological Manufacturing Systems) [4]-[6], and HMS (Holonic Manufacturing Systems) [7]-[11].

Many researches have been carried out to deal with the distributed architectures for planning and control of the manufacturing systems [12]-[13]. However, there are only a few researches dealing with the human operators as the autonomous components of the distributed manufacturing systems [5].

Human operators are newly considered, in this paper, in order to establish more flexible and robust real-time scheduling method for the ADMS. Improvement of human operator's motivation is also considered in the real-time scheduling method.

2 REAL-TIME SCHEDULING METHOD FOR HMS [7]-[8]

2.1 Basic structures of Holonic Manufacturing Systems

The Holonic Manufacturing System (HMS) has been proposed and discussed in an international cooperative research consortium called HMS consortium, in order to develop a new autonomous distributed architecture of the manufacturing systems, which are applicable to very small batch productions.

In the previous papers [7]-[8], a real-time scheduling method based on the utility values and the dispatching rules have been proposed and applied to the HMS, in order to determine the suitable combinations of holons for the machining processes and the transportation processes.

The HMS consists of a set of holons of automated manufacturing equipment, such as CNC machine tools and automated guided vehicles, which carry out the manufacturing processes without the human operators. The holons in the HMS are divided into four classes based on their roles in the manufacturing processes and the scheduling processes.

(a) CNC machine tool (CMT) holons: They transform the job holons in the manufacturing process. In the scheduling process, they evaluate the utility values for the candidate job holons which carry out the machining processes in the next time period.

(b) Automated guided vehicle (AGV) holons: They transport the job holons in the manufacturing process. In the scheduling process, they determine the job holons which carry out the transportation processes in the next time period by using dispatching rules.

(c) Job holons: They are transformed by the CMT holons from the blank materials to the final products, and transported by the AGV holons in the manufacturing process. In the scheduling process, they evaluate the utility values for the candidate CMT holons which carry out the machining processes in the next time period, and they select the AGV holons which carry out the transportation processes in the next time period by using dispatching rules.

(d) Coordination holon: It selects a most suitable combination of the CMT holons and the job holons for the machining processes in the next time period, based on the utility values sent from the CMT holons and the job holons.

2.2 Real-time scheduling method for machining processes

In the previous research [7], a real-time scheduling method based on the utility values has been proposed to determine a suitable combination of the job holons and the CMT holons which carries out the machining processes in the next time period.

At the time *t*, all the 'idling' holons have to determine their machining schedules in the next time period, as shown in the Figure 1. The individual job holons and CMT holons firstly select all the candidate holons for the machining processes in the next time period, and determine the utility values for the individual selected candidates, based on their own objective functions.

The utility values have the range from 0 to 1, which indicates the ratio of the improvement of the objective function values for the cases where the holon selects the candidate holons for the next machining process. It is assumed that the individual job holons and CMT holons have one of the objective functions shown in Table 1 for evaluating the utility values.

All the job holons and the CMT holons send the selected candidates and the utility values of the candidates to the coordination holon, in the second step. The coordination holon determine a suitable combination of the job holons and the CMT holons which carry out the machining processes in the next time period, based on the utility values. The decision criteria of the coordination holon is to maximize the total sum of the utility values of all the holons.

2.3 Real-time scheduling method for transportation processes

In the previous research [8], a real-time scheduling method based on the dispatching rules has been proposed to select a suitable combination of the job holons and the AGV holons which carries out the transportation processes in the next time period.

At the time *t*, all the 'idling' job holons, which have already decided their next machining schedules, have to select their transportation schedules in the next time period, as shown in the Figure 2.

All the job holons firstly select suitable AGV holons which are 'idle' by using SPT (Shortest Processing Time) rule. Some collisions may occur among the selections of the job holons. For example, more than one job holons select a same AGV holon for their next transportation processes, as shown in Figure 2(a).

If an AGV holon is selected by more than one job holons, the AGV holon selects a most suitable job holon by using

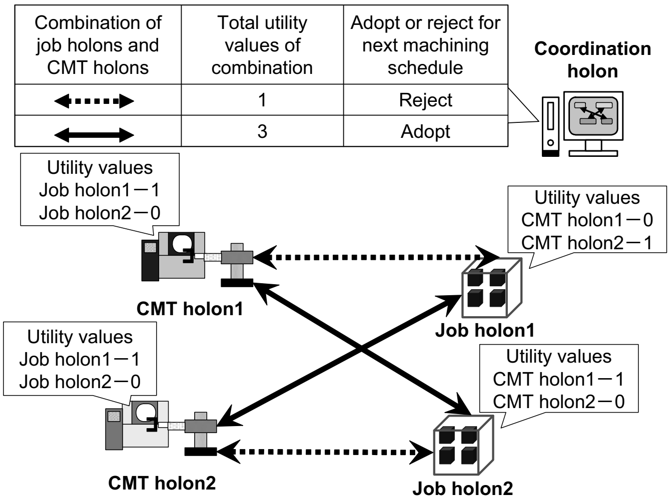


Figure 1: Real-time scheduling method for machining processes.

Table 1: Objective functions of holons.

Objective functions		Objective function values
CMT holon	Efficiency	ΣMachining time / Total time
	Machining accuracy	Σ(Machining accuracy of CMTs – Required machining accuracy of jobs)
Job holon	Flow-time	Σ(Machining time + Waiting time)
	Machining cost	Σ(Machining cost of CMTs)

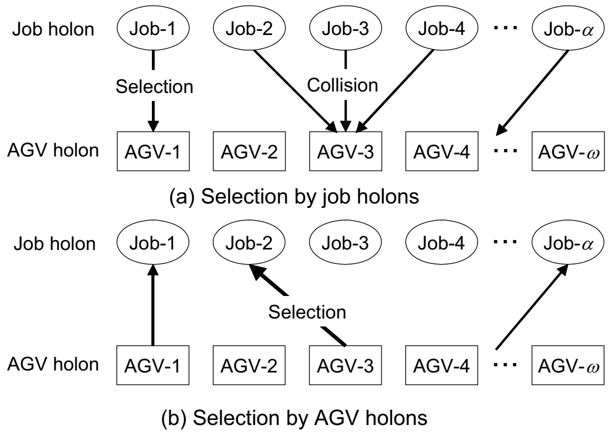


Figure 2: Real-time scheduling method for transportation processes.

SPT/TWKR (SPT/Total Work Remaining) rule, in order to avoid the collisions, in the second step, as shown in Figure 2(b). The job holons and the AGV holons select most suitable ones by applying their own dispatching rules.

3 REAL-TIME SCHEDULING METHOD CONSIDERING HUMAN OPERATORS

3.1 Manufacturing systems including human operators

This section deals with the ADMS including the human operators based on the HMS proposed in the previous papers, aimed at establishing a real-time scheduling method for the manufacturing systems with human operators.

The target manufacturing systems consist of the following autonomous components. They are, the human operators, the manual machine tools (MMTs), the handcarts, the CNC machine tools (CMTs), the automated guided veichles (AGVs), the preparation stations, and the jobs.

The scheduling system considered here generates suitable manufacturing schedules for all the components autonomously and real-timely. The following conditions are assumed for the components and the manufacturing processes.

(1) The human operators can carry out the machining processes, the transportation processes and the setup processes based on their skills. The operation time of the machining processes and the setup processes, and the machining accuracy depend on their skills.

(2) The human operators have the priority against the other automated manufacturing equipment in the real-time scheduling process, which determines the manufacturing processes to be carried out in the next time period.

(3) The machining processes are carried out by either the CMTs or the combination of the MMTs and the human operators.

(4) The setup processes are required before the machining processes to change the fixtures and the fixturing positions. The setup processes are divided into the external setups and the internal setups. The external setups and the internal setups are required before executing the machining processes by the CMTs and the ones by the MMTs respectively.

(5) The external setups are carried out by the combination of the preparation stations and the human operators. In this case, the transportation processes are needed between the CMTs and the preparation stations.

(6) The internal setups are carried out by the human operators at the MMTs, just before the machining processes by the MMTs.

(7) The transportation processes are carried out by either the AGVs or the combination of the human operators and the handcarts.

The followings summarize the manufacturing processes of the jobs to be manufactured.

J_{ik}: k-th machining process of the job i. (i = 1,..,α), (k = 1,..,β).

JFI_{ik}: Internal setup process for the machining process J_{ik}.

JFE_{ik}: External setup process for the machining process J_{ik}.

CR_{ik}: Type of the CMT, which can carry out the machining process J_{ik}.

MR_{ik}: Type of the MMT, which can carry out the internal setup process JFI_{ik} and the machining process J_{ik}.

CR_{ikm}: m-th candidate for type of CMT CR_{ik}. (m = 1,...,γ).

MR_{ikm}: m-th candidate for type of MMT MR_{ik}. (m = 1,...,δ).

FR_n: n-th candidate of the preparation stations, which can carry out the external setup process JFE_{ik}. (n = 1,...,μ)

JAC_{ik}: Required machining accuracy of the machining process J_{ik}. It is assumed that the machining accuracy is represented by the levels of the accuracy indicated by 1, 2, and 3, which mean rough, medium, and high accuracy, individually.

The human operators' capabilities and skills are given by the following parameters.

$SKL_{ik,p}$: The skill of human operator p (p = 1,…,ρ) for the type of MMT MR_{ik} and external setup process JFE_{ik}. It is assumed that the skill of human operators is represented by the grade indicated by 1, 2, and 3, which mean low, middle, and high grade, individually.

$MRT_{ik,SKL}$: Internal setup time and machining time in the case where the human operator p with the skill $SKL_{ik,p}$ carries out the internal setup process JFI_{ik} and the machining process J_{ik} by using the type of MMT MR_{ik}.

$MRAC_{ik,SKL}$: Machining accuracy in the case where the human operator p with the skill $SKL_{ik,p}$ carries out the machining process J_{ik} by using the type of MMT MR_{ik}. $MRAC_{ik,SKL}$ is also represented by the levels of 1, 2 and 3.

$MRCO_{ik,SKL}$: Internal setup cost and machining cost in the case where the human operator p with the skill $SKL_{ik,p}$ carries out the internal setup process JFI_{ik} and the machining process J_{ik} by using the type of MMT MR_{ik}.

$FT_{ik,SKL}$: External setup time in the case where the human operator p with the skill $SKL_{ik,p}$ carries out the external setup process JFE_{ik}.

$FCO_{ik,SKL}$: External setup cost in the case where the human operator p with the skill $SKL_{ik,p}$ carries out the external setup process JFE_{ik}.

$TrT_{p,A,B}$: Transportation time in the case where the human operator p carries out the transportation process from position A to positon B.

$TrCO_{p,A,B}$: Transportation cost in the case where the human operator p carries out the transportation process from position A to position B.

The followings give the capabilities of the automated manufacturing equipment including CMTs and AGVs.

CRT_{ikm}: Machining time in the case where the CMT CR_{ikm} carries out the machining process J_{ik}.

$CRAC_{ikm}$: Machining accuracy in the case where the CMT CR_{ikm} carries out the machining process J_{ik}. $CRAC_{ikm}$ is also represented by the levels of 1, 2 and 3.

$CRCO_{ikm}$: Machining cost in the case where the CMT CR_{ikm} carries out the machining process J_{ik}.

$TrT_{q,A,B}$: Transportation time in the case where the AGV q (q = 1,…,ω) carries out the transportation process from position A to positon B.

$TrCO_{q,A,B}$: Transportation cost in the case where the AGV q carries out the transportation process from position A to position B.

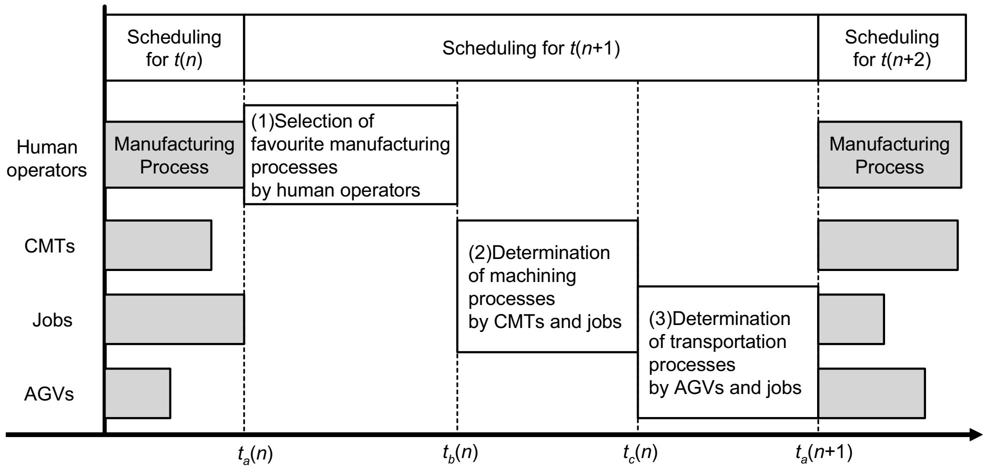


Figure 3: Real-time scheduling method considering human operators.

3.2 Real-time scheduling method

A real-time scheduling method is proposed, in this research, to select suitable combinations of the human operators, the manufacturing equipment and the jobs for the manufacturing processes. The proposed scheduling method consists of following three steps shown in Figure 3.

Step1 Selection of favourite processes by human operators

The human operators select their favourite manufacturing processes which they want to carry out in the next time period, based on their preferences. However, it is a very time consuming process to select the most favourite manufacturing process from all the candidate machining processes, the setup processes and the transportation processes by the human operators.

Therefore, human interfaces and a coordinator are proposed for the human operators to select the favourite manufacturing processes, as shown in Figure 4. The human operators are required to input the data describing their favourite manufacturing processes to the human interfaces, before executing the real-time scheduling processes. The information of the input data is described in section 3.3. The human operators select their favourite processes based on the following procedures.

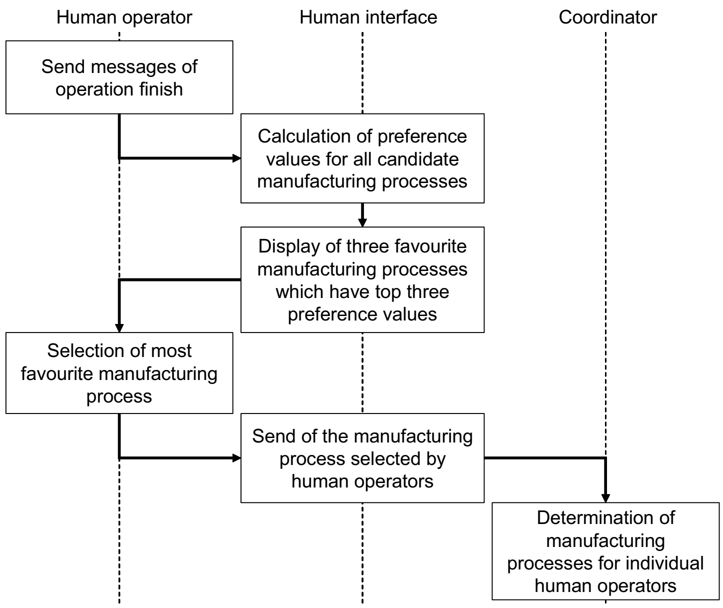


Figure 4: Selection of favourite manufacturing processes.

(1) When the human operators finish their operations, they send a message to the human interface, in order to get the information about the next operations.

(2) The human interfaces generate all the candidate manufacturing processes for the operators and calculate the preference values for all the candidate manufacturing processes. The calculation processes of the preference values are described in section 3.3.

(3) The human interfaces display three most suitable manufacturing processes to the human operators, which have higher preference values.

(4) The human operators select a most favourite manufacturing process from all the candidate manufacturing processes.

(5) The human interface sends the manufacturing process selected by the human operators to the coordinator.

(6) The coordinator determines the manufacturing processes of the individual human operators, based on their selections. The manufacturing processes of the human operators are determined according to the sequence of the operator's messages received by the coordinator, in order to avoid the conflict among the manufacturing processes selected by the human operators.

Step2 Determination of machining processes by CMTs and jobs

The CMTs and the jobs select suitable combinations for the next machining processes. This process is carried out by using the utility values described in section 2.2. It is assumed that the coordinator has the same function with the coordinator holon described in section 2.1.

Table 2: Preference value of human operator *p*.

	Machining processes					Transportation process	Setup process
	MR_{i1}	…	MR_{ik}	…	$MR_{i\beta}$		
n	1	…	k	…	β	β+1	β+2
Job 1	x_{11p}	…	x_{1kp}	…	$x_{1\beta p}$	$x_{1\beta+1p}$	$x_{1\beta+2p}$
Job 2	x_{21p}	…	x_{2kp}	…	$x_{2\beta p}$	$x_{2\beta+1p}$	$x_{2\beta+2p}$
…	…	…	…	…	…	…	…
Job *i*	x_{i1p}	…	x_{ikp}	…	$x_{i\beta p}$	$x_{i\beta+1p}$	$x_{i\beta+2p}$
…	…	…	…	…	…	…	…
Job α	$x_{\alpha 1p}$	…	$x_{\alpha kp}$	…	$x_{\alpha\beta p}$	$x_{\alpha\beta+1p}$	$x_{\alpha\beta+2p}$

Step3 Determination of transportation processes by AGVs and jobs

The AGVs and the jobs select suitable combinations for the next transportation processes. This process is carried out by using dispatching rule-based method described in section 2.3.

These steps are repeated until all the manufacturing processes are finished in the manufacturing systems.

3.3 Preference values

The human operators are required to input the data dealing with their favourite manufacturing processes to the human interfaces, before executing the real-time scheduling processes. Input data to the human interface are following.

(1) Types of favourite MMTs.

(2) Preferences of transportation processes and setup processes.

(3) ID of favourite jobs.

(4) Skill for individual types of MMTs and external setup processes $SKL_{ik,p}$.

The human interface generates all the candidate manufacturing processes for the human operator *p*, based on the information of (4) and the status of the manufacturing systems. Following this, the human interface calculates the preference values of all the candidate manufacturing processes x_{inp} based on the information of (1), (2) and (3).

x_{inp} means the preference values for the cases where the human operator *p* carries out the manufacturing process *n* of the job *i*, as shown in Table 2. The manufacturing process *n* is the machining processes by using type of the MMT MR_{ik} ($n = k = 1,\ldots,\beta$), the transportation process ($n = \beta+1$), and the external setup process ($n = \beta+2$). The human interface calculates x_{inp} by using following equation.

$$x_{inp} = A_{np} + B_{np} + C_{np} + D_i \qquad (1)$$

where,

$$A_{np} = \begin{cases} 1: \text{ if the human operator } p \text{ prefers the type of the MMT } MR_{ik}. \\ 0: \text{ otherwise.} \end{cases}$$

$$B_{np} = \begin{cases} 1: \text{ if the human operator } p \text{ prefers the transportation processes.} \\ 0: \text{ otherwise.} \end{cases}$$

$$C_{np} = \begin{cases} 1: \text{ if the human operator } p \text{ prefers the external setup processes.} \\ 0: \text{ otherwise.} \end{cases}$$

$$D_i = \begin{cases} 1: \text{ if the human operator } p \text{ prefers the job } i. \\ 0: \text{ otherwise.} \end{cases}$$

After calculating x_{inp}, the human interface retrieves the manufacturing processes which have top three preference values for the candidate manufacturing processes which may be selected by the human operator *p* in the next time period. The human operators select a most favourite manufacturing process displayed by the human interfaces, as shown in Figure 3 and 4.

4 CASE STUDY

Some case studies have been carried out to verify the effectiveness of the proposed methods. The ADMS model considered in the case studies has 6 CMTs, 3 AGVs, 4 human operators, 8 MMTs, and 5 preparation stations.

The CMTs and MMTs are classified into 3 types which have same machining functions CR_{ik} and MR_{ik}. It means that the each type of CMTs and MMTs can carry out the same machining processes J_{ik}. The individual CMTs have the different objective functions and the different machining capacities, such as the machining time CRT_{ikm}, the machining accuracy $CRAC_{ikm}$, and the machining cost $CRCO_{ikm}$.

The individual human operators have the highest skill $SKL_{ik,p}$ for all the types of the MMT. Therefore, the individual human operators have the same manufacturing capacities, such as the internal setup time and the machining time $MRT_{ik,SKL}$, the machining accuracy $MRAC_{ik,SKL}$, the internal setup cost and the machining cost $MRCO_{ik,SKL}$, the external setup time $FT_{ik,SKL}$, and the external setup cost $FCO_{ik,SKL}$.

As regards the jobs, 30 jobs are considered in the case study, which have the different objective functions and the machining sequences.

4 cases are considered by changing the preference of the individual human operators as shown in the following.

Case 1: All human operators prefer all the types of the MMT and don't prefer the transportation processes, the external setup processes and the jobs.

Case 2: All human operators prefer the transportation processes and don't prefer all the types of the MMT, the external setup processes and the jobs.

Case 3: All human operators prefer the external setup processes and don't prefer all the types of the MMT, the transportation processes and the jobs.

Case 4: All human operators prefer all the types of the MMT, the transportation processes, the external setup processes and the jobs.

Additional 2 case studies have been also carried out to consider the cases with and without malfunctions of the CMTs.

Figure 5 shows the real-time scheduling results from the view point of the total make span. In the figure, the horizontal axis gives the cases with and without malfunctions, and the vertical axis shows the total make span. As shown in the case without malfunctions, case 3 is the most effective to reduce the total make span, because the CMTs carry out the machining processes faster than the human operators. As shown in the case with malfunctions, case 1 is more effective than case 3 to reduce the total make span, because the human operators carry out the machining processes by using the MMTs instead of the CMTs.

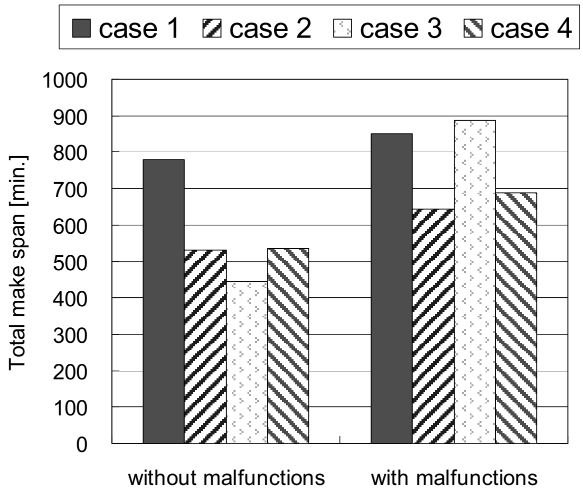


Figure 5: Results of case study.

5 CONCLUSIONS

A real-time scheduling method is proposed to consider the human operators, aimed at establishing more flexible and robust control structures of the autonomous distributed manufacturing systems. The following remarks are concluded.

(1) A real-time scheduling method is proposed, in this research, to select suitable combinations of the human operators, the manufacturing equipment and the jobs for the manufacturing processes.

(2) A systematic method is proposed to calculate the preference values of the human operators for selecting the favourite manufacturing processes.

(3) Some case studies of the real-time scheduling have been carried out to verify the effectiveness of the proposed methods. It was shown, through case studies, that the proposed methods are effective to improve the total make span in the case with the malfunctions of the CNC machine tools.

6 REFERENCES

1. ISO, 1994, ISO 10303 Industrial automation systems and integration -Product data representation and exchange-, ISO.
2. ISO, 2001, ISO 14649 Industrial automation systems and integration -Physical device control - Data model for computerized numerical control-, ISO.
3. Moriwaki, T., Sugimura, N., 1992, Object-oriented modeling of autonomous distributed manufacturing system and its application to real-time scheduling, Proc. of the ICOOMS '92, 207-212.
4. Ueda, K., 1992, An approach to bionic manufacturing systems based on DNA-type information, Proc. of the ICOOMS '92, 303-308.
5. Ueda, K., Vaario, J. Fujii, N., 1998, Interactive manufacturing: Human aspects for biological manufacturing systems, CIRP annals - Manufacturing technology, 47/1:389-392.
6. Ueda, K., Hatono, I., Fujii, N., Vaario, J., 2001, Line-less Production System Using Self-Organization: A Case Study for BMS, CIRP annals - Manufacturing technology, 50/1:319-322.
7. Iwamura, K., Tanimizu, Y., Sugimura, N., 2003, Real-time scheduling for holonic manufacturing system - Coordination among holons using effectiveness values -, Proc. of the 17th international conference of production research, CD-ROM
8. Sugimura, N., Tanimizu, Y., Iwamura, K., 2004, A Study on real-time scheduling for holonic manufacturing system, CIRP journal of manufacturing systems, 33/5:467-475.
9. Iwamura, K., Okubo, N., Tanimizu, Y., Sugimura, N., 2006, Real-time scheduling for holonic manufacturing systems based on estimation of future status, International journal of production research, 44/18-19: 3657–3675.
10. Iwamura, K., Seki, Y., Tanimizu, Y. and Sugimura, N., 2005, A study on real-time scheduling of holonic manufacturing system -Multi-objective optimization in scheduling process-, Proc. of the 38th CIRP international seminar on manufacturing systems, CD-ROM.
11. Iwamura, K., Nakano, A., Tanimizu, Y., Sugimura, N., 2007, A study on real-Time scheduling for holonic manufacturing systems -Simulation for estimation of future status by individual holons-, LNAI 4659, Holonic and multi-agent systems for manufacturing, Springer, 205–214.
12. Hino, R., Moriwaki, T., 2002, Decentralized job shop scheduling by recursive propagation method, JSME international journal series C, 45/2:551-557.
13. Kaihara. T, Fujii, S., 2006, A study on dynamic resource allocation mechanism with multi-agent systems for pareto optimal solutions, Proceedings of 2006 international symposium on flexible automation, CD-ROM.

Production Scheduling System with Dynamic Lot Size in Case of Considering Set-up Time

Minoru Koike[1], Akira Tsumaya[2], Ryota Matsui[3], Eiji Morinaga[3], Hirofumi Wakamatsu[3] and Eiji Arai[3]

[1] Department of Systems Design, College of Industrial Technology, Japan

[2] Department of Mechanical Engineering, School of Engineering, Kobe University, Japan

[3] Department of Materials and Manufacturing Science, Graduate School of Engineering, Osaka University, Japan

Abstract

The dynamic scheduling management system architecture that is based on the parts and information packets unified technology is developed, and variable lot size production with the parts and packets unified technology is proposed. We assumed that each part or unit has a data career, which handle such as identification number, position of the part, state of the part. Using these data, progress of the project process is compared with the master schedule. If the difference arises between them, scheduling system dynamically re-schedules with dynamic lot size to optimize the processes. A pilot system that realized our proposal is developed.

Keywords:

Production Scheduling, Lot Production, Dynamic Lot Size, Set-up Time

1 INTRODUCTION

An enlarged and complicated production system is needed to satisfy various requirements to production in recent years. Because understanding and managing a whole system in such circumstance become difficult, the concept of an autonomous & distributed production system has been proposed. Numerous intelligent and autonomous production systems, which are focused on product facilities and constructed intelligent system using network structure, have been proposed so far. By using the network system, each production facility can understand production status and can respond to the status. These methods are suitable to adopt an automation factory [1] [2] [3] [4] [5] [6].

The lot production system is generally adopted in automation factories. In the system, the production lot size is an important factor for productivity. It is possible to reduce work-in-process inventory in a small lot size production while setup time and production cost per products is decreased in a large lot size production. Therefore, the control of the production lot size is important to perform efficient production.

Although the conventional autonomous & distributed production system is managing production process by using information obtained from facilities, the important information are those from whole/part of the product. Therefore, acquiring information directly from parts of the products is needed for more flexible production management.

Recently, RFID technology like IC tags is widely utilized, and can be available from a practical viewpoint of cost, size, and memory size. Therefore, the information about various production activities may be given / read / written by using IC tags which stuck on parts, and the production system which manages the dynamic scheduling for variable lot size based on IC tags can be proposed.

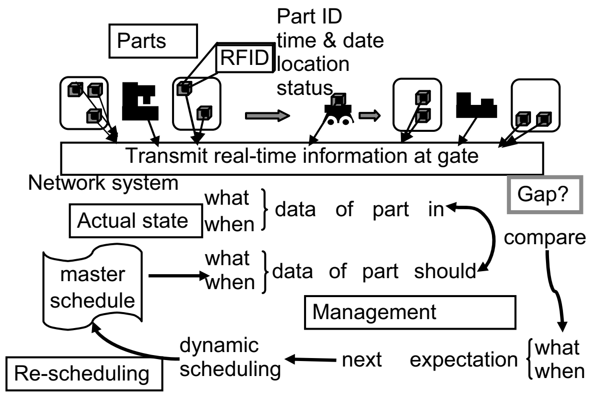


Figure 1: The concept of a dynamic production scheduling system using parts and packets unification technology.

2 CONCEPT OF PARTS & PACKETS UNIFICATION SYSTEM

The concept of a dynamic production scheduling system using parts and packets unification technology is shown in Figure 1. The RFID tag is attached in the parts, and ID number and the present status are written in it. Status is rewritten according to becoming a unit from materials. The RFID tag attached on a unit transmits ID number, current position, and status through network system whenever it passes through the gate provided in the inside of the factory, the construction site, in the transfer route between them, and the entrance of a warehouse, etc.

On the other hand, as for the production planning side, progress management of the production project is performed by the scheduling system after a master schedule is drawn up. The scheduling system can show the position and status in which each part / unit should exist. By comparing with real-time information on each part / unit transmitted through network system to the information on the master schedule, a gap between the master schedule and an actual state is detectable. According to the quantity of a gap, re-scheduling is performed when the system judges it is needed.

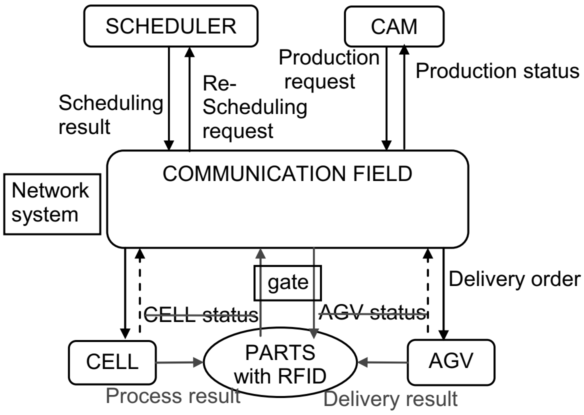


Figure 2: The concept of manufacturing system architecture using parts and packets unification technology in a factory.

3 SYSTEM ARCHITECTURE

The architecture of dynamic management for manufacturing system using parts and packets unification technology is explained on this section. Figure 2 shows the system architecture in the case of parts processing and assembly processing in a factory. The network system built by the conventional intelligent manufacturing system has combined production facilities, such as production cells and stations. Meanwhile, in the proposed system, the information that each part have in RFID tag can be exchanged with network system through the gate provided in each production facilities. For example, at the gate provided in the outlet of a processing facility, the status of parts is rewritten according to the contents of processing.

If the memory capacity of RFID tag will become large in the future, it is also possible to compose a network system only with RFID tags attached to the parts by storing processing processes and work information in RFID tag, and exchanging information about them with production facilities. However, since the present RFID tag has not satisfactory memory capacity, the proposed system hybridized the network system which combines intelligent production facilities, and parts and packets unification technology. Process and work information are transmitted to production facilities from the scheduling and CAM systems through the network. The part ID in the RFID tag is checked and after processing process, the status of the part is rewritten and also is sent to CAM system.

4 VARIABLE LOT SIZE APPROACH

In a production process on lot processing, the lot size is decided on setup time and work-in-process inventory. When a lot is bigger, setup time and cost per products are smaller while work-in-process inventory is increased. Consequently, it is needed to control the lot size according to setup time and work-in-process inventory [7] [8]. Here, we propose variable lot size production approach that permits the changing of lot sizes in each process.

Figure 3 shows the concept of variable lot size production approach. We assume that each part has several production processes to become final products. At the start of manufacturing, the same kind of parts composes are gathered as a component of a sub-lot. When one of the processes is finished, the lot is divided into original sub-lot. Each sub-lot sends data of its state, following process, and timestamp to the manager. When the manager received the data, it selects the combination of sub-lots according to following process of each sub-lot and facility's setup time [9] [10].

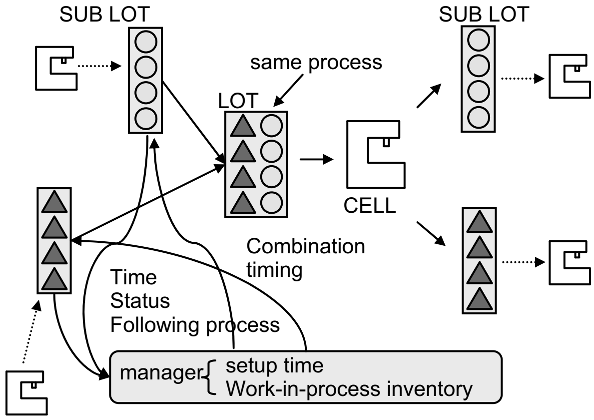
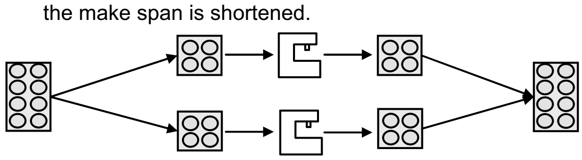


Figure 3: The concept of variable lot size production approach.

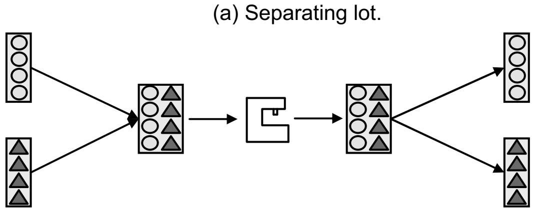
5 CASE STUDY

A pilot system that realized our proposal is developed and applied to two case studies. (Figure 4)

1. The System separates the lot when the division of the lot can reduce work-in-process in the bottleneck process. As a result, the make span is shortened. (Figure 5, 6)
2. There is a process in which the processing time isn't influenced by the lot size. The System integrates processes when the integration of the processes can shorten the processing time in the process. As a result, the make span is shortened.



(a) Separating lot.



(b) Integration processes.

Figure 4: Separating lot and integration processes.

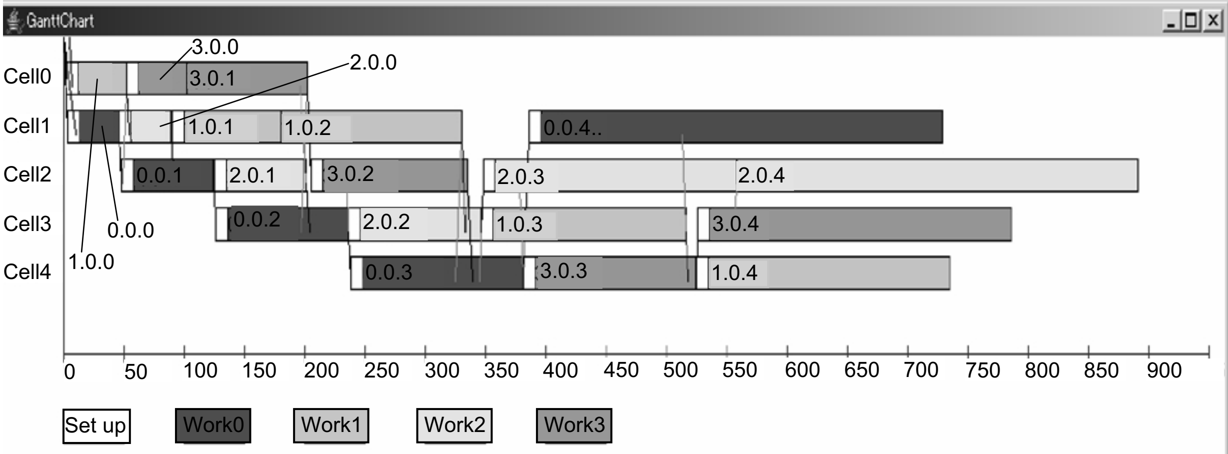


Figure 5: Gant chart in case without separating lot.

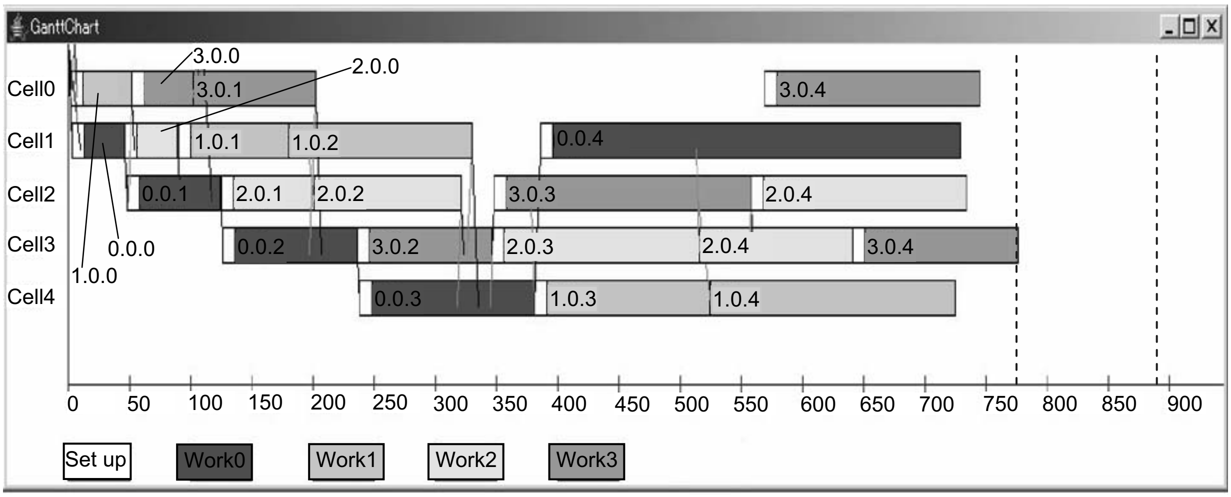


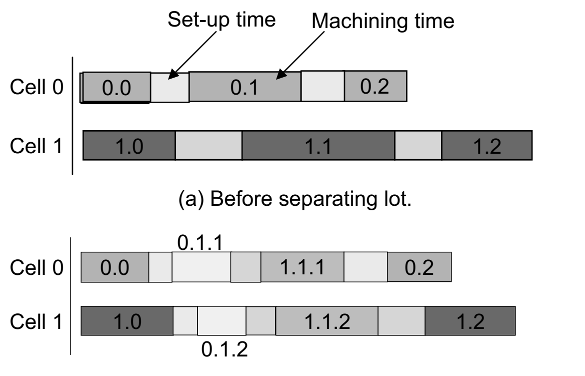
Figure 6: Gant chart in case using separating Lot.

6 PRODUCTION SCHEDULING SYSTEM WITH DYNAMIC LOT SIZE CONSIDERING SET-UP TIME

The differences in the set-up times are not considered when the scheduling is executed shown at the case studies in the previous section. That is, the set-up times are managed at the fixed time without regard to the kind of the next machining process and the relationship between machining processes before and after the target process .

In this section, the production scheduling with dyanamic lot size condiering the differences in the set-up times is discussed. The case in the separating lot is described in the pages that follow.

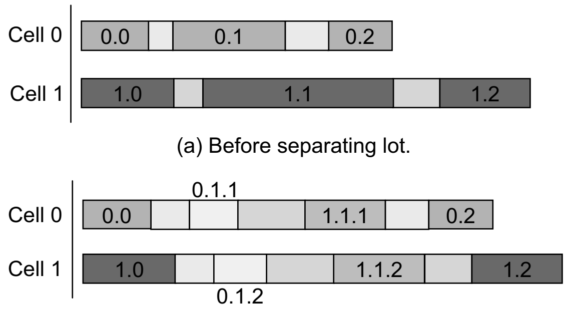
Separating the machining process '1.1' is assumed as shown in Figure 7(a). When the set-up time is fully short compared to the machining time, the effect of the separating lot appears, and the total machining time can be shortened as shown in Figure 7(b). Next, when the set-up time is long compared to the machining time, separating the machining process '1.1' is assumed as shown in Figure 8(a). There is some possibility that the end-time of the last machining process is late by the separating lot as shown in Figure 8(b).



(a) Before separating lot.

(b) After separating lot.

Figure 7: Separating lot in case with short set-up time.



(a) Before separating lot.

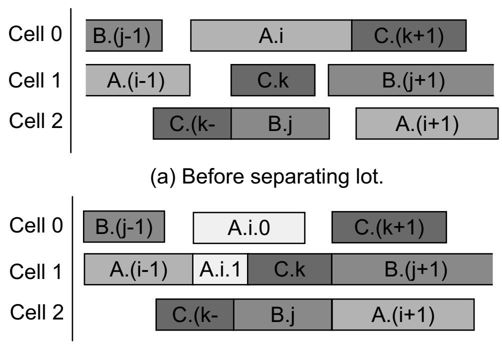
(b) After separating lot.

Figure 8: Separating lot in case with the long set-up time.

In the former discussion, we assumed that the set-up time of Cell 0 and Cell 1 are same without the relation to the former machining process. If the master schedule is constructed concerning the machining processes before and after the target process for the shortening of the set-up time, the set-up times both before and after the inserted process can be longer in the cell inserted the separated machining process (Cell 1 in Figure 8).

In this way, the separating lot is not always contributed to the shortening of the total processing time when the differences in the set-up times are considered.

Then, the lot is separated unequally. Two new jobs are generated by the separating lot. The cell which has had the original lot is called 'Cell A'. The cell where the new job is inserted is called 'Cell B'. The slack times of the jobs scheduled before and after the new job inserted on the 'Cell B' are totalled. The lot size which can execute both set-up and machining in the total time is assigned on the 'Cell B' as shown in Figure 9. When the total time is enough, the lot is separated so that two estimated end times of machining of two jobs both 'Cell A' and 'Cell B' approach.



(a) Before separating lot.

(b) After separating lot.

Figure 9: The unequall separating lot considering set-up time.

7 SUMMARY /CONCLUSION

The following conclusion can be drawn.

- The concept and the architecture of parts and packets unified system in the production system were proposed.
- It is thought that the parts and packets unified system is available in production system.
- A production scheduling system with dynamic lot size was proposed.
- The system changes lot size dynamically, using parts and packets unified system.
- The effectiveness using the production scheduling system with dynamic lot size was discussed.
- The production scheduling system with dynamic lot size considering set-up time was discussed.

8 ACKNOWLEDGMENTS

The authors appreciate Mr. Hirokazu Watanabe and Mr. Takafumi Nakayama, who are graduates of Osaka University, for their help in the preparation of this paper.

9 REFERENCES

[1] Ueda, K., 1992, An Approach to Bionic Manufacturing Systems Based on DNA-Type Information, Proc. International Conference on Object-Oriented Manufacturing Systems, 305-308.

[2] Ranky, P.G., 1992, Intelligent Planning and Dynamic Scheduling of Flexible Manufacturing Cell and Systems, Proc. 1992 Japan-U.S.A. Symposium on Flexible Automation, 415-422.

[3] Wiendahl, H.P., Garlichs, R., 1994, Decentral Production Scheduling of Assembly Systems with Genetic Algorithm, Annals of the CIRP, 43/1:389-396.

[4] Fujii, S., Kaihara, T., Morita, H., Tanaka, M., 1999, A Distributed Virtual Factory in Agile Manufacturing Environment, Proc. 15th Conference of the International Foundation for Production Research, II: 1551-1554.

[5] Sugimura, N., Tanimizu, Y., Yoshioka, T., 1999, A Study on Object Oriented Modeling of Holonic Manufacturing System, Manufacturing System, 27/3:253-258.

[6] Shirase, K., Honjou, S., Wakamatsu, H., Tsumaya, A., Arai, E., 2002, Co-operative Scheduling Based on HLA, Proc. 5th International Conference on Design Information Infrastructure Systems for Manufacturing 2002, 25-33.

[7] Tsubone, H., Ohba, M., Uetake, T., 1996, The impact of lot sizing and sequencing on manufacturing performance m a two-stage hybrid flow shop, Int. J. Proc. Res., 34/11:3037-3053.

[8] C. N. Potts and K. R. Baker, 1989, Flow shop scheduling with lot streaming, Opns. Res. Letter. 8:279-303.

[9] Arai, E., Tsumaya, A., Watanabe, H., Wakamatsu, H., Shirase, K., Takata, M., 2004, Dynamic Management Architecture for Production Systems Based on Parts and Packets Unification, Experiences from the Future - New Methods and Applications in Simulation for Production and Logistics, 491-500,

[10] Watanabe, H., Tsumaya, A., Wakamatsu, H., Shirase, K., Arai, E., 2004, Dynamic Scheduling System with Variable Lot Size Approach Using Parts and Packets Unification, Proceedings of 2004 JAPAN-U.S.A. Symposium on Flexible Automation (JUSFA2004), CD-ROM, JS_038(4 pages),

The Impact of Lot Splitting in a Single Machine Scheduling Problem with Earliness - Tardiness Penalties

Chinyao Low[1] Yuling Yeh[2]

[1]Department of Industrial Management, National Yunlin University of Science and Technology, Taiwan

[2]Department of Industrial Engineering and Management, Nan-Kai Institute of Technology, Taiwan

Abstract

This paper consider a scheduling problem in which each job in the system can be split into several sublots and need to be processed on a single machine. During processing, a setup time is incurred whenever there is a switch from processing a job in one sublot to a job in another sublot. All sublots belong to the same job have a common due date. The objective in this research is to minimize the total earliness and tardiness penalties. The approach was performed using both a mathematical programming model for small size problems and a heuristic algorithm for medium to large problems. The study shows the benefits from job splitting improve monotonically with decreasing sublot size and that job splitting can provide improvements as long as the setup to processing time ratios less than 25%.

Keywords:

Scheduling; Sublot; Earliness;Tardiness; Job splitting

1 INTRODUCTION

Modern technologies of manufacturing provide an opportunity to process jobs in batches and open a new application area of research on scheduling. The scheduling models in this area lead to a new topic of problems which split a job into some sublots and sequence the jobs in each sublot to meet the required performance criteria. The motivation for early and tardy penalties comes from just-in-time (JIT) philosophy. In general, the earliness cost can be considered as a holding cost for finished goods, deterioration of perishable goods, and opportunity costs. The tardiness cost may be the backlogging cost, which includes performance penalties, lost sales, and lost goodwill. Lot streaming problems have received tremendous attention for the last fifteen years [1-13]. Most of researchers have assumed that there is no setup time between two adjacent jobs, i.e. another job can be processed immediately after the machine finishes processing one job. However, this assumption is not always valid in real world applications. Hence, the setup time has taken into account for the problem. The addressed operation scheduling problem is formally stated as follows. Consider a N jobs production system, each job i has a demand TQ_i to be processed on a single machine. Each job i has a unit processing time P_i and can be splitted into B consistent batchs. A setup time S_i is required for each job i when the preceeding job is not the same. For a job i the costs per unit time of being early E_i or late L_i are known. The earliness and tardiness costs for a job are assume to be linear; that is, if job i is finished t time units after its corresponding due date, a penalty $L_i \times t$ is incurred, and if it is completed t time units before its corresponding due date, a penalty $E_i \times t$ is also incurred. The objective is to minimize the total penalty, which is defined as sum of the penalty cost of each job.

2 MATHEMATICAL MODEL

2.1 Notation

To facilitate the problem formulation, we first introduce some notations that will be used throughout the paper as follows.

Subscripts:

i: job type index, $i = 1,2,3,\cdots,N$

j: sublot index, $j = 1,2,3,\cdots,B$

Parameters and sets:

O_{ij} : sublot j of job i

TQ_i : demand of job i

S_i : setup time of job i

P_i : processing time per unit of job i

D_i : due date of job i

H: a large positive number

E_i : early penalty per unit time of job i

L_i : late penalty per unit time of job i

Decision variables

BQ_{ij} : number of units in sublot O_{ij}

F_i : flow time of job i

d_i^+ : tardiness of job i, $d_i^+ = \max\{0, F_i - D_i\}$

d_i^- : earliness of job i, $d_i^- = \max\{0, D_i - F_i\}$

ST_{ij} : starting epoch of setup task for O_{ij}

X_{ij} : 1, if O_{ij} need setup before processing; 0, otherwise

$y_{iji'j'}$: 1, if $O_{i'j'}$ processed immediately after O_{ij} ; 0, otherwise

2.2 Model

The objective of the addressed scheduling problem is to minimize the total penalty in the system, which is the summation of early and tardy penalties of each job. Thus, the model may be stated as follows.

Minimize:

$$\sum_{i=1}^{N}(E_i \cdot d_i^- + L_i \cdot d_i^+)$$

subject to

$$\sum_{j=1}^{B} BQ_{ij} - TQ_i = 0 \qquad \forall i \tag{1}$$

$$F_i = ST_{iB} + S_i \cdot X_{iB} + P_i \cdot BQ_{iB} \qquad \forall i, j \tag{2}$$

$$F_i + d_i^- - d_i^+ = D_i \qquad \forall i \tag{3}$$

$$ST_{ij'} \geq ST_{ij} + S_i \cdot X_{ij} + P_i \cdot BQ_{ij} \tag{4}$$

$$\forall i, j = 1,2,3..., B-1$$

$$ST_{ij} \geq ST_{i'j'} + S_{i'} + P_{i'} \cdot BQ_{i'j'} - H \cdot Y_{iji'j'} \tag{5}$$

$$ST_{i'j'} \geq ST_{ij} + S_i + P_i \cdot BQ_{ij} - H \cdot (1 - Y_{iji'j'})$$

$$X_{i1} = 1 \qquad \forall i \tag{6}$$

$$B \cdot (N-1) \cdot X_{ij} \geq \sum_{i'=1}^{N}\sum_{j'=1}^{B}(Y_{i(j-1)i'j'} - Y_{iji'j'}) \tag{7}$$

$$\forall i, j = 2,3,4,..., B \quad i \neq i'$$

Constraint (1) describes that the batch size of all sublots for each job satisfies its corresponding demand. Constraint (2) defines the flow time of each job. Constraint (3) defines the distance between the flow time and corresponding due date of a job, where d_i^- and d_i^+ stand for the earliness and tardiness of a job *i*, respectively. Constraints (4) and (5) ensure that any two sublots can not be processed simultaneously. Constraint (6) states that the first sublot of each job *i* must be setup. Constraint (7) checks whether the sublot needs to setup before processing. The model can be easily solved using some commercial mathematical programming software such as Cplex. However, the above formulation can be used to solve the problems containing only a few jobs; therefore, other solution approaches, including heuristics that provide quick and near-optimal solutions, are very desirable.

3 DEVELOPMENT OF THE HEURISTIC

In this section, the algorithm begins at a schedule generator to guarantee that a favorable initial solution can be obtained for the addressed problem. Next, applying the tabu search procedure to proceed the solution improving process.

3.1 Starting Solution

For solving the addressed problem, we first assume the jobs are on no lot splitting condition, then solve it near optimally. The procedure is developed in two phases: a dynamic rule called the MET (Minimum Earliness-Tardiness Penalties First) rule is applied first, to obtain an initial sequence, and then tabu search procedure is applied to make further improvements.

Phase I- Minimum Earliness-Tardiness Penalties First Rule (MET)

In the phase I, the initial job sequence is developed. The sequential job assignments start from the first position and proceed forward toward the last position. The assignments are terminated till the last position is assigned a job. The procedure is described in detail as follows.

Step 1: Calculate the penalty for each unscheduled job *i* as

$$PTY_i = \begin{cases}(C_i - D_i) \times E_i & \text{if } C_i > D_i \\ (D_i - C_i) \times L_i & \text{if } D_i > C_i\end{cases}$$

where C_i = machine available time + total processing time of job *i*

Step 2: Select the job *i* with minimum penalty PTY_i, and place it in the earliest possible position in the sequence.

Step 3: Remove the selected job, and update the machine available time for scheduling next

The above steps continuous until all jobs are scheduled.

Phase II- Tabu Search Improvement Procedure

The general framework of our tabu search algorithm is listed as follows.

Step 1. *Initialization*:

Step 1.1. Obtain a feasible starting solution *S* of MET (Section 3.1).

Step 1.2. Initialize parameters $n_{s\max}$, $n_{i\max}$, $N(\Omega_S)$, $N(\Omega_L)$

Step 1.3. Set counter n_{uic} = 0, n_{ic} = 0 and tabu lists $\Omega_S = \phi$, $\Omega_L = \phi$

Step 2. While (*TotalS* < $n_{s\max}$) or (n_{bsc} < 5), do the following.

Step 2.1. While (*Total* < $n_{i\max}$) or (n_{uic} < 20), perform the following.

Step 2.1.1. (Neighbor Search) Pick a neighbor S^{c1} of *S*. Where $S^{c1} \not\subset \Omega_S$

Step 2.1.2. Compute $\Delta = f(S^{c1}) - f(S)$

Step 2.1.3. If $\Delta \leq 0$, set $S^T = S^{c1}$, and $n_{ic} = n_{ic}$ +1, and n_{uic} = 0

Step 2.1.3.1. Compute $\Delta_b = f(S^T) - f(S^*)$,

Step 2.1.3.2. If $\Delta_b < 0$, set $S^* = S^T$, $n_{bsc} = n_{bsc}$ +1.

Step 2.1.4. If $\Delta > 0$, n_{uic} = n_{uic} + 1, n_{ic} = 0

Step 2.1.5. Update tabu list Ω_S

Step 2.1.5.1. If n_{ic} > 10, increase the tabu list size.

Step 2.1.5.2. If n_{uic} > 10, reduce the tabu list size.

Step 2.1.6. *Total* = *Total* +1

Step 2.2. Determine the "most distance" solution S^{c2} of S, where $S^{c2} \not\subset \Omega_L$

$$S^{c2} = \sqrt{\left(f\left(S^{c2}\right) - f\left(S^{*}\right)\right)^2}$$

Step 2.3. Let the "most distance" solution S^{c2} be the starting solution of stage *TotalS* +1

Step 2.4. Reset counter n_{uic} = 0, n_{ic} = 0 and tabu lists

$$\Omega_s = \phi$$

Step 2.5. Set counter *TotalS* = *TotalS* +1,

Step 3. Return the best solution found S^* .

3.2 Heuristic for the proposed problem

Step 1. *Initialization*

For a schedule obtained from Section 3.2, split each job into *B* consistent sublots.

Step 2. *Solution improving*

Applying Tabu search procedure addressed in the Section 3.1 for the further improving.

4 EXPERIMENTAL ANALYSIS AND DISCUSSION

4.1 Design of testing problems

The design of the test problems is similar to those of previous studies for the weighted tardiness problem (Srinivasan 1971, Ow 1985, Ow and Morton 1989). A tardiness factor, *T*, introduced by Srinivasan to be a coarse measure of the proportion of jobs which might be expected to be tardy in a given sequence. The value of *T* can vary from 0 to 1. A data set with *T* close to 0 has a loose schedule. As we examine data sets with increasing value of *T*, the schedule becomes tighter. For each job, the processing time was chosen randomly between 1 and 50 and the corresponding due date was established by $D_i = TP_i \cdot N \cdot (1 - T)$, where TP_i represents the total processing time of job *i* which is equal to $TQ_i \times P_i$. In addition, the penalties for both earliness and tardiness are assigned randomly between 1 and 5.

4.2 Performance of the METTS

In this article, the mean relative percentage deviation (MRPD) is used as an index to evaluate the performance of heuristics. The index is described as

$$MRPD = \frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{M}\frac{T_{d(ij)} - T^*_{d(i)}}{T^*_{d(i)}} \times 100$$

Where $T_{d(ij)}$ is the total penalty of heuristic *j* of test problem *i*, and $T^*_{d(i)}$ is the optimal (best) solution for test problem *i*.

In order to make sure the solution obtained from non-split assumption is a good upper bound for the addressed problem, we attempted comparison of the results of our implementation of METTS with that of the BF heuristics (Sule 1997) implementation. We considered eight kinds of problem sizes (*n* = 4, 5, 10, 25, 50, 75, and 100, jobs) with three kinds of due date factors, *T* = 0.1, 0.4 and 0.7. A set of 100 test problems was generated for each case. Thus a total of 2100 test problems are examined. The experimental results are summarized in Tables 1 and 2.

The results shown in Table 1 indicate that when problem size equals to 4 or 5, both METTS and BF heuristics can reach the optimum solutions. However, as the problem size increases to 10, METTS can obtain optimal solutions for all test problems and is slightly better than BF heuristic. Besides, to evaluate the computational efficiency between these two heuristics, the CPU time of heuristic is usually concerned as an efficiency index. The average computational times on PC Pentium D for these two heuristics are less than 10^{-6} seconds for small-sized problems ($n \le 10$). The experiment results (Table 2) reveal that, as the problem size increases ($n > 10$), the METTS heuristic performs better than the BF heuristic in all test problems.

4.3 Benefits of lot streaming

To analyze the impact of job splitting in a single machine manufacturing system, some scenarios are defined and analyzed. These scenarios are defined as the ratios of setup time with its corresponding total processing time (S/P ratio) required for a job which are 1/8, 1/4, 1/2, and 1. The results shown in Table 3 reveal that as the ratio of setup time with its corresponding total processing time less than 25%, there are some benefits for the job splitting. However, as the ratio greater than 25%, the objective which is total earliness and tardiness penalties is not improved any more.

5 CONCLUSION

In this research, to analyze the benefits of job splitting in a single machine manufacturing system with minimization of total penalties objective, a mathematical model is first proposed for solving the addressed problem optimally. Following, a tabu search based heuristic is designed for solving the medium to large size problems. The experimental results shown there are some benefits for the job splitting. However, as the ratio greater than 25%, the objective which is total earliness and tardiness penalties is not improved any more.

6 REFERENCES

[1] Baker, K. R., Pyke, D. F., 1990, Solution Procedures for the Lot-streaming Problem, Decision Science, 21:475-491.

[2] Baker, K. R., and Jia, D., 1993, A Comparative Study of Lot Streaming Procedures, OMEGA, 21/5:561-566.

[3] Baker, K. R., 1995, Lot Streaming in the Two-machine Flow Shop with Setup Times, Annals Operations Research, 57:1-11.

[4] Cetinkaya, F. C., and Kayaligil, M. S., 1992, Unit Sized Transfer Batch Scheduling with Setup Time, Computers and Industrial Engineering, 22/2:177-183.

[4] Glass, C. A., Gupta, J. N. D., and Potts, C. N., 1994, Lot Streaming in Three-stage Production Processes, European Journal of Operations Research, 75:378-394.

[5] Hancock, T. M., 1991, Effects of Lot-splitting Under Various Routing Strategies, International Journal of Operations and Production Management, 11/1:68-75.

[6] Kropp, D. H., and Smunt, T. L., 1988, Optimal and Heuristic Models for Lot-splitting in a Flow Shop, Decision Science, 21/4:691-709.

[7] Potts, C. N., Baker, K. R., 1989, Flow Shop Scheduling with Lot Streaming, Operations Research Letters, 8:297-303.

[8] Smunt, T. L., Buss, A. H., and Kropp, D. H., 1996, Lot Splitting in Stochastic Flow Shop and Job Shop Environments, Decision Science, 27:215-237.

[9] Sen, A., Topaloglu, E., and Benli, O, 1998, Optimal Streaming of a Single Job in a Two Stage Flow Shop, European Journal of Operations Research, 110:42-62.

[10] Sen A., and Benli, O., 1999, Lot Streaming in Open Shops, Operations Research Letters, 23:135-142.

[11] Trietsch, D., and Baker, K. R., 1993, Basic Techniques for Lot Streaming, Operations Research, 41/6:1065-1076.

[12] Vickson, R. G., and Alfredson, B. E., 1992, Two and Three Machine Flow Shop Scheduling Problem with Equal Sized Transfer Batches, International Journal of Production Research, 30/10:1551-1574.

[13] Wanger, B. J., and Ragatz, G. L., 1994, The Impact of Lot Splitting on Due Date Performance, Journal of Operations Management, 12:13-25.

Table 1. Comparison of Performance of METTS (N<10)

Environments		Data Sets	Opti.Solu.	Performance in MAPD	
n	*T*			METTS	BF heuristic
4	0.1	100	323.71	0.00 (100)	0.00 (100)
	0.4	100	149.81	0.00 (100)	0.00 (100)
	0.7	100	173.26	0.00 (100)	0.00 (100)
5	0.1	100	550.80	0.00 (100)	0.00 (100)
	0.4	100	226.14	0.00 (100)	0.00 (100)
	0.7	100	266.78	0.00 (100)	0.00 (100)
10	0.1	100	1178.72	0.00 (100)	0.15 (96)
	0.4	100	979.75	0.00 (100)	0.22 (95)
	0.7	100	994.85	0.00 (100)	0.19 (96)

(*) number of optimum solutions found in 100 test problems.

Table 2. Comparison of Performance of METTS (*n*>10)

Environments		Data Sets	Performance (MAPD)	
n	*T*		METTS	BF
25	0.1	100	0.00	8.23
	0.4	100	0.00	5.34
	0.7	100	0.00	8.95
50	0.1	100	0.00	11.54
	0.4	100	0.00	10.48
	0.7	100	0.00	10.30
75	0.1	100	0.00	13.26
	0.4	100	0.00	12.59
	0.7	100	0.00	14.38
100	0.1	100	0.00	14.88
	0.4	100	0.00	14.63
	0.7	100	0.00	16.72

Table 3. Performance of job splitting in four different S/P scenarios

S/P ratio	Number of jobs splitting						
	1	2	3	4	5	6	7
1/8	0.000	10.12	13.35	14.76	15.57	16.39	16.65
1/4	0.000	9.88	13.65	14.54	15.12	15.76	16.27
1/2	0.000	3.24	3.32	3.32	3.32	3.32	3.32
1/1	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Integration of Process Planning and Scheduling Using Multi-Agent Learning

Nobutada Fujii[1], Riku Inoue[2] and Kanji Ueda[3]

[1] Graduate School of Engineering, Kobe University, Japan

[2] School of Engineering, The University of Tokyo, Japan

[3] Research into Artifacts, Center for Engineering, The University of Tokyo, Japan

Abstract

This paper proposes a new method to integrate process planning and production scheduling. It is difficult to determine a proper plan that meets both objectives simultaneously because optimality of the process plan and the production schedule often conflict. A multi-agent learning based integration method is proposed in the study to solve the conflicts. In the proposal, each machine makes decisions about process planning and scheduling simultaneously, and it has been modelled as a learning agent using evolutionary artificial neural networks (EANN) to realize proper decisions resulting from interactions between other machines. To confirm the feasibility of the proposal, computer simulations are conducted. The results show that role-sharing among machines occurs so that high productivity can be achieved. Comparisons with other integration methods demonstrate the proposed method can achieve better performance than the others in the points of minimizing make-span and also the speed to converge the best solutions. Results also show that the proposed method is applicable to the problems including production demand fluctuations.

Keywords:

Process Planning, Production Scheduling, Multi-Agent System, Evolutionary Artificial Neural Networks,

1 INTRODUCTION

Optimality of the process plan and the production schedule often conflict. The quality of the process plan can decrease when the quality of the production scheduling is preferred; on the other hand, the quality of the production schedule can be corrupted as the process plan is prioritized. It is difficult to obtain a proper process plan and a production schedule under such dilemma-like conditions using traditional planning methods, in which the process-plan optimization and production scheduling are determined sequentially. New methods to integrate process planning and production scheduling need to be developed.

Integration methods have been proposed to resolve such dilemmas posed by process planning and scheduling needs. For example, Chryssolouris and Chan [1] and Zijm [2] introduced process plan alternatives and changed the process plan when the scheduling results are not feasible. Palmer [3] used simulated annealing techniques to integrate process planning and scheduling. Sadeh [4] attempted to implement communication functions between a process planning module and a production scheduling module to integrate them. Morad and Zalzala [5] proposed to integrate them by encoding both the process plan and the schedule into chromosomes in a genetic algorithm (GA). These proposed methods can reportedly obtain a good solution to fulfil both objective functions of process planning and scheduling. However, because most of these methods repeat separated optimizations, they sometimes encountered problems by which the solutions do not converge but are instead periodic. Moreover, it remains difficult for the methods to adapt to environmental changes such as altered production requirements. A method that is adaptable to environmental changes must be developed because market demands often cause turbulent fluctuations.

This paper describes a new method based on multi-agent learning which realizes simultaneous process planning and scheduling that resolves the dilemma between process planning and scheduling. In the proposed method, each machine has a learning unit developing evolutionary artificial neural networks (EANN) [6] proposed in the area of evolutionary robotics. Thereby, it can accomplish simultaneous process planning and scheduling as a result of local interactions among machines. It is desired that the proposed method be adaptable to environmental changes through re-learning performed by each machine after environmental changes.

2 INTEGRATION OF PROCESS PLANNING AND SCHEDULING

2.1 Multi-agent based approach

Optimization of process planning or scheduling, which is usually solved as a combinatorial optimization problem, is a typical example of NP-hard problems in which the search space can easily explode. Furthermore, it is more difficult to optimize process planning and scheduling simultaneously under conditions in which production environments include fluctuations. Under such conditions, the process plan and the schedule must be changed adaptively because of production demand fluctuations.

Reportedly, emergent synthesis approaches are effective to resolve such NP-hard problems and are essential to resolve problems including environmental fluctuations [7]. A new method, based on the emergent synthesis approach to simultaneous process planning and scheduling, is proposed by developing multi-agent system adopting EANN in this study.

In the proposed method, process plans and production schedules are generated through interaction among local decisions of machine agents. A machine agent decision includes: (1) production scheduling – selection of a product to process from waiting products in the machine's buffer; and (2) process planning – selection of a machine to be used for subsequent processing of the selected product. Figure 1 depicts a machine's decision about process planning and scheduling. These selections are performed simultaneously when a machine is about to start a process. This simultaneity enables the system to produce a process plan and schedule simultaneously. Each machine learns to make appropriate decisions that suit the state of the production floor using EANN, which is introduced into each machine. In the EANN, the ANN structure is determined through an evolutionary process: weight and threshold values of an ANN are encoded into a gene. These values are updated using GA. It is expected that an effective and adaptable learning system is realized by introducing EANN.

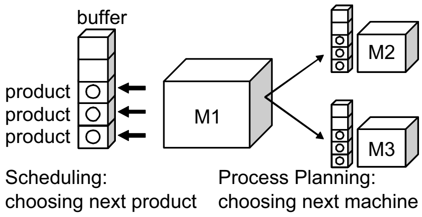


Figure 1: Machine agent configuration

2.2 Setting of EANN in each machine

Figure 2 illustrates the structure of the artificial neural networks (ANN) used in this study. A three-layered feed-forward ANN is employed. Input information to the ANN includes the product types in each buffer, the occupation rate of other machines' buffers, and whether other machines are processing products or not. The input layer has 16 neurons. Outputs from the ANN are selection of a product from the buffer and selection of a machine to be used in the subsequent process: each machine forms the process plan and the schedule simultaneously. The output layer has five neurons; the hidden layer has eight.

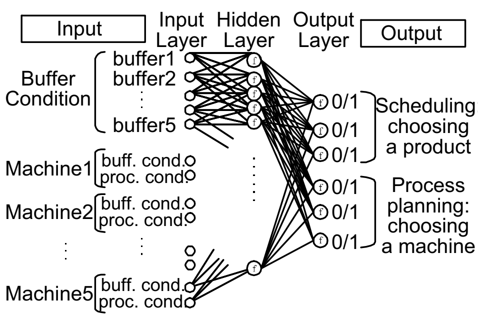


Figure 2: ANN setting in each machine

The weight and threshold values of the ANN are encoded into a genetic string using real values. The number of bits in the gene is set as 318. A multiple-population parallel genetic algorithm [8] is used to maintain high variety among genes. The sub-population size is set to 20. Migration occurs among sub-populations at intervals of 100 generations. These parameters are chosen based on the results of preliminary experiments. An elite strategy is also introduced.

The production result decoded from a gene is evaluated from the viewpoint of productivity. The production objective is set to minimizing the make-span. The fitness value of each gene is calculated as

$$f(t)=\begin{cases}\exp(\alpha * \dfrac{T_{threshold}-t}{T_{threshold}-T_{target}})-P, & (t \le T_{threshold}) \\ 1, & (t > T_{threshold})\end{cases} \quad (1)$$

where t denotes the make-span, T_{target} is the target value of the make-span, and α is a constant number. Also, P represents a punishment calculated by $P = \beta * erro\_no$, where *erro_no* is the frequency of selecting empty buffers by the machine. Finally, β is a constant. A positive evaluation value in the case of a make-span larger than $T_{threshold}$ is introduced to retain variety in the population so that genes having a lower evaluation value can remain.

3 SIMULATION RESULTS AND DISCUSSION

3.1 Problem descriptions

The proposed method is applied to the problem extended from a benchmark problem proposed by Sundaram and Fu [9], which is a job-shop type problem that involves five machines and five job types. Each product has four processes. Table 1 shows the process times (min) achieved by the machines. Each bracketed number is an identification number of the machine which performs the process. The benchmark problem objective is achieving the shortest make-span. In that problem, the make-span can reach 33 min when one job of each type is produced. It was reported that the proposed multi-agent learning based approach can achieve the optimal solutions as well as the previously proposed integration methods [10].

Table 1: Production time used by each machine

Prod.	Proc. 1	Proc. 2	Proc. 3	Proc. 4
1	5(1), 3(2)	7(2)	6(3)	3(4), 4(5)
2	7 (1)	4(2), 6(3)	7(3), 7(4)	10 (5)
3	4(1),5(2), 8(3)	5(4)	6(4), 5(5)	4(5)
4	2(2), 6(3)	8(3)	3(3), 8(4)	7(4), 4(5)
5	3(1), 5(3)	7(3)	9(4), 6(5)	3 (5)

In this study, the benchmark problem by Sundaram and Fu is extended to confirm the effectiveness of the proposed method for more complex and large-scale problems. Concretely, the number of each product to be processed is increased from one to 20: 100 products are produced in all. The search space becomes much larger than the benchmark problem by the extension of the setting.

3.2 Parameter settings of the simulations

The evaluation parameters are set as follows: α = log10, β = 0.01, T_{target} = 500, and $T_{threshold}$ = 800. The parameters for GA are also set as follows: the maximal generation is set to 1000, the population size in a sub-population is set to 20, the elite size is one in a sub-population, and the mutation rate is set to 0.6.

3.3 Simulation results

Computer simulations are conducted to confirm the effectiveness of the proposal. Figure 3 depicts the transition of the make-span during a simulation in which the performance improves as the simulation progresses. Finally, the best solution is achieved at about the 750th generation, even though the system seems to be entrapped with local optima at the middle part of the simulation. The best solution is 500, which is obtained using the proposed method.

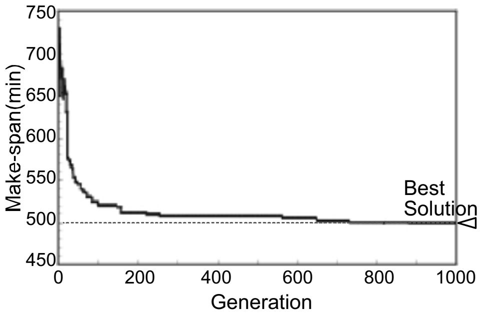


Figure 3: Transition of make-span during a simulation

Figure 4 portrays a Gantt chart obtained from execution of the simulation. The Gantt chart shows that each machine has a certain role: for example, machine 1 processes product types 1, 2 and 5, although the machine can also process product type 3. Furthermore, the machines show tendencies related to time sharing among product types: for instance, machine 1 processes only product type 5 in the early stage of simulation; at the latter stage of the simulation, it processes products 1 and 2.

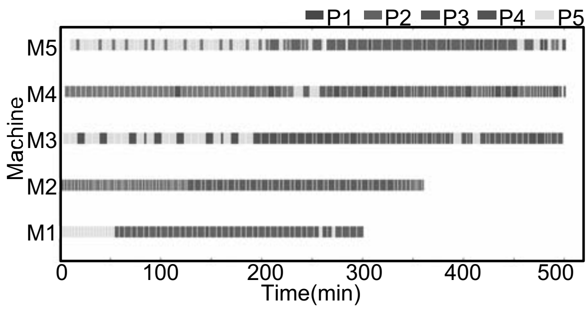


Figure 4: Obtained Gantt chart during simulation execution

Figure 5 depicts the frequency distribution among alternative process plans in each product type. It seems that the product type 1, 2, 4 and 5 uses two kinds of process plans to realize short make-span and changes the two plans properly during the simulation executions. Consequently, the bottleneck machines, machine 4 and 5, can decrease the work-load properly so that the best solution can be obtained where the machine 4 and 5 finished their processes at the same time, 500 steps. These results show that the dynamic role-sharing can occur in the simulation execution. As a result, the role-sharing engenders good performance of the results.

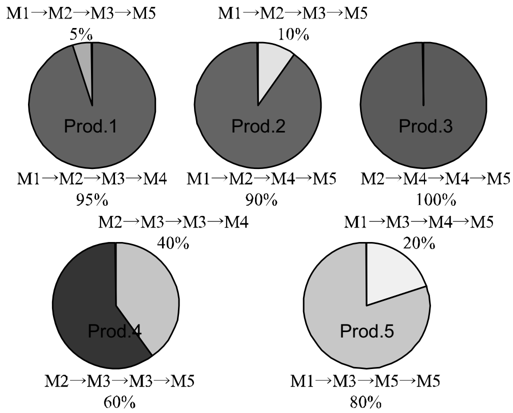


Figure 5: Obtained process plans in each product type

3.4 Comparison with other methods

Comparisons with other methods are made to clarify the effectiveness of the proposed method. Those methods are GA-based simultaneous process planning and scheduling proposed by Morad [5] and the integration method of process planning and scheduling. The Morad method encodes the process plan and the schedule into one genetic string in which the optimization processes are executed simultaneously from the viewpoint of the process plan using plan-resource genetic operation and from the viewpoint of scheduling using order-based and position-based genetic operations. The integration method is realized by alternation of the process plan and scheduling using GA based on Morad's method. Table 2 shows a comparison of the best solutions achieved using the respective methods. The result reveals that the proposed EANN-based method achieves the best solution: 500 min.

Table 2: Comparison of the best solution by each method

	Proposed	Morad	Integrated
Make-span (min)	500	509	513

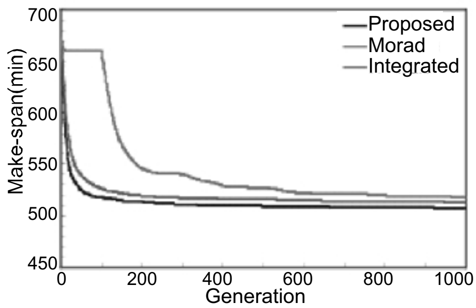


Figure 6: Transition of average make-span in each method

Figure 6 depicts the transition of average make-span among 20 simulation trials using each method. The proposed method not only achieves the best solution at the end of the simulations. It also attains the fastest speed of convergence. These results reveal that the proposed method is applicable to more complex and larger scale problems than other methods.

3.5 Adaptation to environmental change

Demand changes are also introduced to the problem. Table 3 shows the setting of the production demand fluctuations in which three terms of production are used, and in which the product mix is set differently in each term including 1000 generations. Computer simulations comparing Morad's method and the integration method are executed.

Table 3: Setting of production demand changes

Term	Product kinds	Production amount in each kind				
		P1	P2	P3	P4	P5
1	5	20	20	20	20	20
2	4	25	25	25	25	0
3	5	15	15	20	25	25

Figure 7 depicts the transition of the average make-spans of 20 simulation trials that were obtained using each method. The proposed method can obtain better solutions than other methods before and after environmental changes. The result also shows that the incremental value of the make-span by the proposed method remains much less than those of the other methods after the production demand change because the proposed method can use rules that are obtained before the demand fluctuations.

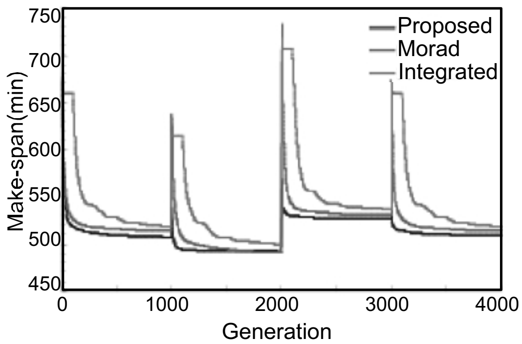


Figure 7: Average make-span in each method

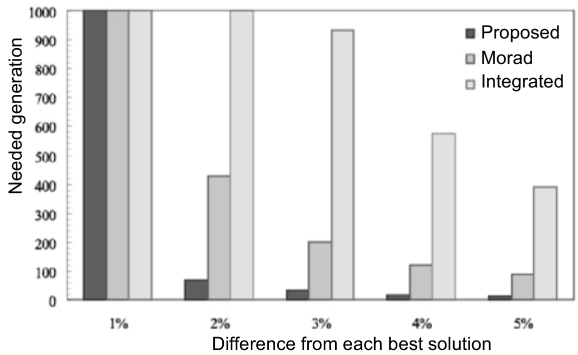


Figure 8: Needed generation to obtain the solutions that are different from 1 % to 5 % from best solutions

Figure 8 illustrates needed generation in which each method can obtain the solutions differing from one percent to five percent from best solutions after the first environmental change at the 1000 generation. The result shows that the proposed method can achieve the solutions differ two percent from the best solution by a hundred generations; the convergence speed of the proposed method is the quickest among the methods. These results support that the proposed multi-agent learning based method using EANN can adapt to environmental fluctuations.

4 SUMMARY

This paper presents a new approach to integrate process planning and scheduling based on multi-agent learning approach: machines employ EANN to produce simultaneous decisions about process planning and scheduling. The system objective is achieved by multiplication of the result of each machine's learning.

Computer simulations were executed using extended data from a benchmark problem. The simulation results showed that the proposed method achieves the best solution. The simulation results in the comparison with other methods demonstrate that the proposed method achieved better solutions than the other methods, and that it also adapts to production demand fluctuations.

5 REFERENCES

[1] Chryssolouris, G., Chan, S., 1985, An Integrated Approach to Process Planning and Scheduling, Annals of the CIRP, 34/1: 413-417.

[2] Zijm, W.H.M., 1995, The Integration of Process Planning and Shop Floor Scheduling in Small Batch Part Manufacturing, Annals of the CIRP, 44/1: 429-432.

[3] Palmer, G.J., 1996, A simulated annealing approach to integrated production scheduling, Journal of Intelligent Manufacturing, 7/3: 163-176.

[4] Sadeh, N.M., Hildum, D.W., Laliberty, T.J., McA'Nulty, J., Kjenstad, D., Tseng, A., 1998, A Blackboard Architecture for Integrating Process Planning and Production Scheduling, Concurrent Engineering, 6/2: 88-100.

[5] Morad, N., Zalzala, A., 1999, Genetic algorithms in integrated process planning and scheduling, Journal of Intelligent Manufacturing, 10: 169-179.

[6] Cliff, D., Harvey, I. Husbands, P., 1992, Incremental Evolution of Neural Network Architectures for Adaptive Behaviour, Technical Report CSRP 256, School of Cognitive and Computing Sciences, University of Sussex, Brighton, UK.

[7] Ueda, K., Markus, A., Monostori, L., Kals, H.J.J., Arai, T., 2001, Emergent Synthesis Methodologies for Manufacturing, Annals of the CIRP, 50/2: 535-551.

[8] Cantu-Paz, E., 2000, Efficient and Accurate Parallel Genetic Algorithms, Springer.

[9] Sundaram, R.M., Fu, S.S., 1998, Process planning and scheduling, Computers and Industrial Engineering, 15/1-4: 296-307.

[10] Ueda, K., Fujii, N., Inoue, 2007, An Emergent Synthesis Approach to Simultaneous Process Planning and Scheduling, Annals of the CIRP, 56/1: 463-466.

Manufacturing System Design

Module Structured Production System

Kuniaki Tanaka[1], Nobuo Nakatsuka[1], Hironori Hibino[2] and Yoshiro Fukuda[3]
[1] Technology Development Center HQ, Omron Corporation, Japan
[2] Technical Research Institute, JSPMI, Japan
[3] Hosei University, Japan

Abstract
Flexible, lean and demand oriented production system enabling efficient manufacture of high-variety low-volume products is desired. In order to realize these needs, production systems should enhance flexibility, scalability and reconfigurability. We propose a new production system which names "Module Structured Production System (MSPS)". The MSPS consists of multiple production lines and manufacturing execution system. The production line is comprised of several production cells consist of function units. By changing the production cells and function units for processing products, optimize production line can be organized, rapidly. In this paper, the concept and the structure of MSPS are discussed.

Keywords:
Reconfigurable Production; Facility Design; Modular Design;

1 INTRODUCTION

There is a prediction that the competitive edge of Japan in the world would decrease continuously from now on by rising of the BRICs including China. According to "Goldman Sachs (2003) Dreaming With BRICs: The Path to 2050, Global Economics paper No. 99," for instance, GDP of Japan in 2050 would be expected to be smaller than one-fourth of China, one-third of the United States, and half of India.

The downward trend of Japan's global competitiveness is thought to be unexceptional even for Japanese manufacturing industry possessed an overwhelming advantage. For Japanese manufacturers to maintain and improve global competitiveness in the future, strengthening power against rapid market expansion to global market including China and abrupt price drop by low-cost production such as China is necessary. In addition, measures to avoid environmental impact, strengthening power to social needs such as safety, and security are becoming particularly important issues. [1]

Under such environment, for Japanese manufacturers to maintain and expand more competitiveness in manufacturing, it is necessary to realize highly integrate production system, most lean, rapid and flexible production, and continually strengthen core technology to integration level of product, production of more added-value products are becoming absolute conditions. For these purposes, it is necessary to recognize that a production system responding to wide-ranging customers' and social needs, enabling production of high-mix high-variation product, in other words development and construction of flexible, lean and rapid production system limiting loss to the minimum in manufacturing process are important issues for Japanese manufacturers to realize.

In this paper, as one realistic method to solve the above issues, we propose a Module Structured Production System (MSPS) which enables complete build-to-order production of assembly-type of products.

2 MODULE STRUCTURED PRODUCTION SYSTEM

2.1 Definition of Module Structured Production System

The Module Structured Production System (MSPS) consists of multiple production lines and manufacturing execution system. The production lines are comprised of several production cells which consist of function units. Figure 1 shows structural elements of the system. The function units hold attributes which define their processing procedures. This way, the MSPS becomes a production system having components with a physical hierarchical structure.

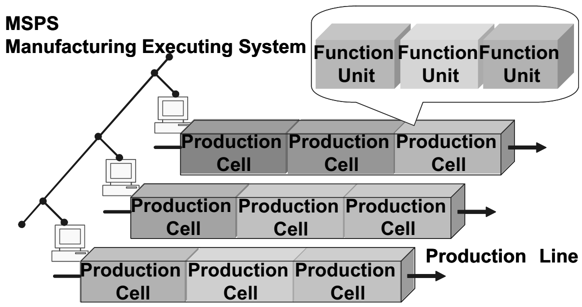


Figure 1: Structural Elements of the MSPS.

Recently there are many proposals of production system using module and cell concept [2][3] presented to increase flexibility, scalability and reconfigurability. The MSPS we propose in this paper is also aimed at increasing flexibility, scalability and reconfigurability. As discussed formerly, by putting individual functions to the production cells that have physically-layered structure and the MSPS execution system, combining these elements to the best structure and synchronizing the combination, the most efficient production system is established.

The production line consists of process groups called multiple production cells that are essential to commercialize product, an essential components of the MSPS. At each production line, materials are charged into the production cell which handles the first process, and by working and assembling them through a series of the production cells, finished product is put through the last production cell.

Each production cell receives mid-products derived from the previous production cell. After adding elemental works such as work, assembly, inspection, and packaging required at this particular production cell to mid-product, it is carried to the next production cell. Production cells consist of one or several function unit(s) and are responsible for one complete structural unit of product.

Function units are incorporated into production cell. This is the minimum physical unit enabling individual elemental works completely. The elemental works mean, for instance, works such as solder coating, soldering, resin coating, screwing, and assembling. In order to correspond to differences of product's shape and work conditions, processing procedure of the function unit and attributes defining work and assembly are set.

The MSPS is an on-site information system which combines information transmitting function and functions required for production execution such as production history management, process condition management, facility operation management, and product quality management. Principally, it consists of a server which has interface with main-frame systems and PC-based client terminals situated at each production line. The client terminal of production line is connected through each cell which consists of production lines and releases required information for production execution at required timing.

2.2 Concept of Product Realizing Module Structured Production System and Production Line Structure

In order to realize the MSPS, production line, production cell, and structure of function unit must be decided hierarchically.

Process-assembly products we are discussing here are consisted with components generally, and the components are consisted with assemblies of several parts. In order to build component with these parts, works such as processing, connection, and assembly are done. When we consider them as production elements, required elemental works for each component can be decided. [4]

On the other hand, the MSPS consists of multiple production lines and manufacturing execution system as described. The production lines are comprised of several production cells which consist of function units. By deciding required function unit for each production elements, composition of parts and production line can be fixed as well-balanced structure.

For breaking down product into suitable components and parts, it is desirable to consider with the following procedures:

1. Definition of objective product groups: It is requisite to determine customers' customizing needs carefully.
2. Using Quality Function Deployment (QFD) [5], deploy required product functions whose product groups defined to function system diagram.
3. Define the best components and parts anticipating future change of the specifications by adding customizing needs from customer with the viewpoint of efficient high-variation production. Here, the method and concept of Group Technology (GT) [6] is effective.

Figure 2 shows structural concept whose product and production line compositions layered and each structural elements matched. As shown, production line is provided for product, production cell for products' component, function unit which is constituent element of production cell for production elements correspondingly. These relationships become the basic concept of product and production line structure to realize a module structured production system.

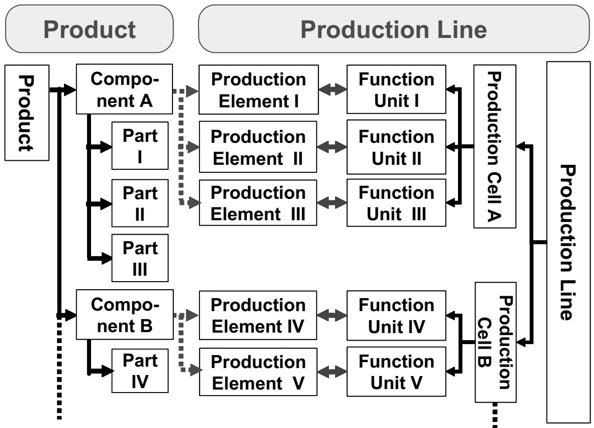


Figure 2: Concept of product and production line structure.

2.3 Structure of Production Cell

Production cell comprises of combination of several function units as described in the former section. Each function unit may be added or removed individually. For changing the combination of function unit, by simply removing a function unit (not requiring special settings), operation starts. Furthermore, function unit sets processing procedure and attributes which define assembly conditions. Figure 3 shows a structural image we propose.

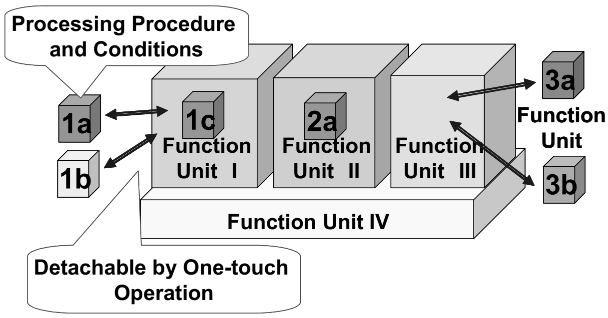


Figure 3: Structural Image of Production Cell.

3 FEATURES OF THE MSPS

3.1 Changeable to Different Model by Flexible Arrangement

The MSPS corresponds to production line of different models efficiently by flexibly arranging the production cells for the best function unit of model to be manufactured. In other words, for manufacturing various models flexibly and performing the best production efficiency, it is necessary to change combination of the function unit and production cell optimally at any time.

There are seven patterns of production lines to change product to be manufactured by different combination of the function unit and production cell as shown in Figure 4.

1. Changing Attributes of the Function Unit

When objective products belong to the same product group, it is general that they have the same shapes but different specifications and performances. For example, when work, control parameter of assembly machine (temperature profile of reflow machine) or attaching jigs are changed, processing procedure and conditions of the function unit need to be changed. But this could be solved by changing attributes of the function unit.

2. Adding More Function Units

When objective product is changed, processing capability of specific function unit would become insufficient and bring loss balance in between previous and next function units. When this happens, by adding the same function unit and combining identical several function units, performance of this function unit would be increased and the tact balance between previous and next function units could be coordinated.

Additional work or assembly not required before changing product may be necessary. In this case, addition of new function unit to the appropriate production cell would respond to the change.

3. Removing the Function Unit

On the contrary of the case 2., by changing the product to be manufactured, performance of the specific function unit would be surplus and bring loss balance in between previous and next function units. This time, by removing exceeded function unit from the function units, performance of this function unit could be decreased and tact balance between previous and next function units could be coordinated.

By changing products, the function unit responsible for work or assembly which will be unnecessary after the change can be removed.

4. Replacing the Function Unit

When product to be manufactured is in the same product family and product series are different, product shape would largely differ in general. In this case, depending on the shape, the best elemental work (method) would be also different. As the function unit is a unit performing single elemental work as previously written, it cannot respond to different elemental work. This time, replacing a function unit could solve the problem.

5. Adding the Production Cell

When structure of product and part configuration is different, change of the function unit would not be sufficient. It is necessary to install production cell matching components configuration into the production line.

6. Removing the Production Cell

On the contrary of the case 5., optimal production line for product to be manufactured would be arranged by removing a production cell.

7. Replacing the Production Cell

Furthermore, by replacing a production cell, optimal production line could be arranged.

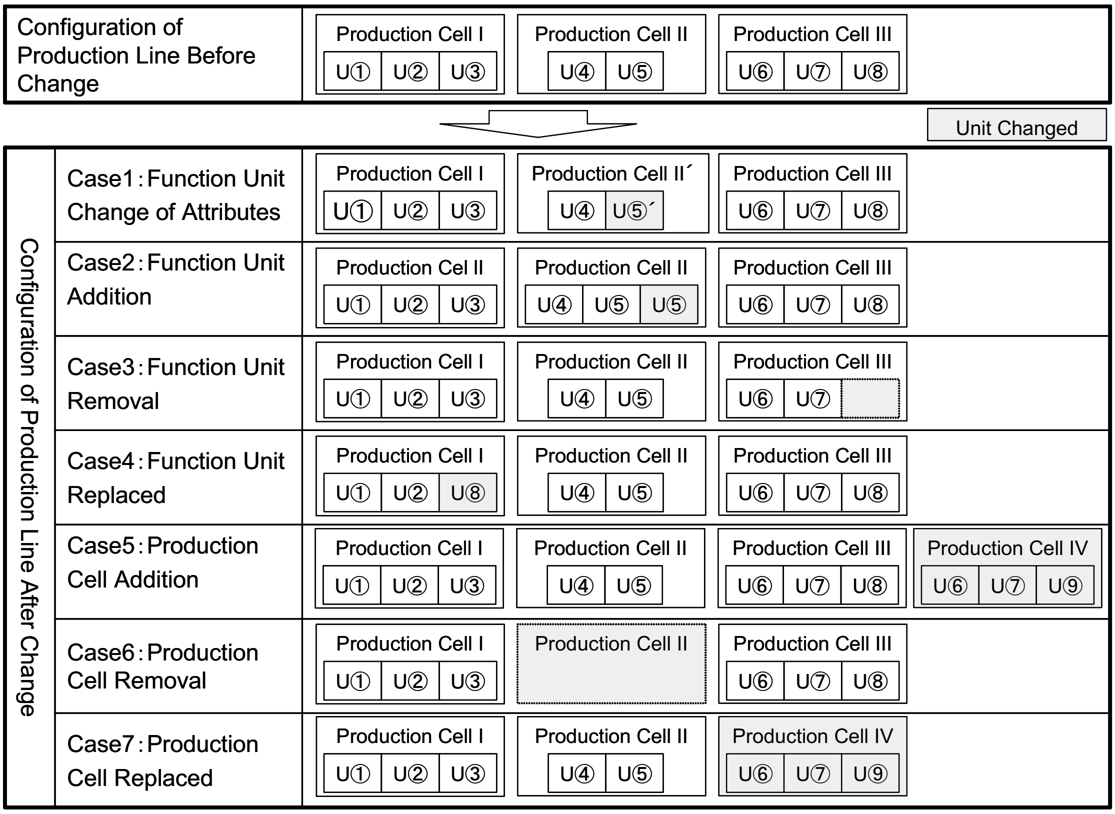


Figure 4: Module Structured Production System, Flexible Production Patterns.

3.2 Flexible Response by the MSPS

In the former section, we discussed the MSPS manufactures various kinds of products flexibly responding to various volumes. For fulfillment of the MSPS aim, efficient manufacture responding to various volumes and various kinds is necessary. The MSPS realizes it mainly by executing its functions. Figure 5 shows an image of the MSPS.

Efficient production system is a system whose productivity does not depend on lot volume. This means productivity does not change when producing a large number of the same product at once or only one piece of product. As formerly discussed, the MSPS has functions receiving customers order directly and dispatching it to the production line and controlling functions required for production such as production history, process progress, facility operation, and product quality. There are two important requirements for the MSPS to produce various volumes.

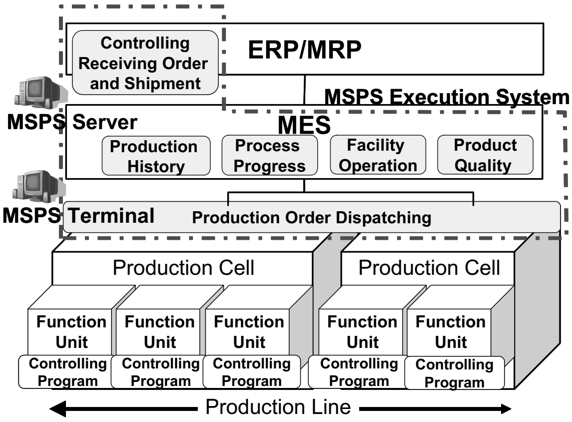


Figure 5: Image of the MSPS.

The first requirement is that accumulation of information should not occur between manufacturing execution system and production line. To meet this requirement, the MSPS realizes a system transmitting required information for production execution between the MSPS terminals and control devices of production cell at required timing. A common production control system consists of main-frame systems such as ERP/MRP working on the factory computer, work station and manufacturing execution system (MES) working in the client server system of personal computer base. Functions of receiving order and shipment, production schedule are processed in the main-frame systems and those relating to production execution are processed by MES. The main-frame systems receive new order online during the day time but further process such as reflecting order to production schedule is done by batch process at night. Therefore loss of more than one day happens until the order reflecting the latest information is put down to the production line. In order to solve this problem, the MSPS incorporates interface function which receives order information from customers directly and dispatches production order to the production line. As a result, loss time of data transmission could be minimized and order information from customers could be transmitted to the production line almost in real time.

The second requirement is that production order responding to the condition of production line should be dispatched at the necessary timing. Especially for small volume of production lot, short dispatching cycle of production order is required. In common production control system, production order is dispatched in conjunction with its scheduling as a function of the main-frame systems. On the other hand, it cannot say that the MES always obtains condition of production line based on the information from each production facility and the main-frame systems obtain the latest condition. Therefore, if customer places an express order, it is difficult to transmit changes and corrections for the best production order in real time reflecting the condition of production line. The MSPS solves this problem by implementing a function of obtaining all the order data from customers and the latest condition of production lines and dispatching production order as necessary as required. Whenever the MSPS server receives order from customer, required information is transmitted through the MSPS terminals which are set at each production line and monitor operation conditions of the applicable line and the best production line for manufacturing order is decided. This enables control of production line in short cycle and makes production line adjustable. Figure 6 shows dispatching image of production order.

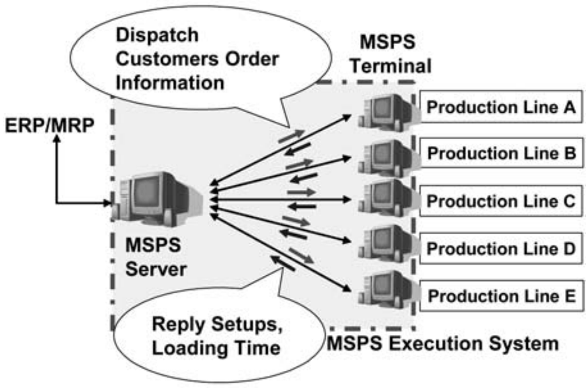


Figure 6: Dispatching Image of Production Order.

4 EXAMPLE OF THE MSPS

The MSPS of programmable logic controller (PLC) and temperature controller are shown below. These products consist of PCB (electronic components mounted on PWB), connector for external interface (terminal board), power supply unit, and chassis. Elemental works required are solder coating, components installation, soldering, assembly, and inspection. When we show PLC product structure and required elemental works in the product structure concept (shown in Fig. 2), it will be as Figure 7.

Production line of PLC consists of multiple production cells comprised of function units corresponding to each work as shown in Figure 7. In Figure 8, relationship of elemental works and function unit as well as structure of production line is shown. Figure 9 shows a layout image of actual production line. Figure 10 shows examples of production cell.

When production line needs to be changed because of customers order change on the PLC production line shown in Figure 9, there are seven patterns of function unit and production cell as shown in Figure 4. Examples are shown below:

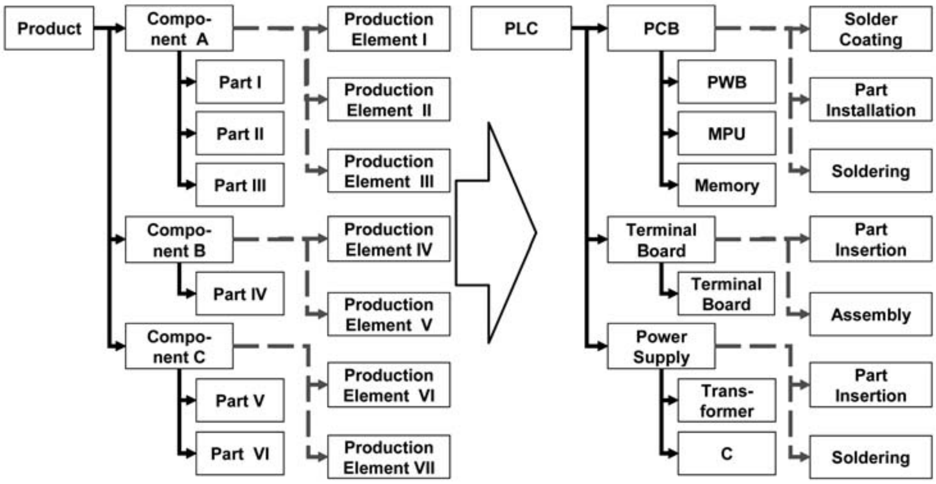


Figure 7: PLC Product Structure and Required Elemental Works (Partial).

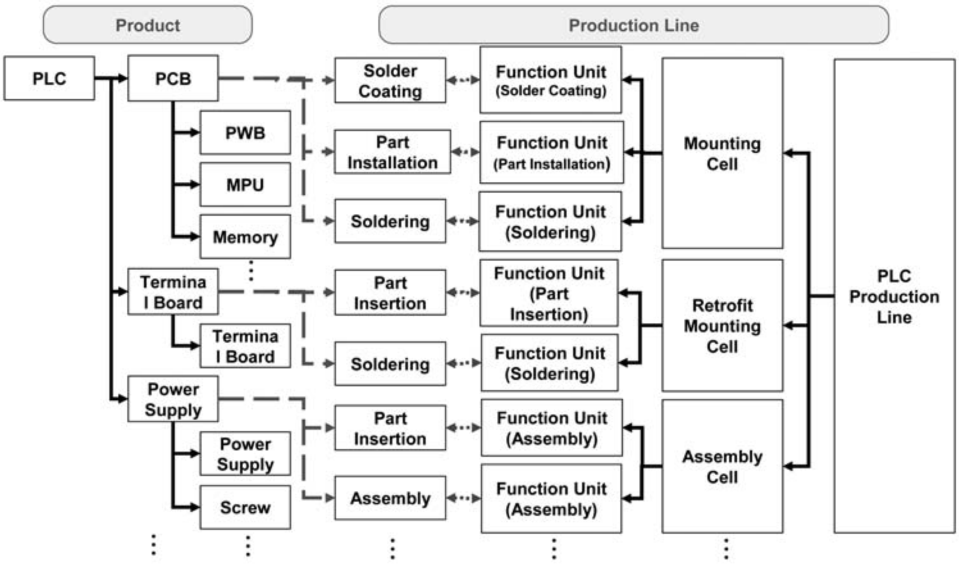


Figure 8: Structure of PLC Production Line (Partial).

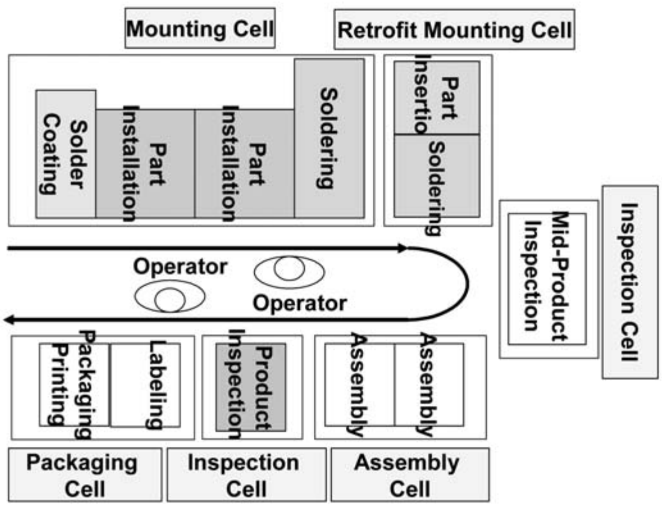
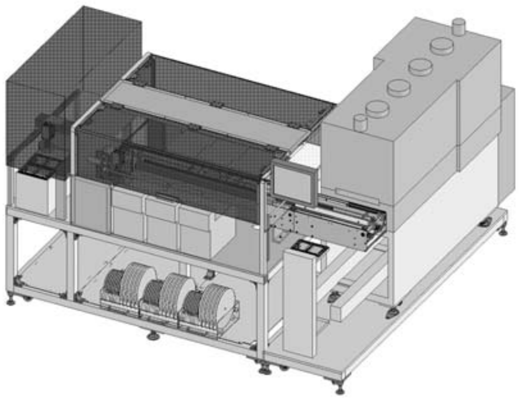


Figure 9: Layout Image of Actual Production Line.



(a) Mounting Cell

(b) Inspection Cell

Figure 10: Examples of Production Cell.

1. Changing Attributes of the Function Unit

When model to be manufactured is changed within the same product group, setup can be adjusted by changing the attributes data of each function unit for many cases. For example, by changing component installation unit of mounting cell, the location data of mounting components are changed. By changing mounting cell soldering unit, heating profile for soldering is changed.

2. Adding More Function Units

Even though the change is made within the same product group, when the shape and size change largely, changing the attribute data of the function unit is not enough; the function unit itself needs to be changed. For example, when number of soldering point increases, solder coating unit of the mounting cell does not satisfy tact time only with dispense method. The method needs to be changed to mask printing which many points are soldered in short time. In this case, solder coating unit of the mounting cell needs to be changed from the function unit corresponding to dispensing method to that corresponding to mask printing method.

3. Removing the Function Unit

When product to be manufactured is changed to different product group, not only function unit but production cell need to be changed. For example, a change to retrofit mounting cell is possible. When retrofit mounting method is changed from manual to robotic, production cell of manual mounting is changed to that of robotic mounting.

5 CONCLUTION

In this paper, we proposed the Module Structured Production System (MSPS) as a system enabling complete built-to-order production of assembly-type of product as an efficient production system of high-mix high-variety products responding flexibly to wide-ranging customers and social needs.

The Module Structured Production System (MSPS) consists of multiple production lines and manufacturing execution system. The production lines are comprised of several production cells which consist of function units. Each production cell consists of several function units and their elements work as production system having physical hierarchical structure.

The MSPS has a feature that its combination can be optimized easily and flexibly depending on the objective product by combination of function unit and production cell. The MSPS also has a feature of transmitting customers order information to the production line without delay. Using these features of the production line and MSPS, efficient production of high-mix high-variation product in the same production system is enabled.

From now on, we will verify efficiency of the MSPS for high-mix high-variation production proposed quantitatively and its effectiveness as a production system which enables complete build-to-order production of assembly-type of products, which is a goal of the MSPS.

REFERENCES

[1] K. Ueda, A. Markus, L. Monostori, H. J. J. Kals, T. Arai (2001). Emergent Synthesis Methodologies for Manufacturing, CIRP Annals

[2] S. M. Ratchev, H. J. Hirani (2003). Requirement Engineering for Re-configurable Assembly Cells, CIRP Journal of Manufacturing Systems, Vol.32, No 4

[3] Y. Shimomura, Y. Fujimoto, S. Yajima, S. Kondo (2005). Proposal of Stratum-Type Manufacturing System based on Cellular Machine, CIRP Journal of Manufacturing Systems, Vol.34, No 4

[4] N. Nakatsuka, T. Hotta, H. Kamoda, Y. Fukuda, S. Sugawa (2006). Development and Operation of a Synchronized Production System Using a Virtual Production Concept, Journal of Japan Industrial Management Association, VOL.57, No.4, pp.272-277

[5] Jack B. Revelle, John W. Moran, Charles Cox (1998). The QFD Handbook, John Wiley & Sons Inc

[6] I. Ham, K. Hitomi, T. Yoshida (1985). Group Technology - Applications to production management, Kluerw-Nijhoff Publishing

Lean Production System Design from the Perspective of the Viable System Model

Christoph Herrmann[1], Lars Bergmann[1], Philipp Halubek[1], Sebastian Thiede[1]
[1] Institute of Machine Tools and Production Technology (IWF), Technical University Braunschweig, Germany

Abstract
Within the design stage of lean production systems, decisions regarding production system design parameters in terms of lean methods have to be taken. While the economic effects of lean methods have been described in practice, there is still a demand for a scientific basis that allows to explain the mode of action of lean methods. The Viable System Model (VSM) supports the analysis of mechanisms within companies, and provides invariant structures for ensuring the viability of organizations. Based on the findings of the VSM, the connection between the lean production system approach and the cybernetic VSM approach are investigated and lean methods are described in terms of attenuating and amplifying variety.

Keywords:
Lean Production; Viable System Model; Variety Engineering

1 INTRODUCTION

The development of Lean Production Systems (LPS) has become a common approach in manufacturing industries [1] [2] and more and more companies design their production system according to lean methods such as Kanban, Just-in-Time, or Total Productive Maintenance. To successfully develop LPS, knowledge on lean methods is required for the design and adaptation processes. As there exists many lean methods as possible production system design parameters, decisions regarding lean methods are of special importance.

In practice, lean methods are often chosen in compliance with the Toyota Production System and without considering company specific strategic and organizational aspects [3]. As a consequence, decisions on the selection and introduction of lean methods often lack an organizational perspective, neglecting the specific company's characteristics like strategic and structural aspects. Against that background, a scientific basis is required, that allows to explain the mode of action of lean methods in an organizational context. Against that background, the objective of the paper is to gain deeper understanding of lean methods within an organizational context. Based on the Viable System Model of Stafford Beer, this paper investigates the way lean methods and their mechanisms are nested and take action within organizations.

2 LEAN PRODUCTION SYSTEMS

The purpose of production systems is to take raw materials, information and resources and transform them into valuable products or services. Nevertheless, a part of the materials and resources is transformed into wasted material and resources. Production systems consist of all required functions, processes, activities, resources, principles, and methods within an enterprise to produce marketable goods, services, or a combination of both. To maximize the efficiency of production systems, the so called LPS approach has been developed. The root of LPS can be backtracked to the Toyota Production System. The founder of Toyota, Sakichi Toyoda and the engineer Taiichi Ohno developed the Toyota Production System (TPS) in the early fifties [1].

The TPS can be regarded as a general framework and philosophy to organize the manufacturing facilities and processes at Toyota and the interaction of these facilities and processes with the suppliers and customers to provide best quality, lowest cost, and shortest lead time through the elimination of waste such as overproduction, wasted time, wasted operator motion, inventory, or production of defect parts [4]. The TPS is maintained and continuously improved through iterations of standardized work and kaizen, following the PDCA (plan, do, check, act) cycle. Toyota was capable of significantly reducing cost and inventory using the TPS, enabling it to become one of the ten largest companies in the world. As the main goal of the TPS is to eliminate waste, TPS is generally known as lean manufacturing.

Due to the success of the TPS and the need of companies to continuously improve competitiveness in global markets, the TPS framework and various of its inherent production system design methods have been copied by many European production enterprises. Most of them did so by copying the concept of the TPS and renaming it to take ownership of it as their own system. As these TPS adaptations show significant accordance, they can be summarized as Lean Production Systems (LPS). But in contrast to the mentioned similarity of different companies LPS, the adaptation, implementation sequence, and the achieved results with these LPS show significant differences.

While LPS comprising of lean methods have been implemented in European companies, the performance as measured by labor productivity improvements and inventory reductions of Toyota still could not be reached by most European companies [5] [6]. Despite the fact that, for example, some German companies have reached a level of best practice in one or more LPS methods, most of them still lack in an overall high performance LPS [7]. A recent review of UK competitiveness emphasizes that European companies still lag behind world class productivity and points out that poor company specific adoption of best practices is a contributor to the UK productivity gap [8]. The poor adaptation of lean methods is caused by multiple factors, such as insufficient deep understanding of the lean philosophy, and deficient comprehension of interrelations between lean methods. Moreover, a lacking holistic organizational perspective neglecting complexity and the connection of management, operations, and the environment impedes an adequate implementation [9].

3 MANAGEMENT AND COMPLEXITY

In order to gain deeper understanding of the relation of management, operations and the environment as a prerequisite for production system design, a systems perspective can be applied. While the operations of an organization are directly linked to the environment, the management is directly connected to the operations. In the context of rapid changing markets, new requirements emerge in the environment that have to be fulfilled by flexible operations. For that, it is necessary, that the management takes right decisions on the production system design, technologies, products, etc. The role of operations that buy, use, develop, and sell technologies and products is essential in this context, as operations represent the link between the environment and the management (Figure 1).

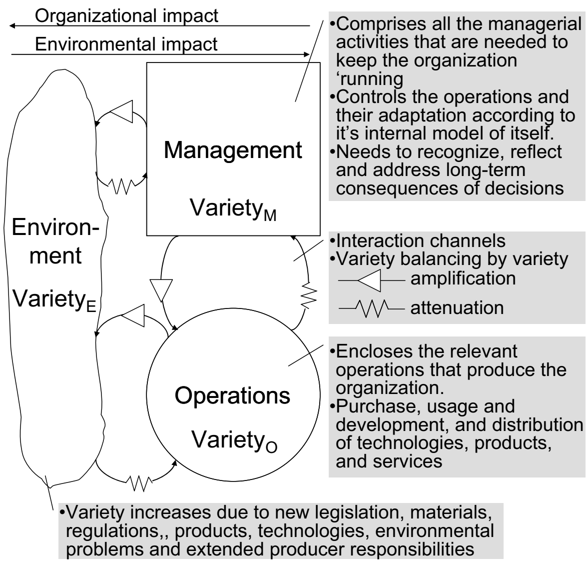


Figure 1: Schematic representation of the need for attenuators and amplifiers.

The environment, the operations, and the management can be regarded as systems with a specific degree of complexity. Since complexity is something perceived by an observer, the complexity of a specific system can be described as a measure of the perceived effort that is required to understand and cope with the system [10]. To be able to compare different systems such as organizations regarding their complexity, a measurement for complexity is required. The cybernetic concept of variety proposes a solution for this problem. The term Variety was introduced by W. Ross Ashby to define the number of possible states of a system [11]. Variety in terms of the count of the possible number of states of a system is a measure for complexity. Considering a set of distinguishable elements of a system, variety means either the number of distinct elements, or the logarithm to the base 2 of this number. A certain situation that has a variety of 32 will take five 'yes/no' decisions to eliminate the implicit uncertainty in that variety – because $32 = 2^5$ [12]. Consequently, the concept of variety as a measure enables to compare the complexity of different systems.

The condition for dynamic stability of complex systems under perturbation is described by Ashby's Law of Requisite Variety as it is one important driver for the design of systems and complexity reduction: "Only variety can destroy variety" [11]. In order to be effective, a control system must be at least as complex and have as many potential behavior patterns as the system to be controlled. This forms a problem for management because in order to make a system flexible and responsive to change, it is necessary to possess as much variety as the system itself exhibits. With systems that exhibit massive variety, such as organizations, only reducing the environmental variety or increasing the management's own variety enables to cope with this problem [13]. If a system such as an organization is to be stable, the number of states of its control mechanism must be greater than or equal to the number of states in the system being controlled. Therefore, variety needs to be managed actively along all communication channels.

The challenge for management is to balance the varieties of the environment and operations as well as the variety of the operations and the management via appropriate attenuators and amplifiers (Figure 1). Organizational measures that attenuate variety decrease the number of possible states, e.g. by filtering unimportant information. Measures that amplify variety allow to increase the number of possible system states, e.g. by hiring new employee with new knowledge. Table 1 gives a brief overview of well-known measures to attenuate and amplify variety.

Table 1: Measures to attenuate and amplify variety.

Measures to attenuate variety	Measures to amplify variety
Standardization: communication, processes, products, etc. *Filtering & Ignoring*: unimportant or unnecessary information; dealing with exceptions/aggregations only. *Modeling to gain system comprehension:* Model the behavior of organizations and the environment.	*Hiring and Training*: hire new and train existing employee. *Empowering*: delegate and empower employee. *Cooperation*: Cooperate with consultants or companies. *Extending products and services:* Develop, combine, or customize products and services.

4 THE VIABLE SYSTEM MODEL

The Viable System Model (VSM) developed by Stafford Beer comprises of a sufficient and necessary set of organizational functions for the viable organization, whereas viable refers as sustained identity within a particular context [14] [12]. The VSM defines a set of functions for organizations that need to be fulfilled in order to become or maintain viable. The VSM is a composed set of operations, a meta-system, and the environment and it provides structures of organizations including key processes, communications and information flows. A VSM is composed of five interacting subsystems (Figure 2):

- The Systems 1 (S1) stand for everything that is done in the organization, i.e. for its operations.
- The System 2 (S2) comprises all activities and resources involved in the coordination between the operative units.
- The System 3 (S3) is responsible for all activities and resources that bring about the optimizing of the operations of the individual systems.
- System 3* (S3*) is closely linked to S3 an provides information for S3 about the state of S1 operations by sporadic audits.
- The System 4 (S4) stands for all the activities and resources that serve to observe the environment and to gain experience from it and to derive and develop strategies to be developed for the future.
- The System 5 (S5) represents all the normative rules and regulations that apply in the organization, such for example ethical attitudes and normative rules.

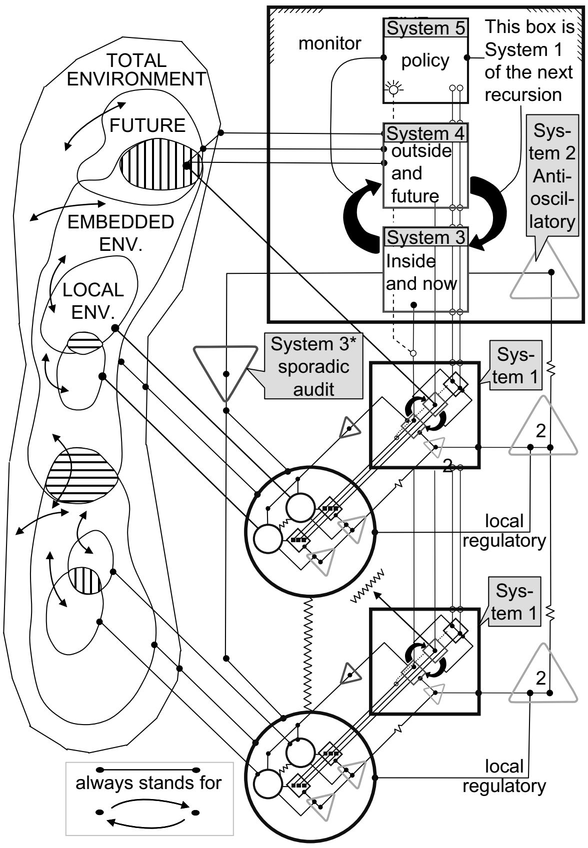


Figure 2: The Viable System Model [14].

The five subsystems link together their functions to provide a configuration of interconnected information loops that convey information generated from and by management activities.

- S1: Information on what the operating systems are doing and information that passes between the operating systems and what is directing them.
- S2: Information that is used to ensure coordination between all the operating systems.
- S3: Information that is used to optimize the operating S1 systems.
- S4: Information that all the operating systems generate from those parts of their environment with which they interact and information from the environment and from the organization that is important to development in the organization and of it.
- S5: Information of a normative nature.

The VSM can be interpreted as two types of systems carrying different types of activities; a system and it's meta-system. The first one consists of several S1 and is concerned with the operational work of producing whatever it is which constitutes the identity of the whole. The second can be described as meta-systemic to the collection of S1. It does not take part directly in the production activities of the lower-level collection as it can be regarded as a controller of both the internal relations between the viable subsystems and the relation of the whole to their environment. The organization of the collection of systems and it's meta-system is important as it shows the way of organization by recursivity.

The VSM is a recursive system, where each VSM contains and is contained within other viable systems [14]. Each viable system comprises several other systems, each viable in itself, and is in turn part of a viable system. All viable systems can be managed autonomously based on the principle of cell division. To make this possible, they must have the functions and information loops describe above. The *Law of Cohesion for Multiple Recursions of the Viable System* defines the conditions for the cohesion of different recursive levels: The S1 variety accessible to S3 of Recursion x equals the variety disposed by the sum of the meta-system of Recursion y for every recursive pair [14].

Based on the insights of the VSM Beer identified three so called Management Axioms as well as four principles of organization that describe the interaction in and between organizations in terms of balancing the variety of information and balancing of autonomy and cohesion [14]. To cope with the variety of a system, the variety of a controller can be amplified or the variety of the system to be controlled can be attenuated. For that, different pattern of variety engineering can be distinguished [12]. The attenuation (AT) of variety can be achieved by:

- *Structural attenuation* (AT1); e.g. by minimization and filtering of input information or delegation of responsibilities.
- *Planning and prioritization* (AT2); e.g. future planning with adequate time horizon and adequate detail level.
- *Operational attenuation* (AT3): e.g. by usage of highly aggregated figures or strict administration with minimal individual decision scope.

To amplify (AM) the management complexity of a controlling system, there exits three general means:

- *Structural increase* (AM1); e.g. by forcing teamwork or job rotation to share knowledge and experiences.
- *Capacitive increase* (AM2); e.g. by employing more managers; employing managers with new qualitative skills or engaging external consultants.

- *Informational increase* (AM3); e.g. by improving the level of information of managers by comprehensive information systems.

Means for attenuating and amplifying variety are often two sides of the same coin. For example, a production plan only contains information for specific assembly operations, and thus attenuates the variety of the product complexity to the worker. At the same time, the production plan increases the variety of the worker, as he can autonomously and goal-directed fulfill his operations in a turbulent environment according to the production plan and without further instructions.

Another very important aspect of the VSM is that the organizational structure needs to comprise of variety amplifying and attenuating mechanisms for adaptation to both its external operating environment and the internal environment on all recursive levels. To maintain their viability, companies as complex organizations within turbulent environments need to develop structures and mechanisms that allow adaptation in all sub-systems and recursions. The mechanisms for organizational adaptation need to be implemented by the structure of the organization and must comprise of the ability to change structures, operations, or the way the operations are linked to the environment.

5 LEAN PRODUCTION SYSTEM DESIGN FROM THE PERSPECTIVE OF THE VSM

5.1 Employee Variety and Autonomous Groups

The VSM as one of the most approved model for the analysis and representation of organizations and their complexity-related questions [15] has a high potential to support the design and diagnosis of lean production system design methods at the structural level. Taylor (1911) claimed that management should precisely define all tasks of a job and define even the exact methodology of the work [16]. Any attempt of the worker to design his own job was not permitted. From a cybernetic perspective, Taylor assumed that a worker was a source of entropy in the workplace [16]. Taylor succeeded in minimizing the variety of work in order to suppress the potential for entropy by simplifying jobs. Thus, the essence of scientific management is the design of low-variety jobs that any man can do. This technique of variety reduction worked good for rather algorithmic repeating procedures where the work does not require heuristics.

In this rather autocratic mode of management, S1 has requisite variety on the horizontal loop and S3 has requisite variety on the vertical loop (Figure 3). But the S1 management unit has no freedom as its requisite variety in the operational circle is supplied to it by S3 [14]. As a consequence, the potential variety of S1 can not be utilized to cope with external variety and the S3 management is extremely involved into the daily business.

Within the turbulent market environment where customers demand more customized product variants, the variety flowing into the companies dramatically increases. This proliferating variety can only be reduced by invoking and using the employee's variety and by limiting the number of product variants. Management must try to focus and use the employee's variety instead of suppressing it. The conversion from human variety suppression to variety enhancement is of central importance and a major element within the lean production system approach. Workers with a shared corporate value system and corporate goals embedded have become an extension of management as they are working on the problems and variety of the markets, adding their variety on the side of the management [17]. By that, the worker supports the management to control the markets. For that, management must try to use the employee's variety and participation is not only desired but apparently essential.

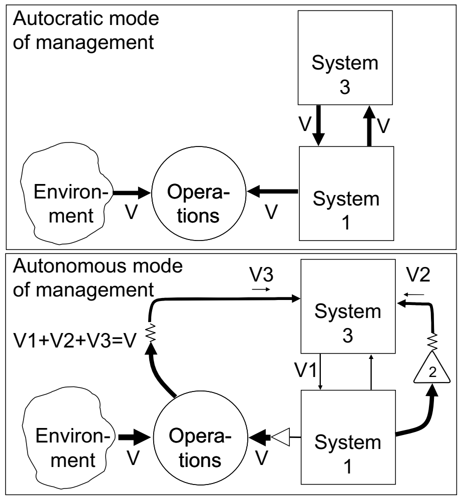


Figure 3: Autocratic and autonomous management [18].

In the autonomous mode of management for the same problem, meta-systemic intervention is minimized, as S1 has requisite variety on the horizontal loop by the amplification of its own variety plus variety stemming from S3 (Figure 3). S3 itself has requisite variety on the vertical plane, since the sum of V1+V2+V3 equals V.

The employee involvement in terms of an autonomous mode of management like claimed in the Toyota Production System approach, also called 'excellence through people' seems to rest on the mechanisms of control variety amplification in compliance with Ashby's Law. Figure 4 show the cybernetic structure of a production company and an autonomous working group [18].

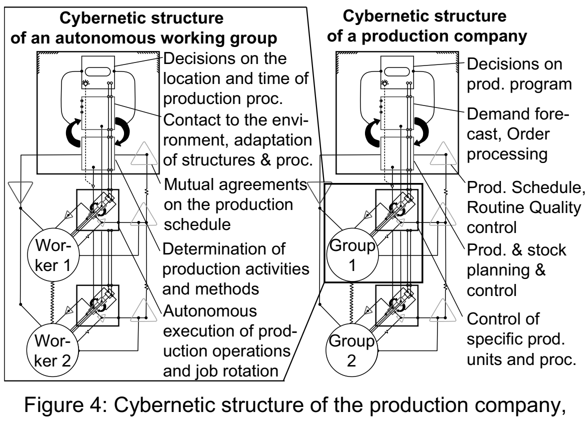


Figure 4: Cybernetic structure of the production company, according to [18].

5.2 Structural Design of Production Systems

According to the VSM and the autonomous mode of management, the structure of production systems in terms of *segmentation* can be derived (Figure 5). Production segments are organizational sub-units of the production system with dedication to integrated processes. A process-oriented system design individually integrates certain collections of activities of the managerial tasks within production segments to fulfill manufacturing processes. According to the VSM structure, the objective of segmentation and structural design of S1 production units should be aligned with the integration of all functions required for viability: operational functions, audits and coordination functions, decision making and adaptation functions. Due to that, lean production segments strive to unify direct functions (operational/manufacturing processes) and decentralized indirect functions. These functions should be unified within production units to effectively and efficiently fulfill processes [19] [20].

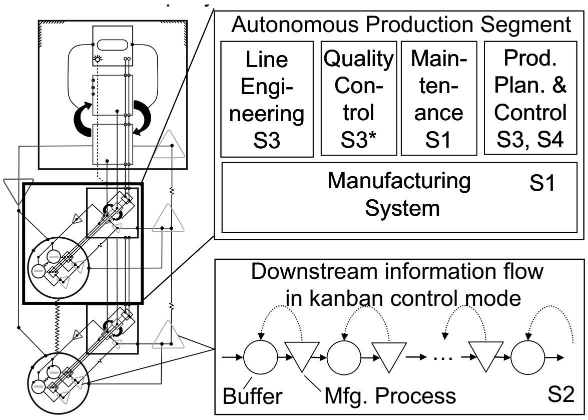


Figure 5: Decentralized production system segmentation.

Autonomous worker groups, decentralized maintenance and quality control processes are important lean methods that allow a segmentation of production systems according to the VSM. The design of information and material flows between production segments according to the Kanban method facilitate autonomous production segment and minimizes the variety for production planning, scheduling, and balancing.

5.3 Mode of Action of Lean Methods

The concept of variety amplification and attenuation can be applied to lean methods. Based on the structure of the VSM, the location as well as the mode of action of lean methods can be depicted. Table 2 depicts the mode of action of selected methods in terms of attenuating and amplifying variety. The Table 2 shows a strong relation between lean methods and the attenuation and amplifying of variety in the production context. Thus, lean methods support the alignment of production system design measures that attenuate and amplify variety. Thereby, lean methods increase operational and managerial variety in a highly effectively way in order to cope with the increasing external complexity. Moreover, lean methods optimize the capacity of information channels and transducers to ensure the transmission and transduction of information relevant to variety selection. To identify possible implementation sequences for lean methods in practice, their potential benefit in terms of attenuating or amplifying variety needs to be assessed and evaluated. As variety in the production context depends on company individual characteristics, company individual implementation sequence are required.

Gharajedaghi underlines the close connection of lean methods with the organizational cybernetics: 'This game emerged slowly but effectively in Japan, when Ohno, chief engineer of Toyota, created the Lean Production System by applying system thinking in the biological context. Using cybernetic principles, he was able to lower the break-even point by an order of magnitude and elevated the competitive game to an incredible higher level' [21].

5.4 Operational Flexibility and Strategic Adaptation

Operational flexibility and strategic adaption remain key factors for the competitiveness of production companies. The VSM allows to explain the mode of action of the fundamental

Table 2: Variety attenuating and amplifying effects of lean methods.

Lean Method	Description	Variety attenuation		Variety amplification	
		Example	Mo-de	Example	Mo-de
Total Productive Mainten-ance	Regularly cleaning, inspection and repair of machines by qualified worker to maximize the overall equipment effectiveness.	Minimization of different, unpredictable, unwanted machine states to reduce the required planning and coordination variety of S2 due to more predictable performances.	AT1	Increased worker knowledge allows to maintain and repair machines, to cope flexible with new problems to predict machine performance and behavior. Actual information increases S3 variety.	AM1 AM2
Kanban	Supply of materials and products by the pull principle so that the material is supplied when and where it is needed.	Minimization of different, unpredictable, unwanted states of material stocks and process input speeds minimizes the required planning and coordination variety of S2 due to close-loop information channels between different S1.	AT1 AT2 AT3	Increased tolerance against unforeseen problems with the material flow or processes within Systems 1. Shortened and closed-loop information channels increase availability of real-time information for System 3* audits.	AM1 AM3
Standards	Harmonizes information and process flows by standardized processes and information representation	Minimization of unwanted states of information handling and flows minimize the required filtration, error, and correction handling variety of S1, S2 &S3.	AT1 AT2 AT3	Standardized information schemes and representation increases availability of real-time information access and audit for System 3* audits.	AM3
Kaizen	Kaizen comprises a continuous improvement paradigm and corresponding process.	Kaizen goes along with a systemic and long-term thinking of the whole organization and all processes and, thus, minimizes variety for S1 process, S2 coordination.	AT1	More detailed and actual information of the current situation in terms of better plans and methods in S3, and a better mental model in S4 increases their variety.	AM1 AM2

nature of operational flexibility and strategic adaptation by relating these capabilities to a specific structure and functions that need to be fulfilled in order to ensure both, flexibility and adaptation. To support the design of flexible and adaptable production systems, the recursive structure of the VSM and the concept of variety can be drawn on. Chapters 5.1 to 5.3 have shown that lean methods simultaneously attenuate (AT1-AT3) and amplify variety (AM1-AM3) in a manufacturing context in order to reduce the efforts of variety handling. Lean design methods foster the design of decentralized information channels and close-loop control and information loops. By this means, lean methods also provide prerequisites for operational flexibility and strategic adaptation. The recursive structure of the VSM exhibits that the overall flexibility and capability to adapt the company depends on the capability of each sub-system to possess requisite variety (Figure 6).

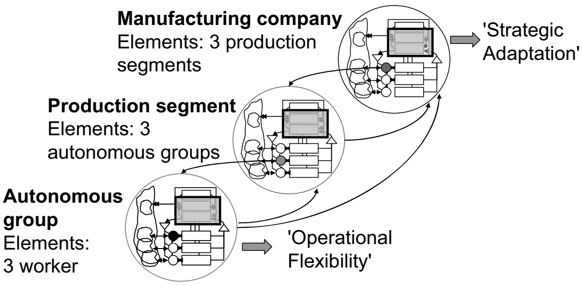


Figure 6: Operational flexibility and strategic adaptation.

While the autonomous mode of management allows flexible reaction of operational production units, standardized and closed-loop information control loops over different levels of recursions improve in the quality of the internal (mental) organizational model of the System 4. By that, strategic adaptations can be initiated rapidly and precisely by the opportunity analyzer according to the real organizational capabilities. That way, the capability of adaptation of viable system result in both, ‘operational flexibility’ at lower recursions and the capability of ‘strategic adaptation’ of the whole organizational on higher level of recursions.

6 CONCLUSION

This paper has shown the close relation between lean methods and cybernetic principles. By using the VSM as a proved organizational model, the mode of action of lean methods can be depicted by thinking of production system design problems in terms of variety attenuation and amplification. Moreover, decision support for the structural design and the segmentation of production systems according to the principles of cohesion and autonomy can be provided by the VSM. The insights of the recursive structure indicates that lean resp. cybernetic principles can be applied on all levels of recursions in order to improve operational flexibility and the capability of strategic adaptation. Thus, the VSM provides decision support on the implementation and development of LPS.

7 REFERENCES

[1] Ohno, T., 2006, Toyota production system. Beyond large-scale production, Productivity Press, New York.

[2] Spath, D., 2003, Ganzheitlich produzieren. Innovative Organisation und Führung, LOG_X Verlag, Stuttgart.

[3] Wildemann, H., Baumgärtner, G., 2006, Suche nach dem eigenen Weg: Individuelle Einführungskonzepte schlanker Produktionssysteme, In: Zeitschrift für wirtschaftlichen Fabrikbetrieb, 101/10.

[4] Liker, J. K., 2004, The Toyota way. 14 management principles from the world's greatest manufacturer, New York, NY, McGraw-Hill.

[5] Spear, S. J., 2004, Learning to lead at Toyota, In: Harvard Business Review 82/5:78-86.

[6] Bauer, J., Birdi, K., Denyer, D., Leseure, M., Neely, A., 2003, Adoption of Promising Practice: A Systematic Review of the Evidence. Report for the Department of Trade and Industry Innovation Review, Advanced Institute for Management (AIM)., London.

[7] Becker, M., Korge, A., Scholtz, O., 2002, Ganzheitliche Produktionssysteme - Erhebung zur Verbreitung und zum Forschungsbedarf.

[8] Porter, M. E., Ketels, C. H., 2003, UK Competitiveness: moving to the next stage, Department of Trade and Industry, London.

[9] Womack, J., 2005, Leading The Change to Lean, Aachen.

[10] Backlund, A., 2002, The concept of complexity in organisations and information systems, Kybernetes:30-43.

[11] Ashby, W. R., 1964, An Introduction to Cybernetics, Methuen, London.

[12] Beer, S., 1981, Brain of the Firm, Wiley, Chichester.

[13] Jackson, M. C., 2000, Systems approaches to management, Kluwer Academic/Plenum Publishers, New York.

[14] Beer, S., 1979, The Heart of Enterprise, Wiley, Chichester.

[15] Malik, F., 2006, Strategie des Managements komplexer Systeme, Haupt, Bern.

[16] Waelchli, F., 1989, The VSM and Ashby's Law as illuminants of historical management thought. In: Espejo, R., Harnden, R. (eds.), The viable system model. Interpretations and applications of Stafford Beer's VSM, Wiley, Chichester:51-75.

[17] Kanter, R. M., 1983, The change masters. Innovation for productivity in the American corporation, Simon and Schuster, New York.

[18] Gomez, P., 1978, Die kybernetische Gestaltung des Operations Managements. Eine Systemmethodik zur Entwicklung anpassungsfähiger Organisations-strukturen, Haupt, Bern, Stuttgart.

[19] Eversheim, W., Kubin, G., Munz, M., Pelzer, W., Schweitzer, G., 1997, Ganzheitliche Unternehmensreorganisation - Ein erfolgreicher Weg zur Standortsicherung, In: VDI-Z Integrierte Produktion/7/8:18-25.

[20] Warnecke, H. J., 1992, Die fraktale Fabrik, Revolution der Unternehmenskultur, Springer, Berlin u.a.

[21] Gharajedaghi, J., 2006, Systems thinking. Managing chaos and complexitya plattform for designing business architecture, Elsevier Butterworth Heinemann, Amsterdam.

Implementation Strategy for a Flexibility Oriented Production System

Felipe A. Gómez Kempf[1], Jan C. Aurich[1] and Olaf Drews[1]

[1]Institute for Manufacturing Technology and Production Systems, University of Kaiserslautern, Germany

Abstract

The continuously growing global industrial environment urges small and medium sized enterprises (SMEs) to work on implementing company specific production systems. However, flexibility requirements in particular have not yet been taken into account by existing approaches which deal with developing and implementing production systems. The presented implementation strategy provides manufacturing SMEs a step by step proceeding to achieve enterprise specific flexibility goals. The approach described in this paper has been already developed and validated in collaboration with four manufacturing enterprises. Significant results of implementing the flexibility oriented production system in these SMEs are described in order to verify the feasible applicability of this approach.

Keywords:
Flexibility; Production System; Lean Production; SME

1 INTRODUCTION

Manufacturing enterprises are continuously developing and improving their production systems in order to be able to face today's highly unpredictable and competitive market environment [1, 2]. While many large sized enterprises succeed to implement lean production systems [3, 4, 5], many small and medium sized enterprises (SMEs) have not yet been able to do so [6, 7]. Previous experiences show that copying or simply taking over existing production systems is not a feasible solution for any enterprise, especially for SMEs due to their structural differences compared to large sized enterprises. Existing production systems and therefore developed implementation strategies do not fulfill the specific flexibility requirements SMEs have to cope with. SMEs, which mostly act as suppliers of larger enterprises, are obliged to shorten processing times, and at the same time they must offer a lot of different product variants, which usually have to be manufactured in small batch sizes. The resulting highly order related production structures must continuously be developed and improved in order to guarantee an efficient and flexible costumer related order processing.

Since SMEs cannot pass on to their suppliers the flexibility requirements of their costumers, the approach of a flexibility oriented production system has been developed [8, 9, 10]. This paper focuses mainly on the implementation strategy SMEs have to follow in order to implement successfully the flexibility oriented production system. Since providing time input and financial resources often constitutes a critical bottleneck for SMEs, the presented implementation strategy includes a continuous questioning of the profitability of the taken measures.

2 FLEXIBILITY ORIENTED PRODUCTION SYSTEM

Four interrelated levels (prerequisites, elements, category groups and methods) constitute the basic framework of the flexibility oriented production system (Figure 1). While the prerequisites ensure the goal oriented implementation, the elements (Flexibility Achievement, Goal Definition, Standards of Methods and Evaluation of Qualification) ensure the flexibility orientation of the approach. In this paper, flexibility is defined as the ability of an enterprise to adapt itself to changing costumer requirements. Modular and combinable category groups, associated to enterprise specific methods, round off this production system.

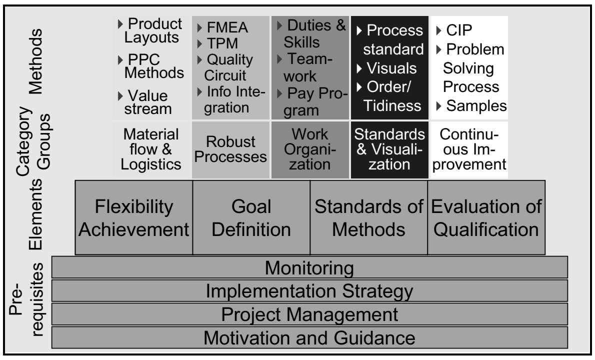


Figure 1: Approach overview.

This standard approach has to be adapted by each SME individually. The step by step procedure SMEs have to follow in order to implement an individually configured flexibility oriented production system is described below. In order to keep the adaptation effort low, the configuration mainly deals with choosing and combining some of the methods included in this approach. The approach accumulates a choice of SME valid methods and tools (e.g. appropriate material flow layouts, visually supported work instructions, as well as clean and safe workplaces). The choice of methods has been influenced mainly by their simplicity of implemention and usage in SMEs. However, upgrading and enhancing, as well as downgrading the production system by adding or excluding methods (if the flexibility goals require it) can be done individually by the respective SME if necessary. This revision and reconfiguration process must be carried out in regular time intervals.

3 IMPLEMENTATION STRATEGY

The implementation strategy consists of five steps, which in turn are divided in three phases (Figure 2): design phase, introduction phase and utilization phase.

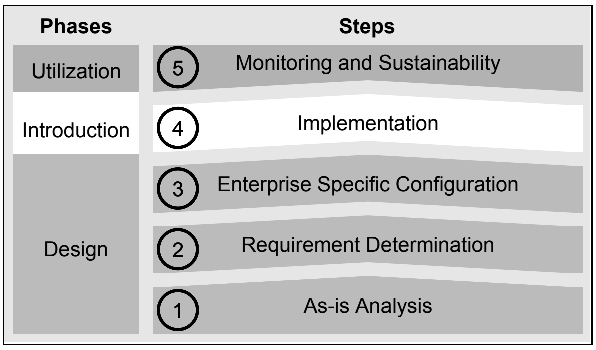


Figure 2: Implementation Strategy.

3.1 Design phase

This first phase comprises the following three steps: As-is analysis, requirement determination and an enterprise specific configuration.

As-is analysis

This is the first step to implement the flexibility oriented production system. Within this analysis the essential requirements for a successful implementation are defined. For this purpose the project team responsible for the implementation has to be precisely built. A communication concept has to be defined in order to clarify the implementation to all the workers. Furthermore, the enterprise specific particularities about the methods used for the order processing, product design and costumer-/provider relations, which are essential for the accomplishment of the subsequent implementation steps, have to be verified.

As a result of the as-is analysis, the need for action regarding the enterprise specific flexibility goals is defined. Additionally, starting points for the production system and the related improvement potential are determined. Deficits as well as inflexibilities (inability to adapt to changing costumer requirements) along the order processing structures are determined. The main aims, which are pursued in the as-is analysis are the following:

- Definition of the flexibility goals.
- Identification of potentials and waste on the shop floor.
- Analysis of the used technological and organizational methods.

For deriving the operative flexibility goals from the current long and middle term goals of the enterprise, a so called flexibility transformer has been developed. At this stage, a first evaluation of the corresponding flexibility types is carried out. Specific flexibility score cards are therefor systematically used disclosing their current status, as well as their upper and lower limits (Figure 3).

When identifying potentials and waste types, the following characteristics of the enterprise must be taken particularly into account: the order processing and the products manufactured, the demand rate and the manufacturing type, the factory layout, the machinery and the inventory stocks, the number of workers, and the information paths.

By evaluating the deficits and waste types, the potentials along the order processing become apparent. As a result, the main focuses of the production system can be identified subsequently.

Through the analysis of the used methods, the extent of existing parts of a manufacturing system are examined. The existing manufacturing concepts are evaluated afterward, focusing mainly on flexibility and modularity. Deficits in the use of methods, as well as the workers' competencies must be identified and documented.

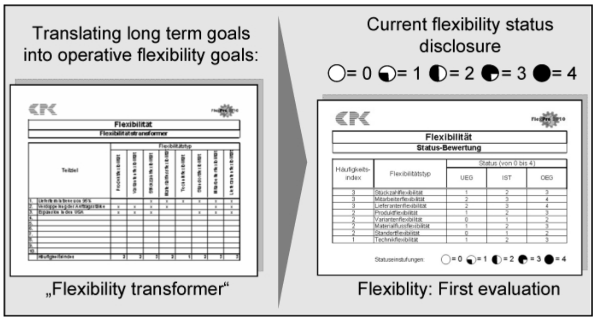


Figure 3: Deriving of the operative "flexibility goals"

Requirement determination

In this step the identification of the enterprise specific requirements follows. It is based on the previously identified deficits (for example regarding the application of organizational methods and tools in the shop floor). A prioritization of the potentials defined in the as-is analysis is carried out. This is necessary for the weighting taking place in the next step, in which the main category groups of the enterprise specific production system are established. The main actions which must be taken are listed below:

- Weighting and categorizing the potentials of the as-is analysis.
- Definition of the main focuses of the flexibility oriented production system.
- Evaluation of qualification.

The deficits detected in the analysis must be first categorized in a gross mode, before a concrete categorization of the ongoing potentials can be completed. These categorized groups of potentials and deficits have to be weighted thereafter. This analysis, categorizing and weighting of the shop floor potentials must be carried out on a regular basis (every 12 or 18 months for example). Based on this weighting, the main focuses of the initial enterprise specific production system can be derived. For SMEs it is recommended to start with three category groups (Figure 1) of methods as main focuses of the production system. This definition of main focuses can be proposed by the project team, but must as well be adopted by the workers and the management board.

Furthermore, a systematic determination of workers' competence requirements has to be fulfilled. Task related qualification measures must be therefor developed and applied. A three phase approach is recommended: specification, analysis and documentation. In the specification phase the tasks of the workers in the main focuses of the production system are concretized. The following analysis phase, breaks down the category groups into worker tasks and worker competencies, determining the actual way of solving concrete worker activities. Furthermore, the links between workers' tasks and competencies with the elements of the production system are considered. For this purpose a simplified (SME oriented) competence rating matrix has been developed. In this matrix, the methods are opposed to the tasks that are to be solved. Hereby the worker or teams can be evaluated in a scale between 0 and 4. The documentation phase ensures the possibility of further actions.

Enterprise specific configuration

This step comprises the procedure to adapt the general flexibility oriented production system to an enterprise specific one, by adapting it to the specific flexibility needs of the respective enterprise. The first (as-is analysis) and the second step (requirement determination) of the implementation strategy provide the input needed for the configuration.

At this stage the following procedures must be fulfilled:

- Determination of the method combination.
- Primary visualization of flexibility operating figures.

On the basis of the earlier determined main focuses of the production system, as well as the flexibility requirements of the enterprise, methods for the implementation of these flexibility features must be chosen. By means of standards of methods (especially developed for this production system) and a so called "method compass", which provides a clear method overview for the operators, methods can be chosen and implemented, depending on their respective flexibility type. The determination of the method combination should be elaborated and presented to everyone involved in the implementation.

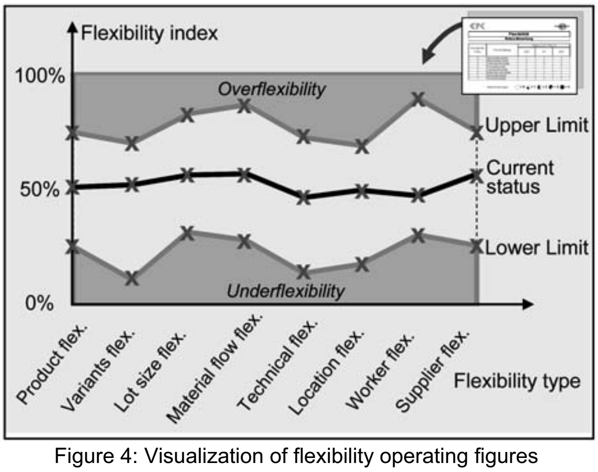


Figure 4: Visualization of flexibility operating figures

The flexibility achievement results from the developed approach of flexibility visualization (Figure 4). Eight flexibility types are considered: product, variants, lot size, material flow, technology, location, worker and supplier flexibility. The values of the relevant flexibility types, which are identified by the project team in collaboration with the division managers, are visualized at first. The upper and lower flexibility limits are visualized next, using the specially developed flexibility score cards. Each flexibility type has its specific flexibility score card, in which predefined questions help the project team to determine the actual flexibility operating figures. The visualization is implemented by the aid of a visualization control chart. A quarterly updating of the flexibility operating figures is recommended.

3.2 Introduction phase

This phase consists of the implementation step. All of the previously defined measures have to be realized at this point. At the beginning, gross measures have to be completed, in order to set the basis for further detailed actions. In addition, a structured roadmap for the implementation must be defined. The project team, the enterprise's middle management and the workers must define concrete measures, documented in standardized work packages. Everyone in the plant must understand and take responsibility for what has to be done. The overall overview of the ongoing work in progress of the implementation should be visualized. Weekly reports or statements regarding the status of the work packages should be centrally evaluated by the project team. Useful and easy to employ project management methods have been documented in a simplified manner. Moreover, some hints and tools for non experts have been summarized. Nonetheless, the project team should be familiar with project management basics in order to be able to initiate a proper implementation.

Basically, the following three activities bundle the actions which must be taken in this step:

- Definition of a project master plan.
- Determination of concrete measure descriptions.
- Step by Step implementation by example pilot projects.

The project master plan includes concrete measure descriptions for especially chosen pilot projects. Costs, due dates and the involved resources are planed in it. When determining the concrete measure descriptions, the project team must consider the possibility of goal conflicts between the various measures defined. The step by step implementation is divided as follows: Pilot projects, Roll out, and Continuous improvement. At first, small divisions of the enterprise develop pilot projects in applying the defined measures. When first positive results arise (Quick Wins), which helps to motivate the workers, the rest of the divisions of the enterprise follow with the implementation. Once implemented, a regular continuous improvement of the production system must be done. For this purpose concrete workers have to be chosen and provided with the necessary resources. It is very important, that the management board of the SME firmly and publicly supports the continuous improvement.

3.3 Utilisation phase

Monitoring and Sustainability are the last step of the implementation strategy. This phase is characterized by keeping the production system alive. Monitoring regularly the ongoing activities is one of the main tasks of the project team at this stage, as well as making sure, that the production system is "kept alive" by the shop floor operators. Therefore, operators must be involved in the continuous improvement process of the production system. They must understand why change is necessary and they are urged to give some input of how improvements could be realized. The project team must spread the success of the implementation and reward and hence motivate the operators. This is how the production system is to be kept alive in the enterprise and how it can be continuously upgraded. Quality and profitability have to be improved constantly, in order to achieve the SMEs' long term success.

The project team must also monitor and control costs, due dates, quality and the ongoing implementation progress on a regular basis, quarterly for example (Figure 5). If divergences or even irregularities regarding some tolerance levels (e.g. Flexibility Limits) are noticed, concrete measures must be defined and taken. These measures vary according to suitability. Either the tolerance levels must be readjusted (though only in exceptional cases), or measures must be defined to get back to the predefined tolerance interval. Another action, which must be taken at this stage, is to ensure that everyone involved in the implementation is provided with the necessary information regarding the project status and its current measures.

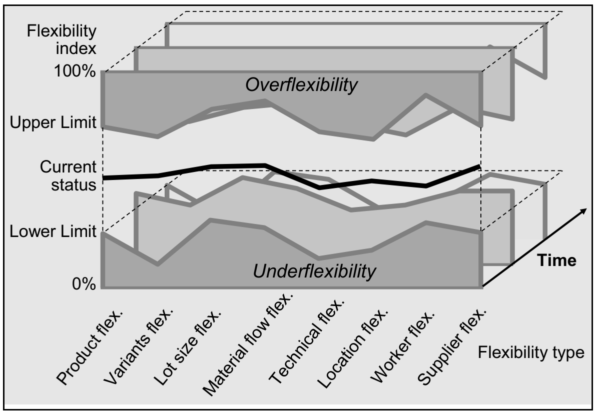


Figure 5: Periodical visualization of the flexibility types

Moreover, the sustainability is only given if following notions are consistently considered and applied firmly:

- Every process has to be designed aiming zero defects.
- Continuous improvement requires periodical modifications of the production system.
- The production system can not be "kept alive" in an SME, if the shop floor operators do not actively support it.
- Any production system (the flexibility oriented production system included) is just a means to an end, it is not an end itself.

4 VALIDATION

The implementation strategy presented in this paper takes into account the particularities of SMEs. The step by step procedure was developed paying attention to simplicity and clarity. The flexibility oriented production system and its implementation strategy have been validated in a joint research project involving four manufacturing SMEs. In each participating SME, the approach has been adapted to the specific flexibility needs of the enterprise. Different emphasis has been put respectively on the main category groups of the flexibility oriented production system. To name one example, one of the manufacturing SMEs (a tool manufacturing enterprise) involved in the joint research project, defined the category groups Robust Processes, Work Organization and Standardization & Visual Management as the main focuses of the enterprise specific production system. However, the applied implementation strategy has been very similar throughout the consortium.

Some of the measurable achievements among the participating enterprises have been:

- Reduction of stock of inventory.
- Achievement of higher machine availabilities.
- Decrease in throughput times.

5 SUMMARY

Taking into account the specific flexibility needs of SMEs a validated implementation strategy (consisting of five steps) for a flexibility oriented production system has been described. The procedures presented have been crosschecked regarding simplicity, clearness and usability in order to enable SMEs' workers to implement and continuously improve the production system themselves. Finally, a few examples of successful achievements due to applying the implementation strategy have been described.

6 ACKNOWLEDGEMENTS

Sincere thanks to all partners in the joint research project FlexPro 2010, who have contributed to the implementation strategy described in this paper. FlexPro 2010 is funded by the Ministry of Education, Science, Youth und Culture Rhineland Palatinate under financial support by the European Fund for Regional Development (EFRD).

7 REFERENCES

[1] Black, J. T., 2007, Design rules for implementing the Toyota Production System, International Journal of Production Research, 45,16: 3639-3664.

[2] Yasuhiro Monden, 1998, Toyota Production System: An Integrated Approach to Just-In-Time, Third edition, Norcross, GA: Engineering & Management Press.

[3] Ohno, Taiichi, 1995, Toyota Production System: Beyond Large-scale Production, Productivity Press Inc.

[4] Womack, J. P., Jones, D. T., Roos, D., 1991, The Machine That Changed the World: The Story of Lean Production, HarperBusiness.

[5] Shingo, Shigeo, 1989, A Study of the Toyota Production System from an Industrial Engineering Viewpoint, Productivity Press.

[6] Herrmann, C., Bergmann, L., Crespo, I., Development of an Integrated Modernization Concept for SME, 2006, 39th CIRP International Seminar on Manufacturing Systems, Ljubljana, 457-463.

[7] Spath, D., ed., 2003, Ganzheitlich produzieren: Innovative Organisation und Führung, Log_X.

[8] Aurich, J. C., Drews, O., Fuchs, C., 2006, Produktionssysteme für den Mittelstand, wt Werkstattstechnik online, 96, 302-307.

[9] Aurich J. C., Gómez Kempf, F. A., Drews, O., Wagenknecht, C., 2006, Flexibilitätsorientiertes Produktionssystem für den Mittelstand: Ein Konzept zur Gestaltung und Einführung, ZWF, 12, 709-714.

[10] Aurich, J. C., Gómez Kempf F. A., Drews, O., Wagenknecht, C., 2007, Flexibility Oriented Production Systems - Concept for SMEs. 40th CIRP International Seminar on Manufacturing Systems, Liverpool.

A Study of Design Factors for Information Supporting System in Cell Production

Ye Zhang[1], Feng Duan[1], Jeffrey Too Chuan Tan[1], Kei Watanabe[1],
Nuttapol Pongthanya[2], Masao Sugi[1], Hiroshi Yokoi[1], Tamio Arai[1]
[1] Department of Precision Engineering, School of Engineering, The University of Tokyo, Japan
[2] Department of Systems Innovation, Faculty of Engineering, The University of Tokyo, Japan

Abstract
With the innovation of mechanism, operators are released from the physical assembly tasks. Most of current assembly tasks are the cognitive tasks; therefore providing assembly supporting information can improve the assembly process. However, which kinds of assembly information fit the operators and how to provide the assembly information is convenient to the operator have not been investigated thoroughly. To solve these problems, an information supporting system is developed and several operators are employed to execute a cable-insert task. The results show that providing multimedia based assembly information in the assembly area can reduce assembly mistakes and mental work load while accelerate the assembly process.

Keywords:
Assembly Performance; Assembly Information; Instruction Format; Mental Work Load

1 INTRODUCTION

Traditional line manufacturing systems cannot satisfy the current customers' changing tastes effectively. Therefore, cell production system is proposed. In the cell production, the assembly tasks are done by a human operator from start to finish [1]. Since the operator plays a significant role in the cell production system, the productivity mainly depends on the operator's performance. However, human operators, especially the beginners, often make assembly mistakes. Moreover, due to the aging problem in Japan, it is difficult to maintain the number of the experienced human operators. As a result, without effective innovations, cell production is difficult to be maintained.

With the rapid development of computer science, it is possible to provide multimedia based assembly information to the human operators during the assembly process. Many researchers utilized display or projector to provide information to guide the human operators during the assembly process. Ou [2] employed Augmented Reality to support human operators by assembly instructions, which improved the human operator's assembly performance. Nitta [3] proposed a Head Mounted Display (HMD) to provide information to the human operators in a nuclear power plant. Based on HMD, human operators can easily find out the work place and know the assembly steps conveniently. Although these researches successfully support the human operator in information aspect, however, all these researches mainly focused on how to design an information supporting system.

Which kinds of assembly information fit the human operators and how to provide the assembly information effectively to the human operators have not been investigated thoroughly. To solve these problems, based on Norman's [5] cognition science and Ishii's [6] visual scope theory, the assembly information is record in the PowerPoint Slides and provided to the operators through LCD TV. A group of operators are employed to execute a cable-insert task under the assembly information guidance to find out the factors that influent the assembly performance.

The rest paper is organized as follows. The current assembly task and its characteristics are introduced in section 2. The information supporting materials and the methods are explained in section 3. A sample assembly task is proposed in section 4 to investigate the effects of different assembly information formats and the supporting modes. Finally, the conclusions and future work are given in section 5.

2 ASSEMBLY TASK IN CELL PRODUCTION

2.1 Assembly system

In the assembly system, a human operator works in the designated place. The assembly information, such as assembly flow chart or assembly part's picture, is listed in the manual. During the assembly process, the human operator executes the assembly task based on the manual's requirement. The diagram of the assembly system is shown on Figure 1.

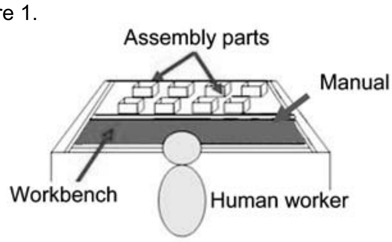


Figure 1: Diagram of the assembly system.

2.2 Characteristics of the assembly tasks

In cell production, the assembly tasks are not the physical tasks that the human operators need to lift or carry heavy parts. Most of assembly tasks are the cognitive tasks; therefore, the human operators need to find the parts, remember the information from the instruction, and install the parts into the correct assembly positions.

In order to understand the difference between the skilled operators and the unskilled operators, Fujino [4] analyzed their assembly performance in the real assembly tasks. The results show that the differences between the skilled operators and unskilled operators are that the skilled operators know "What the most efficient assembling order is", "Where the assembly parts are" and "How to assemble the parts".

According to Fujino's research, during an assembly process, the human operator should understand the following items listed in Table 1. Which kinds of assembly information fit the human operators and how to provide the assembly information effectively to the human operators are the two important factors that influence the human operators' assembly performance.

Table1 Points that a new operator should understand

Items	Content
Assembly Introduction	What is the next assembly step? e.g. Insert the cable
Assembly Instruction	The content of assembly task e.g. Pull out cable 1 from the cable kit
Characteristics of Parts	The cable's color, length, etc.

3 ASSEMBLY INFORMATION

3.1 Cognitive engineering

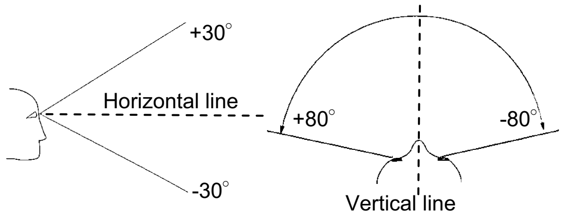
Norman [5] pointed out that humans have two types of cognition: experimental cognition and reflective cognition to deal with the information from the outside world. Experimental cognition means that human beings can recognize and respond to the surrounding things efficiently without much effort. While in the reflective cognition, human beings need to compare and consider before they make the decision. In the reflective cognition, new ideas and new actions are generated. Different to the experimental cognition, the reflective cognition induces the mental burden to the human beings.

Both the experimental cognition and the reflective cognition are necessary to humans. However, as far as assembly task is concerned, considering the assembly efficiency, it would be better to let the human operators obtain assembly information by experimental cognition as much as possible. Therefore, the assembly information formats, such as keywords, pictures and flash marks, should be used rather than abstract manual.

3.2 Visual scope

Besides providing the assembly information to the human beings, how to provide information to the human beings is another important research topic. Ishii [6] pointed out that provide information to the human beings within their visual scope is an effective way to accelerate the cognition process. Visual scope is briefly illustrated in Figure 2. There are two kinds of visual scopes: vertical scope (±30°) and horizontal scope (±80), shown in Figure 2 (a) and (b) respectively.

During the assembly process, the assembly information should be provided in the human visual scope as much as possible. This will reduce the eyes' movement and accelerate the understanding process. Meanwhile, the assembly information of the related assembly steps should be provided to the human operator in the same area.



(a) Vertical scope (b) Horizontal scope

Figure 2: Visual scope.

3.3 Assembly information format

With the development of computer science, besides the traditional paper based assembly manual, multimedia based assembly information can be provided to the human operators. In this study, the utilized information formats are analyzed in the following.

Keyword

Providing too much information at the same time increases the operators' burden; therefore, unnecessary information, such as tool's specification, should be ignored. The significant assembly information (e.g. assembly step) is written in keyword. During the assembly process, the keyword to explain each assembly step can be provided to the human operator through a projector or LCD TV.

Picture

According to cognitive point, recognizing the assembly parts' pictures is an experimental cognition process. The vivid appearance of assembly parts can be presented to the human operator by their pictures. The pictures should be taken from the operators' line of sight during the assembly process. This can let the operators understand the information easily.

Emphasis indication

According to Fujino's opinion [4], "Where the assembly parts are" influences the assembly performance. Using emphasis indication method, such as flash marks, to point out the assembly parts or the assembly position is an effective way to improve the assembly performance.

Voice

Assembly information can be saved in to audio files and provide to the human operators during the assembly process. This may accelerate the assembly process, since the human operators can focus on the assembly task without unnecessary motions, such as using hands to change the manual's page.

Video

To let the unskilled human operator understand the assembly process easily, videos can be used to provide assembly details vividly. Duan [7] record and extracted the advantages of the human operators' assembly motions, and then employed the synthesized assembly video to train new human operators to improve their assembly performance.

4 EXPERIMENTATION

4.1 Experimental setting

Although the mentioned assembly information formats can guide the human operators to execute the assembly tasks; however, which kinds of formats fit the human beings has not been investigated thoroughly. To investigate the effects of different instruction formats and the information supporting way, a sample experiment is set up, shown in Figure 3.

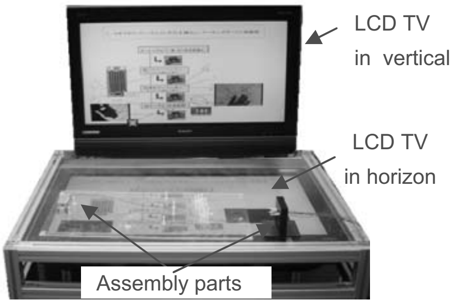


Figure 3: Structure of the information supporting system.

In this experiment, a LCD TV is employed to provide assembly information to the human operators. To find out the effective direction to provide the assembly information to the human operator, the LCD TV is laid in the front of the workbench and laid at the top of the workbench respectively, shown in Figure 3.

To investigate which kind of information format fits the human operator and how to provide the assembly information is effective, a cable-insert task is employed, illustrated in Figure 4. In the cable-insert task, the human operator must insert the correct cable into the correct hole in the connector. After that, following the cable route, the operator should fix the cable at the workbench. All of the assembly information, including the cable, the connector's hole and the assembly steps, is provided by this information supporting system.

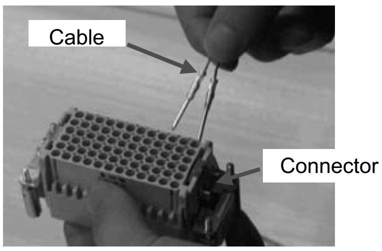


Figure 4: A cable-insert task.

The whole assembly scheme is divided into several simple assembly steps, and the corresponding assembly information is written in PowerPoint slides, illustrated in Figure 5. In each PowerPoint slide, the assembly parts and assembly tools are illustrated by pictures; the assembly positions are pointed out by colourful flash marks; following the assembly flow chart, the videos of the experienced operators' assembly motions will appear to guide the novices to execute the assembly tasks. To make the human operator understand the assembly process easily, the colours of the words in the slides are same with the real assembly parts' colours, such as "blue cable", "grey cable" in Figure 5. During the assembly process, these PowerPoint slides are output into the LCD TV and switched by the footswitch in assembly sequence.

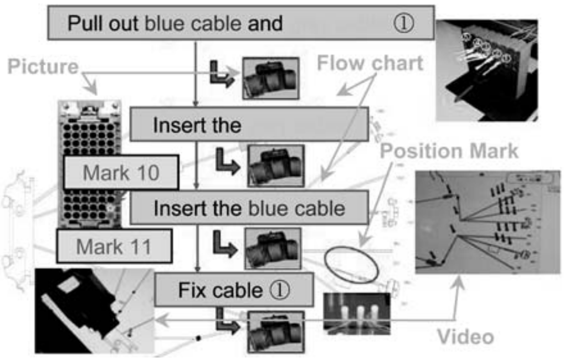


Figure 5: Assembly information of one step.

4.2 Experimental results

A group of healthy males (with ages ranging from 24 to 29 years) are asked to execute the cable-insert task under this information supporting system. The experimental results are shown in Figure 6 to Figure 8. In these figures, "A" means the results of the human operators executing the cable-insert task only based on the assembly manual. "B" and "C" mean the results of the human operators executing the same assembly task under the information supporting system. In "B", the LCD TV is set in the front of the human operators vertically, shown in Figure3. In "C", the LCD TV is laid at the top of the assembly workbench.

During the experiment, the average eye move numbers of the human operators under different assembly information supporting methods are counted, shown in Figure 6. According to Figure 6, in "A" the average eye move number of the human operators is 25 times; in "B" the average eye move number increases to 29 times; in "C" the average eye move number reduces to 15 times. Using the assembly information system may reduce the eye movement; however, the assembly information providing direction is very important, which influence the effect of information supporting system. Based on assembly manual (in "A"), operators can remember several assembly steps; while in "B", operators have to move their eyes to read the information in the vertical TV. Hence, the eye movement number in "B" is the most. According to Ishii's opinion, the information should be provided within the visual scope as much as possible. In this cable-insert task, the operators should insert the cable at the top of the workbench; therefore, providing assembly information by the LCD TV laid at the top of workbench will reduce the eye movement.

The average assembly time of the human operators under different assembly information supporting methods is shown in Figure 7. In "A", the average work time is 213 (s). In "B", the average work time increases 9.4% than that of "A". The reason is that the human operators have to move their eyes to see both the LCD TV and the assembly parts, which slows down their assembly process. In "C", the average work time reduces 16.0% than that of "A". From these results, we can see that providing the assembly information in the working areas can accelerate the assembly process.

The assembly error rate is listed in Table 2. According to Table 2, only depending on the assembly manual, human operators make assembly mistakes. Fortunately, with the assembly information supporting system, the error rate is 0.

Table 2 Assembly error rate

Items	A	B	C
Error rate	3%	0%	0%

Although supporting operators by assembly information can improve the assembly performance, the relationship between the information format and human mental fatigue should be investigated. To evaluate the mental load of different information format, NASA-TLX (Task Load index) method [8] is used. After the operators finished the cable-insert task, they should give their answers to the questionnaires. Based on the NASA-TLX method, the mental work load can be computed. The results are shown in Figure 8. According to Figure 8, the work load of “A” is the highest (62); the work load of “B” and that of “C” reduce to 57 and 38 respectively. This means that based on the multimedia assembly information and providing the assembly information in the working area can reduce the human operators’ mental load.

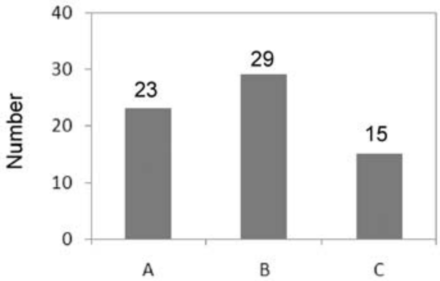


Figure 6: Average eye move number.

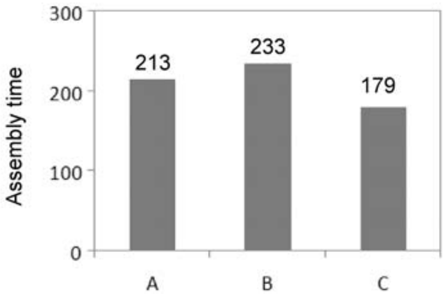


Figure 7: Average assembly time.

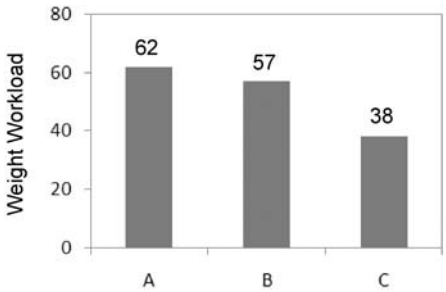


Figure 8: Results of NASA-TLX.

5 CONCULSION AND FUTURE WORK

In cell production, current assembly tasks are mainly cognitive tasks. Providing assembly information can effectively improve human operator’s assembly performance. To find out which kinds of assembly information fit the human operators and the effective information providing method, an assembly information supporting system was set up. Based on the Norman’s cognition science [5] and Ishii’s [6] visual scope theory, the assembly information was made and record in the PowerPoint files. A group of human operators were asked to execute a cable-insert task. The experimental results show that providing the assembly information to the human operator can reduce the human mental work load and accelerate the assembly process. Meanwhile, based on the assembly information supporting system, the assembly error rate reduces. Providing assembly information in the operator’s assembly area is the most effectively method to improve the assembly productivity and reduce the human mental work load.

In the future, the effects of different assembly information formats to the assembly process should be evaluated in quantity. The best way to mix the different information formats, such as text, picture, audio, video, etc., will be investigated.

6 ACKNOWLEDGMENTS

This work is part of “Strategic Development of Advanced Robotics Elemental Technologies” project supported by NEDO. Highest gratitude is expressed to NEDO and research partner − FANUC for the excellent cooperation and technical support.

7 REFERENCES

[1] Seki, S., 2003, One by One Production in the “Digital Yatai” − Practical Use of 3D-CAD Data in the Fabrication, Journal of the JSME, 106/1013:32-36.

[2] Ou.j.et al., 2004, DOVE: Drawing over Video Environment, Multimedia 2003 Demonstration.

[3] Nitta, K., 2003, Using Augmented Reality based Maintenance Work Support System for Nuclear Power Plant, Master Thesis.

[4] Fujino, H., Matsuzaki, T., Nitta, K. Ishii, H., Shimoda, H., Yoshikawa, H., Wei, W., Terashita, N., 2002, Construction of a Supporting System for Inserting Electronic Parts on Printed Circuit Board Using LASER and a Video Camera, HIS, 733-736.

[5] Norman, D. A., 1993, Things that Make Us Smart: Defending Human Attributes in The Age of The Machine, Addison Wesley Pub. Co. Technology/Philosophy.

[6] Ishii, H., Kobayashi, M., Brave, S., Dahlay, A., Gorbet, M., Ullmer, B., Yarin, P., 1998, AmbientROOM: Integrating Ambient Media with Architectural Space(video), in Conference Summary of CHI’98, ACM.

[7] Duan, F., TAN, J.-T.-C., Zhang, Y., Watanabe, K., Pongthanya, N., Sugi, M., Yokoi, H., Arai, T., 2007, Analyze Assembly Skills using a Motion Simulator, Proceedings of IEEE-ROBIO, 1428-1433.

[8] NASA-TLX for Windows U.S., Naval Research Laboratory, http://www.nrl.navy.mil/aic/ide/NASATLX.php.

On Reconfigurable Robotic Working Cells – a Case Study

Mikael Hedelind

School of Innovation, Design and Engineering, Mälardalen University, Sweden

Abstract

Today's global market and increased competition between companies necessitate the development of production systems that are agile and responsive to change. This paper presents the results of a project where a flexible and reconfigurable robotic working cell was developed. There has been a strong focus on ease-of-use, reconfigurability and visual management throughout the project. The resulting robotic working cell has been evaluated as a prototype in a laboratory and developed into a production module; used in production at a manufacturing plant in Sweden. A discussion around the concept of reconfigurable robotic working cells will be presented, together with relevant theory.

Keywords:

Industrial Robotics, Reconfigurability, Flexibility

1 INTRODUCTION

The business environment for manufacturing industries has become more and more intense and difficult as globalization increases; and as a result toughens competition from international competitors. Manufacturing industries has evolved towards becoming more flexible in their product portfolio offering more customized products and at the same time trying to lower their prices while sustaining their product quality. One way of becoming more productive and keeping consistent quality is using automation within the production process.

For quite some time, automation has been considered one way of becoming more competitive by having a more efficient production process, with less labour cost included in the final product. However, to meet with global competition the companies also need a very flexible production system; developing manufacturing systems that can handle small lot sizes and being able to quickly implement changes to become even more efficient or meet with changes in the product design. Flexibility has been a topic within research for quite some time, and the need for flexible systems has been established by several authors, see e.g. [1, 2]. However, the concept of reconfigurability is not as deeply covered within literature. As well, there is a need for more practical application of the two concepts within industry in order to evaluate and to continue the development of the two concepts.

The objective of this paper is to present the results of an industrial case study where the concept of reconfigurable robotic working cells are investigated and realized. The case study was performed as part of a research project where academic partners cooperated with industrial partners in order to solve an industrial problem where an ergonomically unsound manual workstation was to be automated. The solution was a robotic working cell with strong requirements regarding flexibility and reconfigurability. The robotic working cell was developed as a prototype in a laboratory setting and later introduced into real production.

1.1 Methodology

This research has been conducted in the form of a case study, where the researchers have actively taken part in a development project in close cooperation with industry. In general, case study method is the preferred strategy when 'how' or 'why' questions are being posed [3]. It is also useful when the investigator has little control over events, or when the focus is on a contemporary phenomenon within some real-life context, which is relevant for this work.

During this research, a literature review has been conducted and relevant literature have been identified and analyzed. During the project at the industrial partner, direct observation, interviews, and document reviews have been conducted. Throughout the project a lot of experience have been gathered and used as a basis for reasoning and problematising about the problem and solution.

This research has been conducted as a longitudinal case study at the industrial partner, together with employees from that company. The researcher has taken an active part in the development project and to some extent managing the activities in the project.

2 THEORETICAL FRAMEWORK

Below, the theoretical framework for this research is presented.

2.1 Definitions of flexibility and reconfigurability

The definitions of flexibility and reconfigurability that are used in this research have been adopted from [4].

Flexibility is defined in this research as the ability to robustly handle short-term changes quickly and at a low cost *within an existing system*.

Reconfigurability on the other hand means the ability to *adapt the system* rapidly in response to changing needs and opportunities.

2.2 Flexibility within manufacturing systems

The concept of flexibility have been extensively researched and published over time, a study performed by Shewchuk and Moodie identified about 70 different definitions and measurements of flexibility described in litterature [5]. Flexibility is by many authors considered a necessity for coping with internal and external disturbances [6]. Flexibility can also be seen as a measurement for how many different product variants that a certain manufacturing system can handle; what e.g. Browne et al. call product flexibility [7]. Other similar types of flexibility can be [7]: process flexibility, production volume flexibility, and operation flexibility. The measurement of flexibility cannot be performed on an absolute scale, but must be seen as a relative measurement [8]. One area of research where the concept of flexibility has been investigated and developed is called Flexible Manufacturing Systems (FMS), see e.g. [7, 9, 10].

2.3 Reconfigurability within manufacturing systems

The concept of reconfigurability has been introduced within manufacturing systems in several different ways. There is ongoing research in the area of reconfigurable machines, where focus is on how to change machines to perform different operations. Another type of reconfiguration is on a production system level where reconfiguration is performed over a larger scale, e.g. logistical re-routing. The research presented here is focusing on re-configuration of robotic working cells. Reconfigurability as a concept has been identified and researched primarily within the area of Reconfigurable Manufacturing Systems (RMS), see e.g. [11, 12, 13]. Mehrabi, et al. define RMS as [11, p. 405-406]:

'*A reconfigurable manufacturing system is designed for rapid adjustment of production capacity and functionality, in response to new circumstances, by rearrangement or change of its components. [..] Components may be machines and conveyors for entire production systems, mechanisms for individual machines, new sensors, and new controller algorithms.* '

In contrast to the FMS, which is designed to have an over-capacity, the design of the RMS is to enable easy reconfiguration and responsiveness to change, using a modular structure for both hardware and software [12].

The European Robotics Research Network (EURON) state in their research roadmap that flexibility and agility is becoming increasingly important for advanced manufacturing systems, at the expense of e.g. efficiency [14]. The report derives this statement from the change in market where the customer will demand highly diversified product mixes at shorter delivery times, in smaller batches. Mehrabi et al. describe the future need for responsive and agile manufacturing systems as [11, p. 403]:

'The manufacturing systems used for this new approach must be rapidly designed, able to convert quickly to the production of new models, able to adjust capacity quickly, and able to integrate technology and to produce an increased variety of products in unpredictable quantities.'

2.4 Flexibility and reconfigurability within industrial robot working cells

Manufacturing systems consists of many types of cells or working stations. When looking at the definitions adopted for flexibility and reconfigurability one realizes that the industrial robot itself is very flexible, by definition (ISO Standard 8373:1994): '*An automatically controlled, reprogrammable, multipurpose, manipulator programmable in three or more axes, which may be either fixed in place or mobile for use in industrial automation applications*.'

However, it is when the robot is placed into a workstation or cell that it often becomes rigid and hard to adapt. Reconfigurability have been discussed to some extent within the area of automation, e.g. Canny and Goldberg concludes that rapid reconfigurability will be critical for the manufacturing systems and that it will be supported through work cells that are "easily edited" [15]. There is a need for robotic working cells which can be changed over time, as identified as a result of e.g. the MARK projects [16, p. 223]:

'The belief is that true long-term flexibility may only be achieved by leaving the robot solution open for changes. [..] Many simple, strictly task-oriented components with standard interfaces are better than few, very flexible but expensive solutions.'

It is commonly understood that the life-span of an industrial robot in most cases is longer than a single project or products life-span, which means that it is important to be able to re-use the robot over several projects or products, see e.g. [17].

2.5 Complexity in manufacturing systems

Manufacturing systems are often considered to be complex systems [18, 19]. Kuzgunkaya and ElMaraghy [19] points out that as companies strive to increase their range of products and implement reconfigurable manufacturing systems to achieve the agility needed to meet with the uncertainty of today's market; the manufacturing systems become more complex and difficult to manage. McCarthy and Rich [20] discuss different levels of automation and refer to them as *Levels of complexity*, meaning that the more automation used - the higher the complexity. They also discuss these levels of complexity and points out some of the difficulties that companies may experience if they implement manufacturing systems with to high level of automation and complexity. A system that is complex is described as having a static structure and dynamic behaviour which is unpredictable and counterintuitive [18]. One of the *costs of automation extremes* that Miller and Parasuraman [21] describes are reduced situation and system awareness, which in the case of manufacturing systems leads to operators which are out of the loop.

3 INDUSTRIAL CASE

As part of the *Factory-in-a-Box* research project the *Demonstrator 1* sub-project was carried out. For more information on Factory-in-a-Box see e.g. [22]. Demonstrator 1 was carried out as a co operational project where academic partners together with industrial partners worked to solve an industrial problem. The project started in the early 2005, and the first prototype was presented in laboratory environment in May 2006. The prototype showed promising results and an operational version of the cell was introduced into production in December 2006. During the summer of 2007 a first reconfiguration of the cell was carried out.

3.1 Industrial problem

The problem to be investigated at the industrial partner was a manufacturing problem that included the need for increased productivity through making some workstations more efficient and achieving a more consistent product- and process quality. The aim of the project was to investigate how automation could be used in order to achieve the above mentioned needs and to develop a working cell that could match the requirements on efficiency and quality.

The production process that was investigated included a set of workstations where metal sheet plates were assembled into a cabinet. Each cabinet consists of 4 to 6 plates that were joined using hemming, with glue applied in each joint. When done manually, the human workforce had to manually lift the cabinet into the hemming machine, which was ergonomically unsound. Also, the glue had to be applied manually, which is not an operation suitable for humans. The application of glue had to be perfect, as too little glue resulted in a cabinet which was not airtight, and too much glue resulted in waste glue in the hemming station and on the outside of the cabinets which in turn leads to rework and frequent maintenance.

3.2 Objective of Demonstrator 1

The company had identified that industrial robots was a suitable solution based on the size of the cabinets, and their ability to be reprogrammed. The company's production department did not have a lot of experience with automation and industrial robotics, so they sought to use a stepwise automation migration strategy where the level of automation was increased as the company felt comfortable with its current situation, see Figure 1.

The cabinet that was produced in the production line was built-to-order and highly customized. This resulted in a very high number of product variants, which all should be assembled in this working cell.

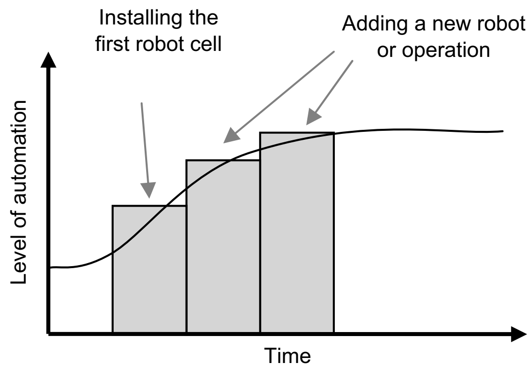


Figure 1: A stepwise automation migration strategy where the level of automation of the production process is increased over time.

3.3 Vision of Demonstrator 1

As the team started to work with the project, a vision of how the manufacturing system at the plant could look like started to emerge. The idea that was introduced was that the shop floor should be constructed as a jigsaw puzzle, where each piece of equipment and workstation should be designed as a piece of the puzzle with well defined interfaces between them. Figure 2 shows a rendered illustration of the vision of the modular manufacturing system developed in Demonstrator 1.

If one can move each piece of equipment, then the working cells may be reconfigured when needed. Realizing this vision should meet with the overall goal of the industrial partner, that they should be able to stepwise increase the number of robots in the working cell or the peripheral equipment and operations performed by the robots.

3.4 Development of Demonstrator 1

During the development project, several different technologies were investigated and developed in order to realize the vision of the modular manufacturing system. Technologies that enable rapid and easy introduction of robots or equipment into the working cell were identified, e.g. wireless I/O communication and automatic calibration of the robot and peripheral equipment, and some technologies were developed e.g. mobile platforms for the robots.

Figure 3 shows an industrial robot, an ABB IRB 6600, on a mobile platform. The mobile platform was designed as a hollow steel plate with a connection for compressed air. The compressed air can be used to fix the platform to the floor, as a large vacuum gripper. Also, the compressed air can be reversed and thus used as an air cushion for easy transport. This design of the mobile platform was realized, as shown in Figure 3; however the platform was only tested to a certain extent in the laboratory since the company did not whish to include this technology in the finalized version of the working cell.

The use of wireless I/O communication pads was very beneficial since it enabled easy changes of layout with little wiring.

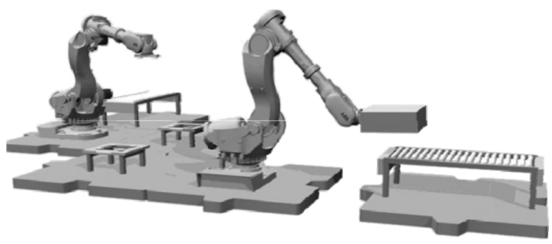


Figure 2: Robots placed on pieces of a jigsaw puzzle.

Figure 3: A robot placed on a mobile platform.

A lot of effort was put into solving how to easily calibrate the robot to find and identify the peripheral equipment in the working cell. If all equipment is easily moved, one has to have a way of calibrating the robot so that it can use the equipment. A system with tactile sensors was used, so that each piece of equipment had a set of metal orbs attached. The robot then had a tool with a similar orb as an end-effector which was used to automatically locate the fixture, using a current in the tool and special search patterns. If the operator jogged the robot close to the fixture, the robot could then search and locate it. This technology was investigated and some work was made together with the developers of the system in order to modify and improve the technology. This technology was used in the prototype, but not migrated to the final version of the cell due to some technical issues in the calculation of the exact position of the fixture.

Other, more commonly used, technologies that have been found to be beneficial as enablers for reconfigurability are e.g. tool changers for the robots. Throughout the project, all the fixtures and grippers constructed have been designed to be as flexible as possible so that they fit both current and future product variants.

All those hardware solutions that enabled the use of industrial robots in this cell were still not enough. Early on in the project, the need for supporting software was identified. No software was available on the market which could be used to program, (re)configure, and control this type of robotic working cell. This lead to an extensive research and development project which resulted in a prototype of a new software tool used for controlling this type of reconfigurable robotic working cell, called the Cell Configurator.

The purpose of the Cell Configurator was to be used as a Cell-PC at the working cell. The Cell Configurator was used to program the overall logical programming of the cell, and then to control the cell using this program logic. The Cell Configurator could connect to several industrial robots at a time and thus be used to control a cell containing multiple robots cooperating to complete a task. The software was also used to control peripheral equipment which could be controlled using I/O signals or software commands. Figure 4 shows a schematic view of how a cell-PC (a) can communicate with a robot controller (b), which in turn is used as an I/O device to control the industrial robots and peripheral equipment (c). The cell-PC can also read and write data from other external systems (d) in the surrounding IT-structure.

The Cell Configurator contains three main user interfaces: (1) a cell designer where the layout of the cell could be drawn using drag-and-drop techniques, (2) a sequence editor used to program the cell using icon-based flowchart programming style, and (3) a production view where the user can start/stop production and see current status together with statistics and maintenance schedules.

During the development of the software, a set of requirements was identified. The requirements originated from the users perspective of the software and focused on ease-of-use and visual management. Areas that were considered as important were: visualisation so that the operator always can see what is happening in the cell; maintenance scheduling that notify the operator when maintenance is due, even by stopping the cell if necessary; systems for assisting the operator in failure recovery; and the presentation of statistics and feedback in forms which are easily understood by the operator. Since the operators at the production site had little experience with industrial robotics, this kind of assistance was of extra importance. For more information regarding the Cell Configurator see [23].

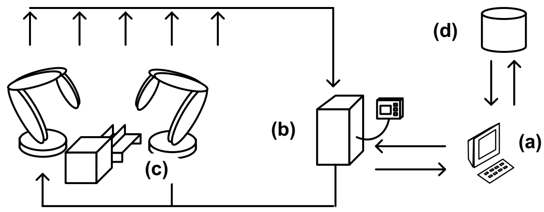


Figure 4: Schematic view of the robot working cell controlled by a cell-PC. [23]

3.5 Results of Demonstrator 1

The Demonstrator 1 project resulted in the installation of a robotic working cell into production in December 2006. The cell was designed to be flexible enough to handle all the available product variants that were currently produced, but a lot of work was also made to ensure that probable future variants could be produced. Enablers for flexibility such as grippers and fixtures designed to handle a wide variety of product variants were designed. Also, automation technologies such as barcode readers were introduced in the cell to enable automated identification of incoming components.

Another important requirement was that the cell should be *future safe*, i.e. that the company should be able to reconfigure the cell in the future, e.g. adding new hardware or operations. This was ensured using hardware solutions for easily moveable peripheral equipment and tools for calibrating the robots. The original cell introduced into production in December 2006 contained two industrial robots and a large amount of peripheral equipment, however the cell was re-configured during the summer of 2007 when a third robot was added to the cell and a new screwing operation was introduced.

One important result of the study was the need for software support tools and the development of the Cell Configurator. This software can be used to connect to a robot controller, automatically read all the available data, I/O signals, program routines etc. and then use this information to build up the overall logical program of the cell. During production the software is used to control and monitor the production process in the cell. The software also enables communication between the robot cell and surrounding IT-structure, e.g. reading production orders from barcode readers or production planning systems.

One of the more important functions of the software support tools of a robotic working cell of this kind is to provide an easy-to-use user interface. As the systems contain more and more functionality and components they become more complex, which in turn leaves the user out-of-the-loop and reduces the users system awareness. The user interface is the communication path between human and technology and should thus be design in a way so that it reduces the perceived complexity and guides the operator to make the correct and well informed decisions.

4 THE CONCEPTS OF FLEXIBILITY AND RECONFIGURABILITY

The project has focused on how to realize flexible and reconfigurable robot automation in an industrial setting, where the users of the developed system are inexperienced within the area of industrial robot automation. There is a scientific relevance in this kind of studies since flexibility, reconfigurability, and agility are concepts of importance for manufacturing systems.

The concepts of flexibility and reconfigurability have been studied, both from a theoretical perspective and from a practical perspective during the development of the robotic working cell. The theoretical perspective of the study has been partly presented in Section 2 of this paper, whilst the practical part is presented briefly in Section 3.

As presented in the theory presented in section 2 there is a need for more flexible and reconfigurable robotic working cells. The type of flexibility that has been primarily studied in this case study is product flexibility, i.e. the manufacturing systems capabilities when it comes to handling different types of products or product variants. Another way to discuss the concept of flexibility is the systems ability to handle unexpected product changes, leading to a more robust system. Flexibility can be measured as the systems ability to handle e.g. product variants of different sizes.

An example could be as in the case presented in section 3 where the robotic working cell could handle cabinets of different heights. In that case the scope of flexibility over the height of the product was between 700mm and 1000mm. The envelope of flexibility is defined as a combination of all the available measurements of flexibilities for a system. Figure 5 shows a diagram over the envelope of flexibility, as defined by two scopes of flexibilities. The flexibility of a system is the systems ability to handle changes, as is, without changing the system. Using the flexibility of a system is rapid, since no, or very limited, changes have to be made. A small configuration could perhaps be made in parameterized control software in order to utilize the flexibility in the hardware. In the case of a robotic working cell, one could assume that the production personnel would like to make a test run of the system with limited speed to ensure that it can safely handle the new product variant, which results in a short time loss when restarting the system for the first time.

The reconfigurability of a system is the systems ability to adapt to changes, rapidly with little effort. A reconfigurable manufacturing system is designed so that it is easy to implement changes in the system, e.g. exchanging, adding, or removing equipment, see Figure 6. Designing a system for reconfigurability is having a strategy for how to extend the systems current capabilities and capacity, along with the technical enablers to achieve the reconfiguration.

Each time a reconfiguration is performed, the systems capabilities change, as do the systems envelope of flexibility.

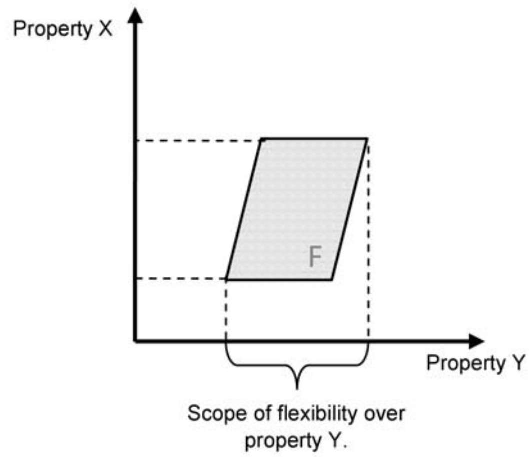


Figure 5: The envelope of flexibility of a system with two properties, X and Y, which have different scopes of flexibility.

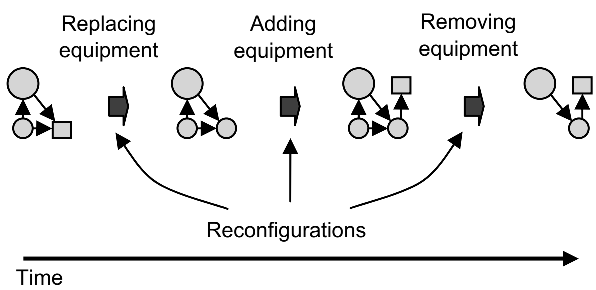


Figure 6: A diagram showing different reconfigurations of a manufacturing system over time.

5 CONCLUSIONS

From the literature a lot of information, measurements, and classifications of the concept of flexibility can be found. However, the concept of reconfigurability is not as deeply covered within literature. As well, there is a need for more practical application of the two concepts within industry in order to evaluate and continue the development of the two concepts. This paper has presented a case study performed within industry where the concept of reconfigurable robotic working cells have been investigated and realized.

Several conclusions may be drawn from this case study. Firstly, there is a need for more reconfigurable robot automation as this will lower the bar that companies have to pass in order to invest in this type of technologies. Reconfigurability allows the company to stepwise invest in automation technologies which may help the company to get accustomed to the technologies. Re-usability of equipment is also an enabler for companies to invest in production equipment which is not dedicated for a certain product or project. This also allows companies to think in terms of life-cycle cost when purchasing equipment.

Secondly, another conclusion to be drawn is the importance of software support. This is shown within literature, e.g. [13], but also very much an experience learned during this research project. The Cell Configurator developed during the project shows promising results in terms of supporting the user in both everyday production and during programming and (re)configuration.

As the production department, where the case study was performed, was not experienced within the area of industrial robot automation, several issues around the complexity of proposed solutions had to be dealt with. The software tool developed had the purpose of reducing the perceived complexity of the system so that the operator did not feel lack of system or situation awareness. Areas that were identified as important included: intuitiveness in user interface, visualisation of current state of production, maintenance scheduling, provide feedback in terms of statistics and production data, and assisting in failure recovery. Other efforts made to reduce the perceived complexity include automatic reading of data from the robot controller to the PC and the implementation of an icon-based programming style.

6 ACKNOWLEDGMENTS

This research have been funded and supported by the Swedish Foundation for Strategic Research through the ProViking programme as part of the Factory-in-a-Box research project. Also, the industrial partners of the project have contributed with time, funding, and competence.

7 REFERENCES

[1] Slack N., 1987, The flexibility of manufacturing systems, International journal of Operations & Production Management 7:35-45

[2] Hauser, D.P., de Weck, P.L., 2007, Flexibility in component manufacturing systems: evaluation framework and case study, Journal of Intelligent manufacturing, 18:421-432

[3] Yin, K.R., 1994, Case study research – Design and Methods, 2nd edition, ISBN: 0-8039-5663-0

[4] Jackson, M., 2000, An Analysis of Flexible and Reconfigurable Production Systems – An Approach to a Holistic Method for the Development of Flexibility and Reconfigurability, Dissertation No. 640, Linköping University, Sweden

[5] Shewchuk, J. and Moodie, C., 1998, Definition and Classification of Manufacturing Flexibility Types and Measures, The international Journal of Flexible Manufacturing Systems, 10:325-349

[6] Narain, R., Yadav, R.C., Sarkis, J., and Cordeiro J.J., 2000, The Strategic Implication of Flexibility in Manufacturing Systems, International Journal of Agile Management Systems, 2/3:202-213

[7] Browne, J., Rathmill K, Sethi SP, Stecke KE, 1984, Classification of flexible manufacturing systems, The FMS Magazine April pp. 114-117

[8] Andreasen, M.M., Ahm, T., 1988, Flexible Assembly Systems, IFS Publications UK, Springer-Verlag.

[9] Kaighobadi, M. and Venkatesh, K., 1994, Flexible Manufacturing Systems: An Overview, International Journal of Operations & Production Management, 14/4:26-49

[10] Sanches, A., 1994, FMS in Spanish Industry: Lessons from Experience, Integrated Manufacturing Systems, 5/2:28-36

[11] Mehrabi, M.G., Ulsoy, A.G., Koren, Y., 2000, Reconfigurable manufacturing systems: Key to future manufacturing, Journal of Intelligent Manufacturing, 11/4:403-419

[12] Mehrabi, M.G., Ulsoy, A.G., Koren, Y., Heytler, P., 2002, Trends and perspectives in flexible and reconfigurable manufacturing systems, Journal of Intelligent Manufacturing, 13/2:135-146

[13] ElMaraghy, H., 2006, Flexible and reconfigurable manufacturing systems paradigms, International Journal of Flexible Manufacturing Systems, 17:261-276

[14] Euron, 2005, EURON Research Roadmap V.1. Available online at: http://wwwiaim.ira.uka.de/euron/cwiki.php, visited: 2008-01-15

[15] Canny, J., Goldberg, K., 1994, "RISC" for Industrial Robotics: Recent Results and Open Problems, Proceedings of IEEE ICRA 1994

[16] Gröndahl, P., Onori M., 2000, Standardised Flexible Automatic Assembly - Evaluating the Mark IV Approach, Assembly Automation, 20/3:217-224

[17] Brunn, P., 1996, Robot Collision Avoidance, Industrial Robot, 23:27-33, MCB University Press, ISSN: 0143-991X

[18] Deshmukh, A.v., Talavage, J.J,, Barash, M.M., 1998, Complexity in manufacturing systems, Part 1: Analysis of static complexity, IEE Trans. 30:645-655

[19] Kuzgunkaya, O., ElMaraghy, H., 2006, Assessing the structural complexity of manufacturing systems configurations, International Journal of Manufacturing Systems, 18:145-171.

[20] McCarthy, D. and Rich, N., 2004, Lean TPM – A blueprint for change, Elsevier Ltd

[21] Miller, C., Parasuraman, R., 2003, Who's in Charge?; Intermediate Levels of Control for Robots We Can Live With, IEEE International Conference on Systems, Man and Cybernetics, pp. 462-467

[22] Hedelind, M., Jackson, M., Funk, P., Stahre, J., Söderberg, R., Carlsson, J., Björkman, M., Winroth, M., 2007, Factory-in-a-Box – Solutions for availability and mobility of flexible production capacity, In proceedings of SPS'07, 28-30 August

[23] Hedelind, M., Hellström, E., 2007, Programming of Reconfigurable Robotic Working Cells, In proceedings of IFAC MCPL, 27-30 September, Sibiu, Romania

Linking Modules for Reconfigurable Manufacturing Systems

Uwe Heisel, Walther Maier and Alexander Bader

Institute for Machine Tools (IfW), Universität Stuttgart, Germany

Abstract

Within the framework of a Transfer Unit supported by the German Research Foundation, the IfW (Institute for Machine Tools) of the Universität Stuttgart is developing new designs for the linkage of manufacturing systems for its partners, the company Krauseco Werkzeugmaschinen GmbH and Daimler AG. Based on the requirements on a reconfigurable production line with new innovative machine concepts, an interlinked production unit for the machining of parts such as connecting rods or cylinder heads has been developed from several so-called „inverse kinematics machines" and an innovative linking module.

Keywords:

Manufacturing System; Reconfiguration; Linking

1 INTRODUCTION

1.1 General approach

A high speed of innovation and the fast change in consumer goods require a rapid and easy adaptability of production facilities and machine tools. Neither dedicated machine tools, frequently used in the past, nor highly productive transfer and assembly lines can be used for low unit cost as well as a high productivity combined with a great flexibility and adaptability. Flexible transfer lines enable a quick retoolability to different workpieces of a spectrum of parts. On the basis of analyses made at a big car manufacturer, it was, however, established that a configuration once planned and executed usually remains unchanged. The reason for this is that the present system boundaries have complex module interfaces, which requires a considerable amount of time and money even at the configuration of a transfer system. This also means that a reconfiguration would at least be just as time-consuming and cost-intensive [1, 2]. The machining sequence as well as the machining tasks are rigidly interlinked in transfer systems and hence can only be changed with great expenditure of time and money. Flexible manufacturing systems (FMS), however, finish the complete product on one machine on several universal machining centres with automated work supply and tool-changing systems mostly at the same time. However, due to the frequent tool change within a machining centre, the non-productive times, the susceptibility to failure and the failure frequency increase so that productivity declines.

1.2 Inverse kinematics machines

Shorter non-productive times can be achieved by switching tool and workpiece in the new „inverse kinematics machines". In these machine concepts the workpiece is moved relatively to the tools. The single-spindle or multispindle heads needed for machining can be perfectly designed for the machining operation and orientated to the machining task. In comparison with highly productive transfer lines, a considerably greater flexibility regarding the adaptiveness of manufacturing systems is achieved by a rapid adaptability or reconfigurability of the spindle systems in the range of a few hours. However, the machine concept has only a limited number of possible manufacturing steps due to its compact design.

1.3 Linking systems

All manufacturing systems, however, have a low adaptability of the workpiece transport and loading systems, the so-called machine linkage, in common. Figure 1 shows advantages of different production and machine concepts for series parts, the automation and linking ability as well as the inferred need for adaptable linking.

	Productivity	Flexibility	Susceptibility	Machining Steps	Linking	
Transfer Lines	++	o	o	++	++	poor adaptability reconfigurability
FMS	+	+	+	++	+	
Machining Centres	o	++	+	++	+	
Inverse Kinematics	++	+	+	o	o	

++ very good + good o less appropriate

Figure 1: Need for adaptable reconfigurable interlinked manufacturing systems

2 SYSTEMS ANALYSIS

The aim of the systems analysis is the exact and fundamental functional description of the linking module, allowing for the determining boundary conditions. A functional analysis here is a process completely describing the functions and their relations, which are systematically presented, classified and evaluated [3]. According to Akiyama [4] the object to be examined, in this case the linking module, is therefore to be analysed for its effect, purpose and concept. The object is to be subdivided into its different components, elements, aspects, etc, and these have to be abstracted, divided, classified and determined with respect to its different characteristics, features, attributes, etc. The systems analysis finally produces as result a verbal formulation of the development task's essential core and the representation of the overall function of the object to be analysed in a functional structure.

2.1 Verbal formulation as abstraction

Within the framework of the analysis phase, the development task is analysed by using the method of abstraction. Based on a functional classification of handling systems, the essential core of the overall function was verbally formulated according to VDI standard 2860 [6]. Regarding the criterion of influencing the program, it is concluded, in consideration of the requirements resulting from the overall task, that the function of the workpiece movement within a manufacturing system can only be realised by a user-definable motion machine with automatic program selection. Here the programs for motion control are automatically selected by external signals of the assigned production facilities.

2.2 Function structure

Apart from the verbal formulation of the overall function, the establishing of a function structure helps to clearly define the essential core of the development task. A functional structure represents particular aspects of the mutual assignment of functions [5]. Physical, ergonomic, organisational, administrative or other reactions underlie verbal assignments, forming the basis of the logical organising principle in a functional structure [4]. In Figure 2, the process of organization in a reconfigurable manufacturing system is illustrated as function structure with symbols according to standard VDI 2860 [6]. The linking module contains the functions holding and loosening a workpiece and the essential function arranging.

After an analysis of system the linking module can be subdivided into a module with the subfunction holding or loosening and a kinematics module with the subfunction arranging. The function arranging here corresponds to the verbally formulated main function moving, but contains the technical aspect with greater exactness in view of the further realisation. From a kinematic point of view, the function arranging consists of an orientation and positioning task. The orientation and positioning describe the spatial arrangement of a geometrically defined object in the three-dimensional axes. The subfunction arranging can be subdivided into the elementary functions shifting and rotating, and swivelling as combination of the two elementary functions.

Taking account of the boundary condition regarding the linking task within a reconfigurable manufacturing system, inverse kinematics machines as production facilities leaves the conclusion that workpieces do not have to be orientated within handling. The reason for this is that the tool spindles on the machine frame can be orientated and fixed relative to the tailstock quill and the workpiece or workpieces, in the case of production facilities. Hence the subfunction of arranging can be described by the elementary function of shifting. Shifting is here the moving of an object from a given position along a straight line by means of translation. The orientation of the object remains unchanged in this process [6].

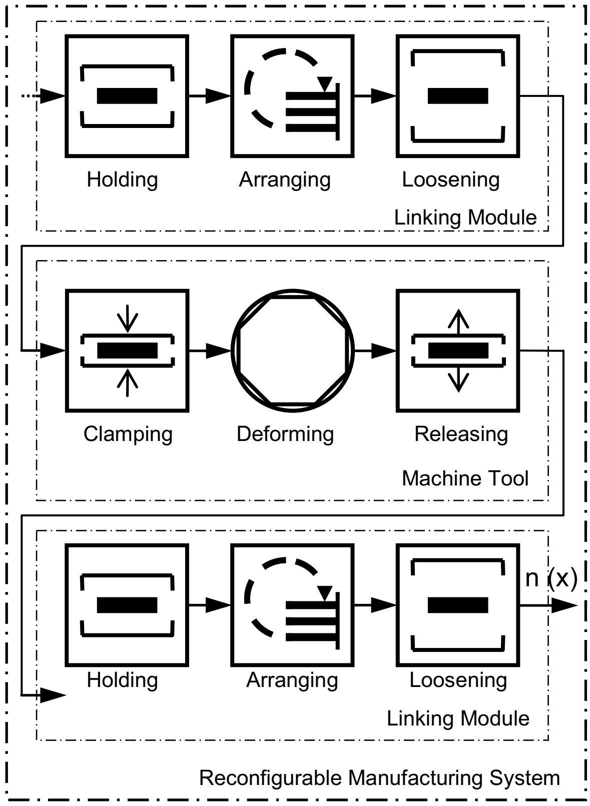


Figure 2: Organisational functional structure of a reconfigurable manufacturing system

3 KINEMATIC CONCEPTS

In the development process it is subsequently necessary to find active principles serving the function of shifting, determined in the systems analysis. The search is based on the condition that the workpieces to be transported are moved as directly as possible, i.e. more or less along a straight line and are not changed in their orientation at the respective transfer position. The external boundary conditions lay down as further demands that no guideway or bar must remain in the working space of the production module. Consequently, the kinematics to be established must be of a projecting type.

The search for a solution was conducted systematically. The solutions, established discursively, were systematically classified (Figure 3). The feature of systematisation is the type of movement. A distinction is made between translational and rotational motions, also taking account of the combination of translational and rotational motions.

The concept T1, which was further developed, represents a vertical lambda kinematics (the components are arranged like the Greek lambda symbol). Due to the fact that the long strut is designed as parallelogram, the kinematics has no rotational degree of freedom at the projecting end of the long struts. Thus the condition of the constant orientation is met.

The positioning along a straight line is met because only the upper nodes of the long struts can be traversed along a linear telescope kinematics. The control of the slide for workpiece transport is linked with the stroke movement. Hence this telescope kinematics also has only one translational degree of freedom, and a change in orientation does not occur.

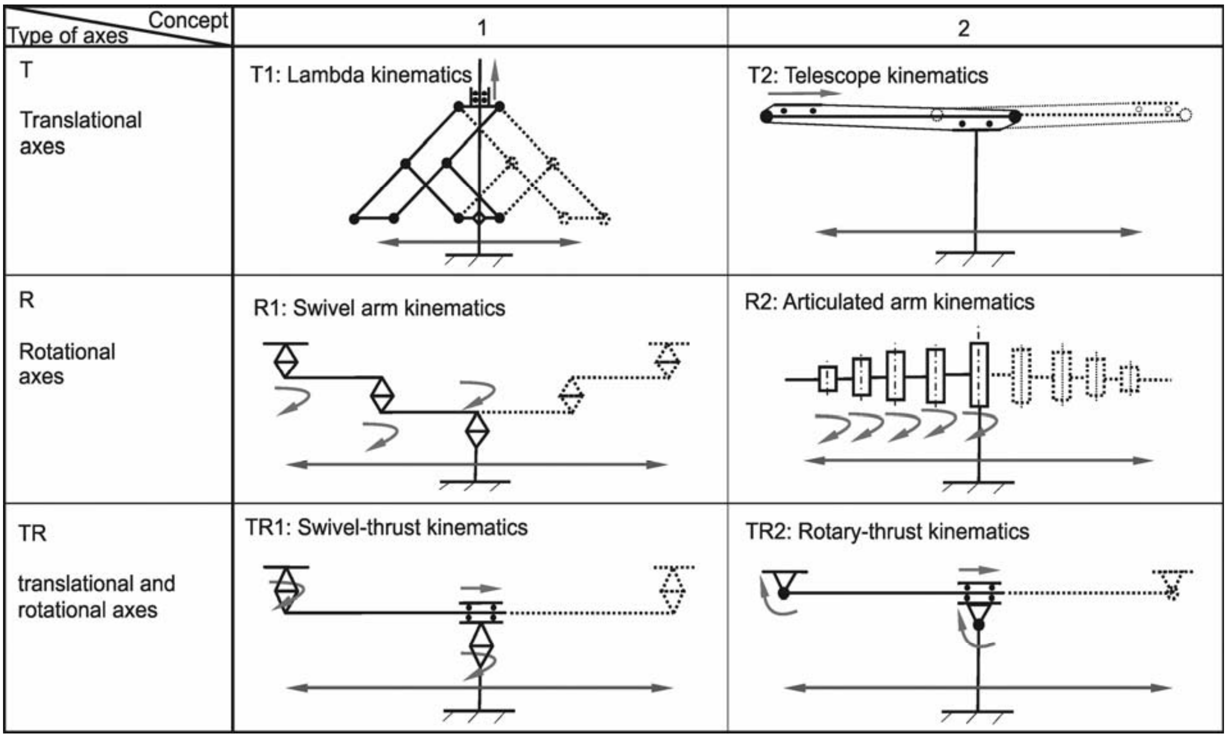


Figure 3: Classification scheme of the kinematics concepts according to the type of axes

Another possibility to position an object along a straight line is to combine rotational axes. The concept R1 of a swivel arm kinematics has three rotational degrees of freedom. The workpiece can be positioned along a straight line through controlled rotary motions. In design R2 the positioning task is realised by an articulated arm consisting of several links identical in design, similar to the segments of an earthworm. Every link here has one rotational degree of freedom, by means of which a workpiece can be positioned approximately along a straight line.

The combination of translational and rotational degrees of freedom was applied in the kinematic concept TR1 and TR2. TR1 is a swivel-thrust kinematics with one linear axis and two axes of rotation. The design TR2 is a rotary thrust kinematics. It also has one linear axis and two axes of rotation, which are, however, varied in orientation compared with design TR1.

4 EMBODIMENT DESIGN OF A MODULE FOR LINKING

As described in Chapter 3, the concept T1 “lambda kinematics“ was chosen as motion kinematics for the linking module. Years ago the advantages of the lambda principle were already recognised during the main time of research on „parallel kinematic machine tools“. E.g. in 2002, a λ-kinematics was applied to a production-oriented parallel kinematic woodworking machine by the IfW within the framework of an international Brite-Euram research project [7]. The concept design for the linking module has the following characteristics:

- simple compact slim structure
- modular design
- horizontal transfer is possible with only one drive
- flexibility through reconfiguration of linkage points
- many standard components.

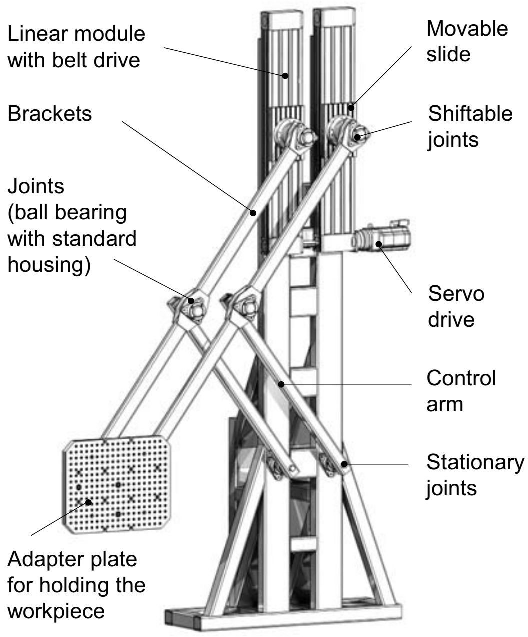


Figure 4: Embodiment design of a module for reconfigurable transport systems

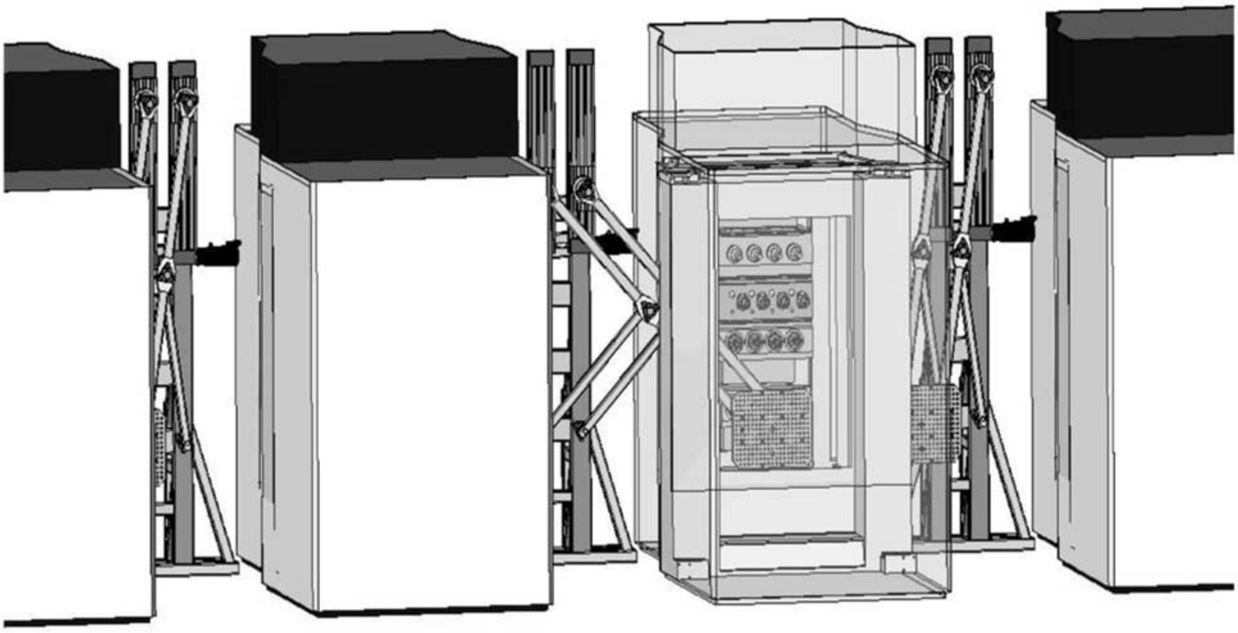


Figure 5: Production line with adaptable linking module

Figure 4 shows the design of a linking module. Corresponding to the long sides of the λ symbol, two brackets serve for transporting the workpiece between two inverse kinematics machines. In the conception, the brackets are designed as parallelogram to block the rotational degree of freedom of the adapter plate. This leads to a structure reducing the load deformation at low dead weight of the bracket. Corresponding to the short sides of the λ symbol, two control arms connect each bracket to the frame via two joints. By shifting the linkage points of the joints, the kinematics can be changed with respect to its positioning range as well as its parameters of acceleration, force and speed. A standard clamping plate is attached to the end of the brackets, built up of two standard profiles, via two further joints. A commercially-available gripper for workpiece transfer can be mounted on this clamping plate. Two further joints are mounted on two linear modules, performing the movement of the linking module. The commercially-available linear modules with belt drive are linked with a servo drive with brake via a coupling. Hence, through the independent linking module, a kind of minimal robot is available for the flexible reconfigurable linkage of production facilities. Figure 5 shows a part of a production plant made up of "inverse kinematics machines" and the newly developed modular linking module. The linking module in the middle is just in transfer position. The tailstock quill (not visible in the figure) of the inverse kinematics machine takes the workpiece and transfers it to the tools opposite, after the linking module has left the working space.

5 SUMMARY

This paper presents the design and development of a linking module for „reconfigurable manufacturing systems". It was found that a **good adaptability of manufacturing systems primarily** requires **adaptable ubiquitously networked machine linking**. The described approach with a linking module, independent of the machines used, meets this demand perfectly.

6 ACKNOWLEDGMENTS

We thank the German Research Foundation (DFG) for supporting the Tansfer Unit "Transformability in Multi Variant Serial Production" (TFB 59).

7 REFERENCES

[1] Heisel, U.; Koren, Y.; Jovane, F.; Moriwaki, T.; Pritschow, G.; Ulsoy, G.; Van Brussel, H.: Reconfigurable Manufacturing Systems. In: Annals of the CIRP (1999) Vol. 48/2.

[2] Heisel, U., Meitzner, M.: Progress in Reconfigurable Manufacturing Systems. The International Journal for Manufacturing Science & Production 6 (2004) Nr. 1 - 2, S. 1 - 8.

[3] Norm DIN EN 1325-1-1996: Value management, value analysis, functional analysis, vocabulary - Part 1: Value analysis and functional analysis; German version EN 1325-1-1996.

[4] Akiyama, K.: Function Analysis - Systematic Improvement of Quality and Performance. Cambridge/Mass., Norwalk/Com, USA: Productivity Press, Inc. 1991.

[5] Pahl, G.; Beitz, W. et al: Konstruktionslehre. Springer Verlag, Berlin, 2003. 5. Auflage.

[6] Norm VDI 2860-1990: Montage- und Handhabungstechnik; Handhabungsfunktionen, Handhabungseinrichtungen; Begriffe, Definitionen, Symbole.

[7] Heisel, U., Maier, W.: Investigation of Truss Structures as Light Weight Elements for the Use in Parallel Kinematic Machines. In: PKS 2002, Proceedings, Herausgeber: Neugebauer, R., S. 715-728.

Global Footprint Design

Günther Schuh[1], Sebastian Gottschalk[1], Philipp Attig[1], Axel Hoeschen[1]
[1]Laboratory for Machine Tools and Production Engineering WZL, RWTH Aachen University, Germany

Abstract
The configuration of global production networks is often times rather the result of coincidence than strategy. Evaluation of production locations and network configurations is mostly based on costs and operative restrictions, but few approaches exist to evaluate possible benefits. "Global Footprint Design", is an integrated heuristic approach to measure and evaluate the strategic value of a specific configuration of a production network with respect to the network's business environment. The strategic value is herein deduced as the degree of correspondence of the network's characteristics to the procurement market on the one hand and to the sales market on the other.

Keywords:
Production Management, Production Engineering, Production Network

1 INTRODUCTION

Even though, relatively strong growth rates within Western and European economies, successful attempts to increase competitiveness through rationalisation and innovation and record growth rates in China and India have somewhat eased the pressure on costs over the last years, going abroad with production sites remains (and again becomes) a major issue. Especially the continuous weakness of the US-Dollar makes production within dollar linked economies particularly attractive.

Production engagement is per definition long term engagement and cannot and must not be based solely on momentary market effects but instead on a strategy to develop and consolidate a corporate production profile with respect to the regional and global distribution of production capacities, a Global Footprint.

2 THE PRINCIPLES OF DISTRIBUTED PRODUCTION NETWORKS

2.1 Aproach

Fundamental research in the field of locally distributed production processes has found a large variety of potentially relevant characteristic business patterns, as well as numerous hypothesis to describe interdependencies between them [2] [4] [5] [6] [9] [10] [16] [17] [18]. It shows, that the distribution pattern of a production network should match the business environment's implications as largely as possible. This "Congruency-Efficiency-Hypothesis" is a widely accepted causal chain of matching system and environment, Considerable coherence can be found between business patterns and environment, such as closeness to the customer, infrastructure or availability of a skilled workforce [3] [7].

The core idea of forming production networks is to separate integrated elements of the value chain to dislocate them into more favourable environments ("cherry picking") [7]. The potential of a networked production relative to a single, integrated production site therefore results from a *heterogeneous Production environment*.

The Design of the local distribution of the elements of the Production Network, thus the design of the Global Footprint, is therefore the mediator of *conflicting characteristics* between sales- and factor environment.

2.2 Characterization of Production Systems in two dimensions of heterogeneity

In the classical Management Theory, the different impact factors of the Production System are being subsumed into the term of **factual heterogeneity** [1] [15]. A task or service may be factually homogenous or heterogeneous, depending on how specific its requirements and needs towards the factual sophistication of the process are.

In addition to the factual heterogeneity, the **local heterogeneity** must be appreciated and described [9]. This local heterogeneity, subsumes the impact factors related to the local position where the production task is undertaken. An example for a locally homogeneous production task may be the production of a highly standardized bulk product, a locally heterogeneous task may be the maintenance service mentioned before, which obviously can only be performed on location.

Since both, factual and local heterogeneity subsume a plethora of different single independent factors, the final evaluation takes only the most heterogeneous factor into consideration, for high requirements towards one factor can in the majority of cases not be compensated by lower requirements towards other factors. The solution here could only be the further separation of the value chain, where only those services and tasks are executed on location which are absolutely necessary. To model and evaluate the production systems characteristics, a two dimensional portfolio is suggested, in which both degrees of heterogeneity are displayed accordingly (see Figure 1).

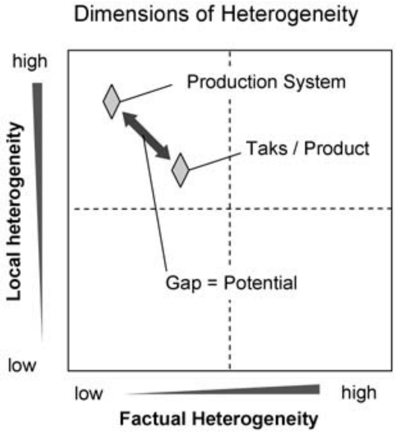


Figure 1 : The Potential of a Network Structure described as the Difference between given and demanded heterogeneity.

2.3 Characterization of a business environment

Likewise, the business environment in terms of sales market and factor market can be displayed in relation to the Production network. In Terms of the sales market, the indication is the limitation to the transport, be it out of factual (e.g. shelf life of perishable foods) or out of economical reasons (e.g. trade barriers, customs, local-content regulations etc). The extrema of the measurement are "locally restricted" and "locally unrestricted".

The factual heterogeneity describes the variety and heterogeneity of the demand, indicating the specificity of consumer requirements. Extrema here are "unspecific", meaning largely homogeneous standard products and "individual", where each customer requires a tailor made product.

Displayed in a Portfolio with four sectors, a typology of the sales markets emerges, displaying four characteristic types (Figure 2):

Type A:

Locally restricted but factually unspecific, this type of sales environment requires for standardized and unspecific supplies from local suppliers, examples can be perishable food ingredients such as meat or fruit, vendor parts under local-content regulations or products being too big to be transported efficiently.

Type B:

Locally and factually restricted, this type of sales environment requires for highly individualised products and services in proximity to the customer. Examples are just-in-sequence Production needs of individual vendor parts, markets with regional or cultural transfer barriers or plant engineering and construction

Type C:

Locally and factually unrestricted, this type of sales environment requires for bulk articles and services with no specific customer requirements and no preference as to where the product comes from and how far it has travelled. Examples include white goods, consumer electronics, most raw materials and certain commodities.

Type D:

Locally unrestricted but factually individual, this type of sales environment requires individual articles and services without customer requirements towards the product's origin. Examples are individually configured cars, machines or other engineering-to-order products and services.

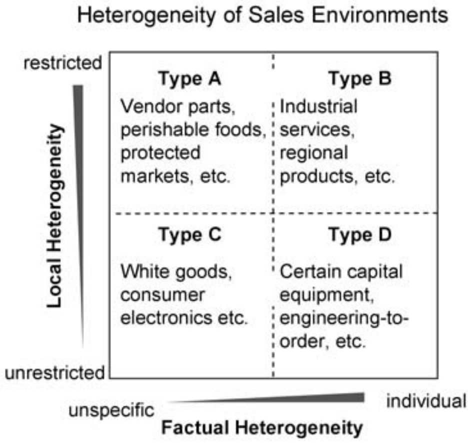


Figure 2: In terms of local and factual heterogeneity, four basic types of sales environments can be differentiated.

To complete the description model for distributed production networks, the factor market has to be described accordingly: *factual heterogeneity* is here the degree of the supplier's ability to provide for individual needs, for instance vendor parts or commodities produced individually. *Factual homogeneity* in turn means the supply of a standardised product, energy for instance.

A good example for *local heterogeneity* of the factor market could be a highly trained and experienced personnel that can not be found anywhere else or the existence of a special infrastructure to allow for heavy weight transportation.

Through systematically describing the difference in heterogeneity between factor market and sales environment, the potential of a production network over a single production site can be derived, since according to the Congruency-Efficiency-Hypothesis, such combinations of factor market, production system and sales environment are expected to be most beneficial, which best match (production-) system and (production-) environment (Figure 3).

3 DESIGNING A COMPANY'S GLOBAL FOOTPRINT

3.1 Global Footprint Design as a strategic approach to design distributed production networks

Gaining these potential benefits requires for a network setup to overcome these differences and form a distributed process in accordance to the business environment. To do so, the general term of heterogeneity must be visualized and examined in both dimensions separately, to derive a congruent strategy.

In terms of local heterogeneity, a perfect match between business environment and production system manifests itself in a minimal transport effort between resource, Production site and customer. In situation where factor market and sales environment already match, the solution is trivial: a local resource to be transformed in a product only locally needed results in a single local production site to minimize transportation effort. The strategic dimension comes into play where the gap between these two offers a large number of different solutions.

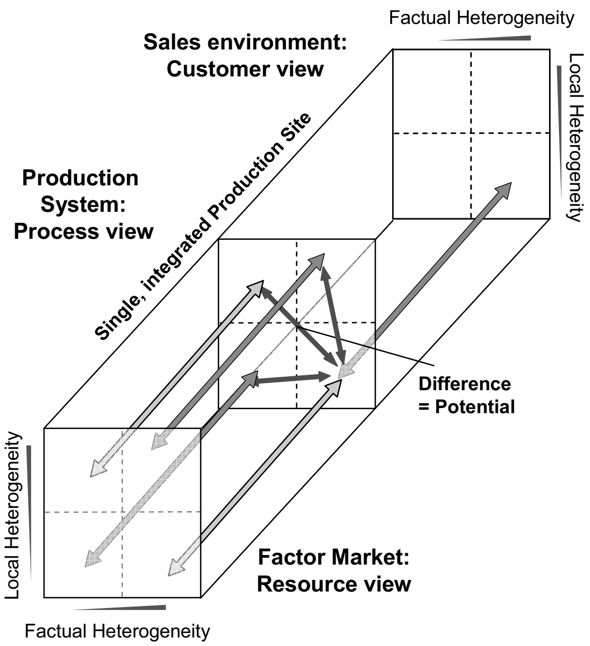


Figure 3: The Production System links the factor market with the sales environment.

In a given case of a white goods producer, only few local restrictions of the sales environment exist, resulting from the customers being largely indifferent as to where the product is produced. The only restriction are the transportation cost, as the relatively low value per volume results in high transportation costs, which in case of North America makes up to 30% of the sales price, leaving the local heterogeneity somewhat restricted.

As of the factual heterogeneity, the product range is highly variant, both in terms of local preferences of the customers as well as a large number of alternative functional designs and a competitive market environment, resulting in wide range of different models under the same brand. In terms of the sales environment, the preferable Global Footprint would be regional (in this case national) production sites, producing the preferred products for the individual market.

In terms of the resource market however, the operation of a large number of factories would boost transaction costs and leave possible economies of scale wasted. That is: with the resources being locally and factually unspecific, a notable gap between the factor markets unspecific and ubiquitous supply and the both, regional and factual specific customer's needs displays a strong potential to be used.

The solution here derived from the Global Footprint Design heuristics is be to separate the production process into standardized, easily transportable parts, such as electric motors and electronic components who themselves have neither factual nor local restrictions, matching the profile of the factor market for labour, energy and raw material. These components can now be easily transported into more specialized local manufacturing plants for white goods, which match both, local and factual heterogeneity of the sales environment.

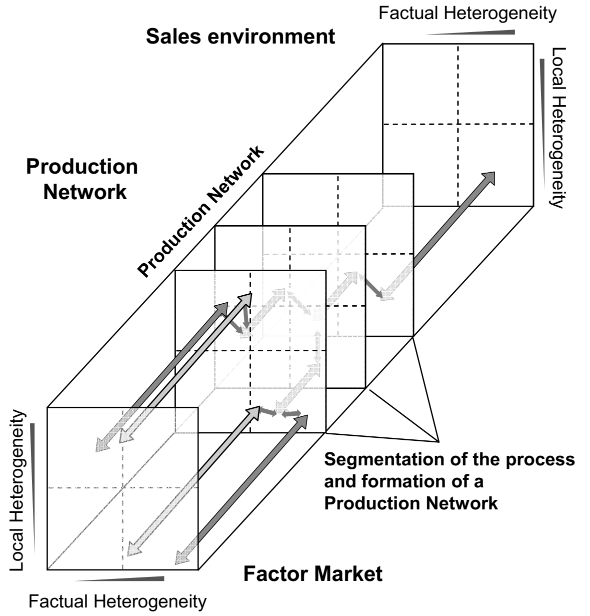


Figure 4: Gaining potential by systematically designing a production network.

A second case is the Global Footprint of an electronic components manufacturer, producing microchips for the computer industry. As for the sales market, neither factual nor local restrictions exist, as the products are small and easily to transport. On the side of the factor market, however, the main restriction is the need for a highly qualified expert workforce, which can only be found in relatively small numbers in few western countries. The resource market is therefore both, locally and factually heterogeneous, leading to a production network of globally distributed production sites in the vicinity of existing industries and highly educated workers, resulting in 13 factories in 3 countries to meet the global demand.

3.2 The Heuristic Approach of Global Footprint Design

The core dilemma of any networked production is finding the individual optimum within a dilemma of complexities [11]: The more spread out a networked production is, the greater its local heterogeneity, thus the more complex (and expensive) is transport and transaction (external view). On the other hand, the more centralised a production system is, the more variance it has to cope with, the higher is its factual heterogeneity, and thus the more complex (and expensive) are its internal processes (internal view). The heuristic approach of Global Footprint design addresses this dilemma by defining the optimal degree of internal (factual) and external (local) heterogeneity with respect to the production network's business environment.

To summarize the given examples, the following heuristic rules can be expressed.

With respect to the factor market:

- As far as no political or environmental transport barriers or transfer restrictions exist, the maximum national centralization is advisable to keep transport and transaction costs low, as long as the network does not exceed the manageable factual heterogeneity with respect to the factor markets capability (the factor market's factual heterogeneity).

- In case of local heterogeneity, multinational decentralisation is to be archived, regarding especially those production tasks that form the locally diverse features, as long as the network does not exceed the manageable local heterogeneity with respect to the infrastructure's capability.

With respect to the sales environment:

- As far as no factual differences between customer's needs exist, a maximum redundancy between production sites is advisable, to provide for volume flexibility and cross-plant economies of scale.
- In case of factual heterogeneity, production sites must be complementary to meet the sales environment's demands in the most efficient way.

A number of Case Studies has shown, that this heuristic approach proves most valuable in designing production networks in accordance to their individual business environment [7] [9]. The model emerges in the combination of the different dimensions of heterogeneity of factor market and sales environment to an integrated evaluation of the "fit" of an existing Global Footprint. Depending on the parameter values, three groups of Global Footprints can be derived: Four matchin types, in which sales environment's and the factor market's heterogeneity match each other in both dimensions, for vondlicting types, in which the sales environment's and the factor market's heterogeneity are of conflictive heterogeneity, and a group of types for which either the local or the factual heterogeneity are conflicting and show a gap, but not both. In such a case, the classification of the Type depends on the priority of either local or factual heterogeneity with respect to the competitive advantage of the company.

4 SUMMARY

4.1 Global Footprint Design in the scope of the strategic management

In accordance with the "Competence-Based-View" according to HAMEL / PRAHALAD [8], which unites the "Resource-Based-View" according to PENROSE [12] with the "Competitive Strategy" according to PORTER [13] [19], Global Footprint Design moves into the scope of the strategic management as an invaluable core competence to develop and maintain a company's competitive advantage in a heterogeneous global business environment [7] [9]. Where Resources and Capital are both equally ubiquitous, the ability and competence to set up and run a globally spread out production network most suitable to meet the customer's needs, Process restrictions and the factor markets characteristics becomes most important [3].

The introduced heuristic approach is the foundation for a basic model to understand and evaluate the quality of a networked production strategy. In accordance to the Congruency-Efficiency-Hypothesis, those production networks are most suitable, which have the best and most extensive fit with the business environment. This fit is hereby described degree of correspondence of factual and local heterogeneity.

4.2 Further Reference

The findings illustrated in this paper are result of the research project "Global Footprint Design", currently conducted at the Laboratory for Machine Tools and Production Engineering (WZL) of the Technical University (RWTH) Aachen, Germany, funded by the Deutsche Forschungs-gemeinschaft (DFG), and of the Study "Global Footprint Design" [7] conducted by WZL in collaboration with Roland Berger Strategy Consultants, published in 2004

5 REFERENCES

[1] Bartlett, C.A., 1989, Transnationale Unternehmung – Aufbau und Management der transnationalen Organisations-struktur, In: M.E. Porter (Hrsg.), „Globaler Wettbewerb". Wiesbaden, Germany: Gabler: 425–466.

[2] Bartlett, C.A.; Ghoshal, S.; Birkinshaw, J., 2004, Transnational Management: Text, Cases, and Readings in Cross-Border Management.

[3] Bhattacharya, A.; Bradtke, T.; Hemerling, J.; Lebreton, J.; Mosquet X.; Rupf, I.; Sirkin, H.; Young, D., 2004, Capturing Global Advantage – How Leading Industrial Companies are Transforming Their Industries (The Boston Consulting Group).

[4] Bleicher, K., 2001, Das Konzept Integriertes Management, Visionen – Missionen – Programme.

[5] Doz, Y.L.; Prahalad, C.K., 1991, Managing DMNCs: A Search for a new Paradigm. In: Strategic Management Journal, Vol. 12, Special Issue, Global Strategy: 145-164.

[6] Gälweiler, A.; Schmaninger, M., 1990, Strategische Unternehmensführung.

[7] Geisbauer, R., Schuh, G., 2004, Global Footprint Design.

[8] Hamel, G.; Prahalad, C. K., 1990, The Core Competence of the Corporation, In: Harvard Business Review, vol. 68/3: 79-91.

[9] Harre, J., 2006, Global Footprint Design.

[10] Jarillo, J.; Martinez, J., 1990, Different roles for subsidiaries: The case of multinationals corporations in Spain. In: Strategic Management Journal 11/1.

[11] Jost, P.-J., 2001, Der Transaktionskostensatz in der Betriebswirtschatslehre.

[12] Penrose, E.T., 1959, The Theory of the Growth of the Firm.

[13] Porter, M.E., 1980, Competitive Strategy.

[14] Porter, M.E., 1998, Competitive Advantage.

[15] Prahalad, C.K.; Doz, Y.L., 1987, The multinational mission: Balancing local demands and global vision.

[16] Sydow, J., 2002, Strategische Netzwerke: Evolution und Organisation.

[17] Ueda, K., 2001, Synthesis and emergence – research overview. In: Artificial Intelligence in Engineering 15: 321-327.

[18] Ulrich, H, 2001, Systemorientiertes Management.

[19] Wernfeld, B., 1984, A resource-based view of the firm In: Strategic Management Journal 5/2: 171–180.

TSUNAMI Effect Prediction Methodology for Critical Resource Analysis

Youichi Nonaka[1], Yuichi Suginishi[1], Attila Lengyel[1], Makoto Ono[1], Koichi Sugimoto[2]
[1] Manufacturing System 2nd Dept., Production Engineering Research Laboratory, Hitachi Ltd., Japan
[2] Department of Mechanical Science and Engineering, Tokyo Institute of Technology, Japan

Abstract
Tremendous growth of the home information appliances requests semiconductor manufacturing to respond *High Product Mix and Low Production Volume* condition. Such condition in manufacturing operations challenges production management to have rapid improvement activities in an environment with uncertain productivity and demand. In this research *Visualized Coefficient of Variation Analysis* (*VCVA*) is extended to analyze the correlation between *TSUNAMI Effect* of productivity fluctuation and manufacturing quality fluctuation. And with the *VCVA* visualization map, *Critical Resource Analysis* (*CRA*) is developed for manufacturing capacity design involved with *TSUNAMI Effect* prediction. Results show the effectiveness of *CRA* for manufacturing capacity estimation tool, and a factory began to use *CRA* in real.

Keywords:
TSUNAMI; Prediction; Quality; Variance; Critical; Resource; Manufacturing System

1 INTRODUCTION

1.1 Motivation

The world wide market of home information appliances was announced 492 Billion US$ in 2004 and will grow to 869 Billion US$ by 2010 [1]. In this market, semiconductor manufacturing is the key since *ASIC* (Application Specific Integrated Circuits) demand has been radical extending for the appliances. To fabricate *ASIC* in the latest semiconductor process technology, *High Product Mix and Low Production Volume* condition is required [2]. To establish this condition, rapid and continuous improvement activities with identifying manufacturing productivity detractors are desired. This research contributes to such improvement activities.

The previous report presented a new method called *Visualized Coefficient of Variation Analysis* (*VCVA*) from variability analysis method in statistics called *Coefficient of Variation* (*CV*). *VCVA* can visualize productivity fluctuation such as throughput, work in process (WIP) level, and manufacturing lead-time on time-line basis [3]. *VCVA* reveals *TSUNAMI Effect* in manufacturing system, a typical case of *Butterfly Effect*, affects manufacturing productivity. And a manufacturing productivity detractor is pointed out at the source point of *TSUNAMI Effect*.

But other subjects about manufacturing productivity detractor identification are remained since productivity fluctuation should involve an impact of manufacturing quality, too. That is, *TSUNAMI Effect* is supposed to be occurred in quality control field. This is the first motivation of this report.

The second motivation is the *TSUNAMI Effect* prediction methodology establishment. Since VCVA can visualize a manufacturing behaviour on time-line basis, a simulation approach with VCVA is supposed to be usable in Production planning and scheduling. But, the present condition of manufacturing simulation tool in industries is mainly applied to *Steady State* [4] condition estimation such as mean value and variation of WIP level, lead-time, failure rate for each process. Since *TSUNAMI Effect* is one of *Transition State* behaviours but affects *Steady State* seriously, the current *Steady State* simulation has a subject *TSUNAMI Effect* is buried to vanish. That is, a scenario-based stress simulation with *VCVA* visualization map is requested for *TSUNAMI Effect* prediction in Production planning and scheduling.

1.2 Approach

This research describes a production system as a network having nested loops for re-entrant processes, and scrap operations. The system has unreliable tools having imperfect random yield rate and uncertainty of process time and time to failure. In this condition, the focus is a manufacturing productivity detractor identification and prediction technology establishment.

In this report, *VCVA* is extended to analyze a correlation between productivity fluctuation and manufacturing quality fluctuation. And, *Critical Resource Analysis* (*CRA*) is developed to simulate *TSUNAMI Effect*, and analyzes some characters for its prediction.

1.3 Related Leterature

Manufacturing variability is studied in the field of quality control. DeVor, Chang, and Sutherland propose a process capability assessment from the viewpoint of statistical quality analysis [5]. Hopp and Spearman describe measures and classes of manufacturing variability [6].

There are also approaches that evaluate manufacturing performance in the viewpoint of flexibility. Vakharia, Askin, and Selim classify flexibility into specific levels to identify significant improvement points at the system design [7]. Parker and Wirth study key parameters for flexibility with use cases [8].

2 EXTENDED *VCVA*

2.1 Target

The previous report focused on productivity fluctuation, so throughput, WIP level at a time unit, and manufacturing leadtime for each process are sampled as the arguments of *VCVA* [3]. In this section, *VCVA* is extended so as to handle manufacturing quality management.

2.2 Methodology

In this research, failure parameter f_{ia} is defined as the amount of failure workpieces of process *i* at time period *a*. And parameter r_{ia}, s_{ia}, c_{ia} is defined as following equation.

$$r_{ia} = \frac{1}{k} \sum_{j=a-k}^{a} f_{ij} \quad (1)$$

$$s_{ia} = \sqrt{\frac{1}{k} \sum_{j=a-k}^{a} (f_{ij} - r_{ia})^2} \quad (2)$$

$$c_{ia} = s_{ia} / r_{ia} \quad (3)$$

In equations, *k* means moving average period at time *a*. That is, equation (1) means the moving average of process *i* failure at time *a* refered by failure parameter f_{ia}, and equation (2) means process *i* failure variance at *a*. In the previous report, productivity indicator p_{ij} is placed in the position of f_{ia} for the evaluation. In the same way, parameter c_{ia} can reflects manufacturing quality fluctuation.

2.3 Application Results

A correlation between productivity fluctuation and manufacturing quality fluctuation by *VCVA* was analyzed to a real assembling manufacturing system. In the system, dozens of assembling process are operated by manual, and an inspection process is placed for each group of process.

Figure 1 shows a result of productivity fluctuation analysis, and Figure 2 shows the failure fluctuation analysis result from the same data. In both figures, a series of manufacturing process in order of the manufacturing routing is indicated in a transverse direction, and manufacturing operational days are indicated in the longitudinal direction. For both figures, the background color classification matrix of c_{ia} is defined in Table 1. In Table 1, *LB* is set as 0.75 and *UB* is set as 1.33 for Figure 1, *LB* is set as 0.50 and *UB* is set as 0.80 for Figure 2. About setting LB and UB remain to be solved for the distribution characters definition of manufacturing system behaviors. In th basic study of coefficient of variability[5], Normal Distribution is assumed as the distribution of the system behaviour, but what phenomena follow Normal Distribution would be in a minority when the distribution of social phenomena or natural population phenomena is considered. On the other hand, in the real application phase, this criterion can be changed systematically by a graphycal user interface to improve the resolution accordingly.

In Figure 1, production fluctuation is converged in the end of month since bright band in lateral direction is observed. With this result, some detail investigation revealed this phenomenon was occurred by the production progress skew, since operators tend to be overworked to achieve monthly target production volume for each end of month. That is, overload of operation caused production fluctuation.

In Figure 2, some failure fluctuations occurred in several processes since red bands in vertical direction are observed. And failure rate fluctuation seems to be converged in the end month, but the convergence is not so obvious compared to the production fluctuation in Figure 1. If correlation between production fluctuation and failure fluctuation is proved, then the production progress skew is verified to affect manufacturing quality. So, to verify the correlation, a comparative statistical analysis is applied in the following.

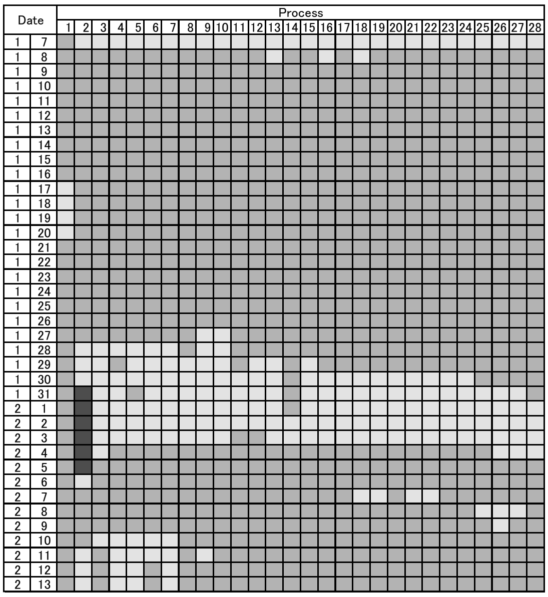


Figure 1: Daily Completion Fluctuation of Each Process

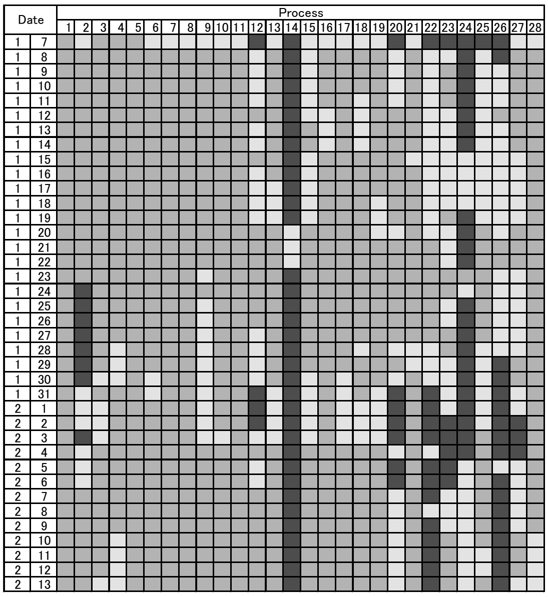


Figure 2: Daily Failure Fluctuation of Each Process

Table 1: Color definition in Visualization Matrix of *VCVA*

Type	Class	Color	Fluctuation
A	$c_{ia} < LB$		Ignorable
B	$LB \leqq c_{ia} < UB$		Middle
C	$UB \leqq c_{ia}$		Heavy

LB ; Lower Boundary, *UB* ; Upper Boundary

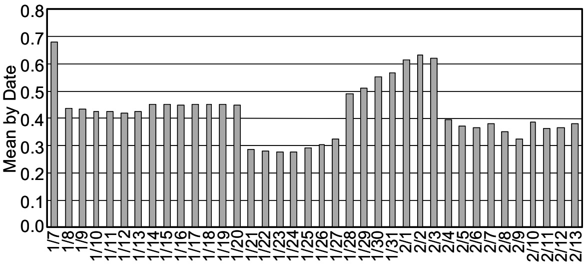


Figure 3: Daily Mean Value of Daily Completion Fluctuation

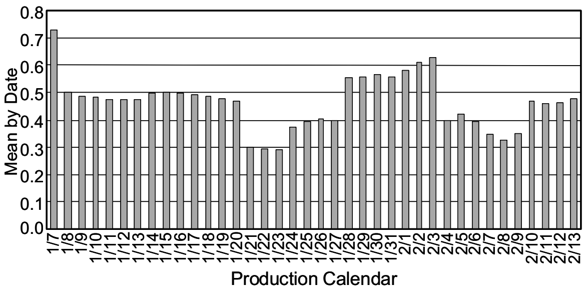


Figure 4: Daily Mean Value of Daily Failure Fluctuation

Figure 3 shows the daily mean value of daily completion fluctuation by *CVA*, and Figure 4 shows the daily mean value of daily failure fluctuation. In both figures, the bar chart of daily mean value is ordered by date. As both figure compared, a convex trend is observed in the end of the month for each figure. That is, this analysis verifies production fluctuation correlates directly with manufacturing quality. At last, since material flow fluctuation is supposed to correlate with failure fluctuation, when manufacturing productivity improvement is executed according to a *VCVA* result, not only material flow fluctuation but also failure fluctuation will be stabilized. From this result, some manufacturing productivity improvement plans are promoted to the sampled manufacturing system to avoid the production progress skew. And a CONWIP [9] type control system is now under evaluation to be applied.

3 CRITICAL RESOURCE ANALYSIS (*CRA*)

3.1 Target

VCVA can visualize manufacturing fluctuations on timeline-basis. With the use of this character, Production planning and scheduling work request a prediction technology of *TSUNAMI Effect*. But, since the present simulation tools in industries are mainly applied to *Steady State* condition, *TSUNAMI Effect*, one of *Transition State* behaviours is buried to vanish in those. That is, a scenario-based stress simulation with *VCVA* visualization map is requested for *TSUNAMI Effect* prediction in Production planning and scheduling.

3.2 Methodlogy

A simulation technology called *Critical Resource Analysis* (*CRA*) was developed to simulate *TSUNAMI Effect*. *CRA* is a kind of discrete event-based simulator, and a number of simulation scenarios can be executed in parallel for calculation time reduction since CRA uses a multi-PC control technology.

1. Input Data Setting from Real Mfg Actual Logging Database
   - Mfg Process Routings
   - Process Time / Failure Rate
   - MTBF / MTTR
   - Production Plan

↓

2. Manufacturing Fluctuation Impact Simulation
   - Simulation Warm-up for WIP Condition Matching
   - Manufacturing Condition Monitoring after A Tool Condition Change

↓

3. Manufacturing Condition Analysis for Each Process
   - Simulation Result Visualization by *VCVA*

Figure 5: CRA Process Flow Chart

Figure 5 shows the *CRA* process flow chart. In the first step, input data to *CRA* is set from a real manufacturing actual logging database. The data contains not only production plan, manufacturing routings for each product type, standard net process time, and standard setup time, but also tool MTTR (Mean Time To Repair), MTBF (Mean Time Between Failure), failure rate, batch/un-batch operation, and so on. As the second step, *CRA* simulates the impact of one or some manufacturing fluctuations set on purpose after the simulation warm-up for WIP condition matching. One or some manufacturing fluctuations are created by tool condition changes such as MTBF change, failure rate change, maintenance scheduling change, input parts scheduling change, and so on. And *CRA* simulates manufacturing condition after the tool condition change. As the final step, *CRA* sends the simulation log to *VCVA* process block, and *VCVA* visualize the simulation result on time-line basis. With these steps, *TSUNAMI* propagation can be predicted by each process on timeline-basis, so some manufacturing action plans will be made to decrease *TSUNAMI Effect*.

3.3 Application Reslut

CRA methodology was applied to semiconductor manufacturing systems to verify the effectiveness of the approach. Figure 6 shows a simulation result with a real manufacturing condition data from a complex job-shop manufacturing system that has over 700 processes. In the figure, the manufacturing operational days is indicated in a transverse direction. In the longitudinal direction, the set of manufacturing process in order of manufacturing routing are indicated. And the daily throughput is set as the productivity indicator. In this case, *LB* is set as 0.75 and *UB* is set as 1.33 of VCVA visualization matrix color definition shown in Table 1.

In this examination, *TSUNAMI* propagation is examined to predict manufacturing performance fluctuation for each process in future by a tool down on purpose in *CRA* simulation. In the figure, a down tool on purpose is placed in the former part of the process routing, and its down period is also placed the former part of the time axis.

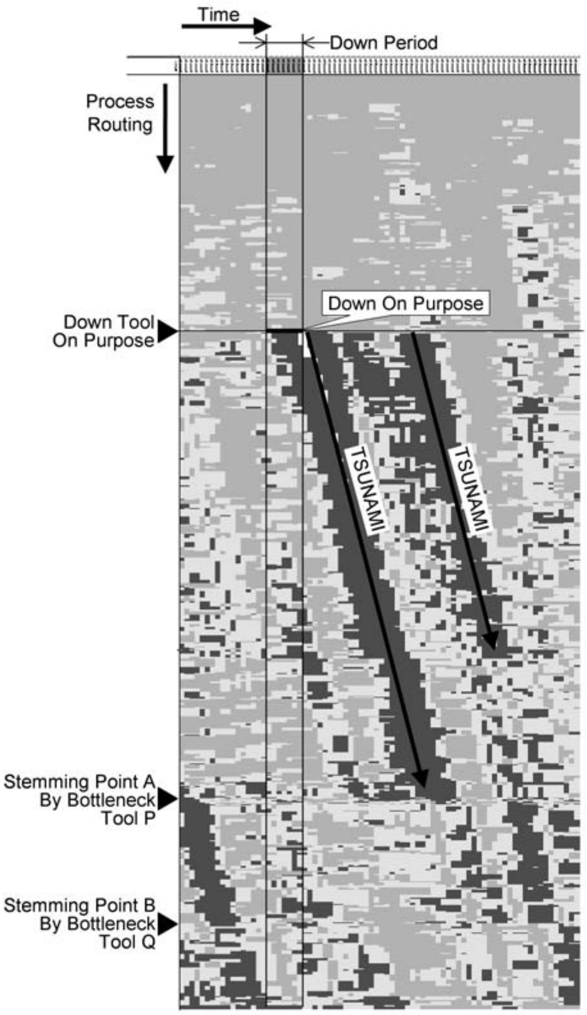


Figure 6: A Simulation Result of *CRA*

In the figure, several clusters can be seen composed by yellow and red. Most of small clusters would occur by tool process time fluctuation, but some clusters are recognized as *TSUNAMI*. As the down tool on purpose is focused, some of large *TSUNAMI* occurred from the down period. And following characters are observed.

1. *Pendulum Swing Phenomenon*

   When the down tool on purpose is focused along with time axis referred as the transverse direction in Figure 6, *TSUNAMI* occurs not only by tool down directly, but also by the aftermath like an oscillation of pendulum swing. That is, once large *TSUNAMI* occurs, Production planning and scheduling should pay attention to the tool down in past times since the available manufacturing capacity might be limited by this *Pendulum Swing Phenomenon*.

2. *Stemming Phenomenon*

   When a *TSUNAMI* is focused along with process routing axis referred as the longitudinal direction in Figure 6, some bottleneck tools in other processes stem *TSUNAMI*. In this research, a tool that have little manufacturing capacity margin is defined as bottleneck tool. In the figure, Bottleneck Tool P and Bottleneck Tool Q stem *TSUNAMI* from the down tool on purpose. That is, even when a large *TSUNAMI* occurs, latter portion of the process routing might have no *TSUNAMI Effect* if a bottleneck tool is placed between *TSUNAMI* and the latter portion.

At last, for *TSUNAMI Effect* prediction, not only the magnitude of *TSUNAMI*, but also aftermath of *TSUNAMI* defined by *Pendulum Swing Phenomenon*, and *TSUNAMI* damming of bottleneck tool defined by *Stemming Phenomenon* must be involved. And *CRA* can visualize these characters on timeline-basis, and would help Production planning and scheduling.

4 CONCLUSION

For a rapid and continuous improvement activity establishment to manufacturing productivity, manufacturing variability visualization technology called *VCVA* (*Visualized Coefficient of Variation Analysis*) was extended to analyze a colleration between productivity fluctuation and manufacturing quality fluctuation. An application result revealed *VCVA* visualized a production progress skew affected manufacturing quality. VCVA has been applied a number of factories even in assembling operation.

And with *VCVA* visualization map, the manufacturing fluctuation prediction methodology called *Critical Resource Analysis* (*CRA*) was developed for Production planning and scheduling. *CRA* revealed *TSUNAMI Effect* has some characters, e.g. *Pendulum Swing Phenomenon* and *Stemming Phenomenon*. In real, *CRA* has begun to be applied as a manufacturing capacity estimation tool when a new product or new tool is introduced to the manufacturing system.

5 REFERENCES

[1] K. Murakami, Ministry of Economy, Trade and Industry, 2004, Road Map for Increasing Profitability of Home Information Appliance Industry, NIKKEI Digital CORE, http://www.nikkei.co.jp/digitalcore/report/041102/01.html.

[2] T. Osada, 2006, Economy of Semiconductor Productivity Improvement - 450mm & 300Prime -, SEMI Technology Symposium 2006 Proceedings, 1-6 - 1-9.

[3] Y. Nonaka, A. Lengyel, and K.Sugimoto, 2007, Monitoring Variance in Manufacturing System Performances, CIRP-MS2007.

[4] S. Gershwin, 1994, Manufacturing Systems Engineering, Prentice-Hall.

[5] R.E. DeVor, T. Chang, and J.W. Sutherland, 1992, Process Capability Assessment, Statistical Quality Design and Control, Contemporary Concepts and Methods, Second Edition, 255-297.

[6] W.J. Hopp and M.L. Spearman, 1995, Variability Basics, Factory Physics, 248-286.

[7] A.J. Vakharia, R.G. Askin, and Hassan M. Selim, 1999, Flexibility Considerations in Cell Design, Handbook of Cellular Manufacturing Systems, 249-274.

[8] R.P. Parker, A. Wirth, 1999, Theory and Methodology: Manufacturing Flexibility: Measures and Relationships, European Journal of Operational Research 118, 429-449.

[9] W.J. Hopp and M.L. Spearman, 1995, Push and Pull Production Systems, Factory Physics, 349-362.

Factory-in-a-box – Demonstrating the next generation manufacturing provider

Mats Jackson[1], Magnus Wiktorsson[2] and Monica Bellgran[3]

[1, 2] Mälardalen University, Sweden

[3] Haldex AB, Sweden

Abstract

Meeting customer demands require manufacturing systems with a high degree of flexibility, low-cost/low-volume manufacturing skills, as well as short delivery times. On top of these challenges, there is a gigantic need within industry for technologies and strategies that will reduce CO2 emissions globally. In this challenging environment there is a need to identify and develop new and improved manufacturing capabilities within the manufacturing industry. The Factory-in-a-Box concept consists of standardized production modules that are e.g. installed in a container and transported by truck or by train. The concept has been developed, exemplified and realized in five industrial demonstrators developed by researchers together with competitive manufacturing companies in Sweden such as ABB Robotics, Bombardier and Pharmadule. The objective of this paper is to discuss the possibility of realizing a Product Service System (PSS) using the results from the Factory-in-a-Box project

Keywords:

Manufacturing System; Product Service System; Flexibility; Speed; Mobility

1 INTRODUCTION

Globalization makes competition within a manufacturing industry more difficult to sustain, and it is recognized that low cost and high quality is not enough to guarantee a firm's competitive position in the market place. To successfully develop and implement the next generation of products and services, industry must be successful in generating new product ideas as well as having the ability to quickly realize these into successful products and competitive production systems.

It is well accepted that the mass-production paradigm based on high-volumes has been replaced by a more flexible and responsive approach [1]. Meeting customer demands requires a high degree of flexibility, as well as abilities to reconfigure operations to suit new demands [2]. Thus, the uncertainty in markets and rapid introduction of new products has created a growing need for flexible, reconfigurable, and responsive manufacturing systems. On top of these challenges, there is a gigantic need within the manufacturing industry for technology and strategies that will reduce CO2 emissions globally.

One possible solution to the above scenario can be the concept of product service system (PSS). PSS can be described as "a system of products, services, supporting networks and infrastructure that is designed to be: competitive, satisfy customer needs and have a lower environmental impact than traditional business models" [3]. A PSS concept is based upon a shift from the traditional production of a product or service towards the delivery of a 'function' to the customer that might, in practice, mean the provision of combinations of products and services that are capable of "jointly fulfilling users needs" [4]. The consequences of this are a range of changes relating to the management of products throughout their life-cycle in an effort to minimize environmental impacts and to identify alternative profitable revenue streams. It is obvious that service activities need to cover the whole product life-cycle; affecting product, manufacturing, and customer use. In conclusion, a competitive service solution for local manufacturing is a manufacturing provider that can deliver e.g. a complete manufacturing offer including installation, ramp-up, maintenance, management and reuse.

In January 2005, the Swedish Foundation for Strategic Research started a research project called "Factory-in-a-Box". The Factory-in-a-Box concept has the key characteristic of a modular production unit that is flexible, mobile and quick to ramp-up. The Factory-in-a-Box concept has been developed, exemplified and realized in five industrial demonstrators developed by researchers together with competitive manufacturing companies in Sweden, such as ABB Robotics, Bombardier and Pharmadule. The objective of this paper is to discuss the possibility of realizing a Product Service System (PSS) using the results from the Factory-in-a-Box project.

2 METHOD

Research presented in this paper was carried out as one part of the recently finalized Swedish research project called Factory-in-a-Box. This paper is based on a literature review, a series of interviews, as well as case studies that have been used to collect data from the development of the demonstrators in the Factory-in-a-Box project.

In general, case study method is the preferred strategy when 'how' or 'why' questions are being posed, when the investigator has little control over events, or when the focus is on a contemporary phenomenon within some real-life context [5], which is relevant for this work.

Case studies are often criticized for lack of statistical reliability and validity. Furthermore, it is argued that it is not possible to test hypotheses. To overcome this dilemma, it is increasingly important to select a representative case and to validate the result continuously, and not simply at the end of the study. It is also important to describe the actual case carefully and only to draw conclusions that are valid exclusively for similar systems.

3 THE FACTORY-IN-A-BOX CONCEPT

The Factory-in-a-Box concept consists of standardized production modules that could be installed in e.g. a container and transported by e.g. truck or train. The modules can then rapidly be combined into production systems that can be reconfigured for a new product and/or scaled to handle new volumes. Production capacity may be provided as a mobile and flexible resource that rapidly can be tailored to fit the needs of a company, at a specific point of time. The emphasis on mobility in the Factory-in-a-Box concept is important in Sweden, where geographic limitations are a reality.

Three examples of customer segments for the Factory-in-a-Box are:

- Company A has a new product design, but without production capacity. The options for the company are to invest in new production capacity or to outsource production. If the company chooses to outsource, chances are that production will be placed abroad and that the company may lose control of their product. Using the Factory-in-a-Box concept, company A may instead lease a production system for the time needed and then return the system to the supplier providing the production modules.
- Company B has a peek in their production, which exceeds their capacity to produce. The company may need to outsource production or make large investments. Using the Factory-in-a-Box concept, company B may lease production capacity as an alternative to outsourcing.
- Company C is a large enterprise, which wants to have full control of the manufacturing of a sub-system, provided by a supplier. Company C can help a supplier by providing a Factory-in-a-Box module to improve quality or handle variation in production volume.

A Factory-in-a-Box could be placed close to product development or customers within the distribution chain. A likely scenario is that the Factory-in-a-Box can be rented or leased from production specialists, i.e. a type of functional sales of production capacity.

The concept of mobile capacity is not new in principle. There are examples of mobile hospitals and welfare centers see e.g. Vårdmobilen (in Swedish), mobile amusement parks, mobile libraries see e.g. Bokbussen (in Swedish), mobile sawmills see e.g. Urban Forest Ecosystem Institute, and mobile gear for musical concerts. However, there are no examples found of adding the mobility factor to a manufacturing system within the engineering industry. Adding mobility to a modular manufacturing system adds another dimension to the concept, and opens up new possibilities. The Factory-in-a-Box concept thus offers a new application of previously used ideas.

There are a number of technical requirements for the Factory-in-a-Box concept in order to realize the key features of mobility, flexibility, and speed. Examples of these requirements are:

- The modules should be easy to transport to the production site as well as easy to move at the site, e.g. having both external and internal mobility.
- Ability to dynamically adapt the degree of automation and flexibility to human interaction.
- Reconfigurability in order to meet changing demands and automatic/semi-automatic configuration of modules and system are prerequisites for scalability when changing production volumes and for fast ramp up of the production.

The use of standardized production modules provides autonomy and reusability. A Factory-in-a-Box installed at a company should be integrated with the company's existing technical production capacity and its present workforce. The intention is to balance automation and manual labour in the Factory-in-a-Box modules. Thus, each configuration will include e.g. automation requirements, operator staff requirements, configuration simulation modules, and case-based experience/knowledge databases.

4 THE FACTORY IN A BOX DEMONSTRATORS

The goal of the Factory-in-a-Box project has been to build five completely operative demonstrators – 5 Factory-in-a-Box production cells –developed in close cooperation between different academia's and industrial partners. The demonstrators have made it possible to exemplify, realize and visalize the Factory-in-a-Box concept; they are practical examples of the usability of the concept in industrial application. All demonstrators are practical solution for a particular function(s) and provide a real business case for the concept. The five different demonstrators are described in the following chapter;.

4.1 Factory-In-A-Box 1 – Automatic assembly with focus on flexibility

A first example of a Factory-in-a-Box module has been developed and demonstrated within ABB Robotics production system – an automatic production module to assemble robot components. The overall goal of this pilot demonstrator was to develop an automatic production module, which assembles robot controller cabinets, meeting the overall Factory-in-a-Box requirements of flexibility, speed, and mobility. The demonstrator has been developed in parallel with an ongoing product development project of a new robot controller at ABB Robotics: "IRC5".

The vision of this demonstrator has been a production system that can be assembled and configured according to the specific needs and that can be delivered to any location. Factory-in-a-Box module #1 has explored this vision and tried to make this a reality. The focus of this demonstrator has been to investigate the following requirements that were specified in the original project plan;

- A module that is easy to transport to the production site as well as to move at the site, e.g. external and internal mobility. In order to attain internal mobility i.e. air cushions can be used for fast and smooth transports of modules or for the entire Factory-in-a-Box.
- Reconfigurability in order to meet changing demands and automatic/semi-automatic configuration of modules and system are prerequisites for scalability for changing production volumes and for fast ramp up of production.
- Reusability of system components and modules together with simple and fast simulation and programming makes conditions for faster and cheaper system solutions and system robustness towards disturbances, especially during ramp-up of production. The reusability also makes it possible to achieve a profitable reduction of production capacity. The reusability makes it possible to reuse the equipment in other applications at the own company or in other companies. Many companies hesitate to invest in new production capacity/equipment because of the financial risk involved in the case of future declining production volumes. This concept could be a solution to that issue.
- Standardized hard- and software interfaces and integrated highly flexible production equipment, integrated metrology, and sensor-based calibration, combined with sensor-integrated robot/equipment control as important prerequisites for flexibility/agility.

The Factory-in-a-Box module at ABB Robotics consists of two robots, a number of fixtures, a gluing station, a folding station, and robot handled tools. Material and components are presented to the robots via conveyors and carriers placed in a lock-system. The cell is designed for mobility using mobile base plates that either is heavy enough to be placed directly on the floor or, as in the case of the robot, a base plate which can be secured to the floor by vacuum or air cushions for fast and smooth transports. Figure 1 shows an illustration made to show that each piece of equipment is thought of as being placed on a piece of a puzzle; the different parts are mobile, with well defined interfaces between them. Figure 2 shows one of the robots and one fixture mounted on base-plates.

Figure 1. Illustration of the concept of Demonstrator 1.

Figure 2. One of the mobile fixtures and one robot on a base-plate developed in demonstrator 1.

The cell is flexible, using flexible equipment, and designed for reconfigurability. Four different types of controllers are today assembled in the cell. Through technical solutions, supporting mobility and flexibility, the requirements of speed are achieved by e.g. quickly reconfiguring the cell at another location and/or introducing a new product. The I/O communication in the cell is established through wireless communication pads to enable easy reconfiguration. The two robots are equipped with swivels and tool-changers so that they may handle all product variants, and also to make it easy to introduce future product variants that require other tools for handling. The Factory-in-a-Box module at ABB Robotics has been commercially developed and put into operation during December 2006.

4.2 Factory-in-a-Box 2 – Welding with focus on mobility

Factory-in-a-box 2 is developed in collaboration with Pharmadule Emtunga (PHEM). PHEM is a supplier of modular facilities to the off-shore, telecom, and pharmaceutical industries. At present the company is striving to implement the same concept in their manufacturing system as they have in their products, i.e. modularization.

Factory-in-a-box 2 is a semi-automated manufacturing cell, which is used for cutting, bevelling, and welding of carbon steel pipes. All machinery will be fitted into a standard container, which also will contain; fume hood exhaust, lighting, computer terminal.

Factory-in-a-Box 2 can perform the following operations;

- Cutting and seam preparation of pipes
- Manual welding when fitting the pipes together
- Orbital welding, as shown in figure 3
- Flexible fixture to handle the different pipes, as shown in figure 4

Orbital welding of carbon steel pipes with straight-angle chamfers has never been done before within the company.

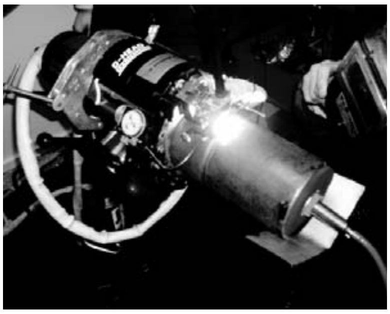


Figure 3. An orbital welding equipment.

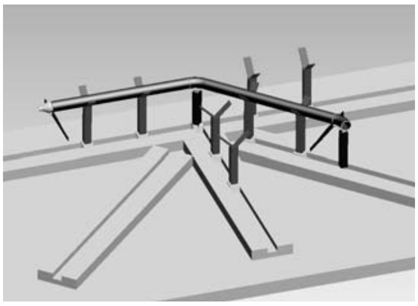


Figure 4. Illustration of the flexible fixtures developed in demonstrator 2.

PHEM has had a business concept in mind when developing the Factory-in-a-Box 2. This kind of mobile orbital welding equipment could make a great asset for a lot of different construction and installation sites where carbon steel pipes are being welded. The Factory-in-a-Box can be leased out to PHEM by suppliers or be owned by PHEM. The final commercial solution will be decided at a later time when the concept has been fully developed.

4.3 Factory-in-a-Box 3 – Foundry with focus on mobility

To ensure applicability of the Factory-in-a-Box project, a third example of a Factory-in-a-Box concept has been developed and demonstrated within Swedish foundry industry in cooperation with the Swedish Foundry Association and Varnäsföretagen. To maintain its competitiveness a foundry must meet increasing demands for efficient production and a good working environment.

Varnäsföretagen AB is a Swedish foundry that produces sand-casted aluminium goods. The company has modern and highly operational facilities for both casting and subsequent machining. However, the process step between deburring and grinding has been neglected, as in many other foundries. This middle step between the casting and machining is important, but often performed manually using handheld tools. This type of work has a tradition of being ergonomically not suitable and is generally also inefficient. This was the motivation for Varnäsföretagen to take part in the Factory-in-a-Box project. The aim of the third demonstrator was, eller om man ska köra konsekvent: has been to create an automatic solution for deburring of casted products.

During 2006 a pre-study at the company has been performed which included a mapping of the production process and a study of the products that are being produced. When the results of the study were analyzed it was concluded that not all operations could be performed within an automatic solution. The ideal solution would be to use several small and cheap standardized automation cells that could perform a couple of operations each. This project focused on developing one cell that can perform a set of operations that can be used for some of the deburring work. This first cell could for example be used to saw off larger pieces from the casting followed by some milling and grinding. These operations are today a burden for the persons doing the manual work.

The final concept that has been developed includes a robot mounted inside a container. The robot is placed on a flexible beam-system that can be moved in and out of the container. A flexible fixture has been proposed that should be able to handle all the products that are supposed to be machined in the cell. The concept has been partly realized in a lab environment, and whether it will be converted into an operational cell is dependent on the company and circumstances outside the research projects control.

4.4 Factory-in-a-Box 4 – Functional sales with focus on flexibility

Factory-in-a-Box 4 has been developed in association with FlexLink Systems. FlexLink's focus is automation of production flow within the following processes: assembly, filling, machining, and packaging. FlexLink will, in this project, use their Dynamic Assembly System (DAS) concept in order to demonstrate the principles of the Factory-in-Box-project – i.e. of flexibility, mobility, and speed – and in a real customer case.

During 2005 and 2006, a pre-study at the company has been performed which has included an investigation of "functional sales" and a mapping of interested companies for this commercial solution. A number of companies have been contacted (over 40 different companies) and possible candidates for a FlexLink and Factory-in-a-Box commercial application where identified.

After identifying one company that was interested, the project has participated in a quoting phase. No order has been placed so far. The Factory-in-a-Box 4 project has investigated the concept of functional sales of manufacturing capacity. The project has shown that industry is interested in this concept, but the circumstances and details in the offering may have to be changed. This demonstrator has shown the importance of further investigating how the sales operations should be designed in a mobile manufacturing system context.

4.5 Factory-in-a-Box 5 – Manual assembly with focus on mobility

This part of the project is conducted in cooperation with Bombardier Transportation. Bombardier Transportation in Västerås is facing a market where the customers are becoming more and more powerful due to the tough global competition. Many customers have strong wishes that part of the production should be carried out locally in order to create new jobs. Instead of building factories, which will be abandoned as soon as the order is processed, Bombardier and the Factory-in-a-Box project aim to develop mobile production facilities, which can be re-located as soon as the production of an order is finished. This can lead to winning orders on markets otherwise closed, which will create job opportunities not only for foreign countries, but also for Bombardier in Sweden since the main part of the order is produced here.

An analysis has shown that a Factory-in-a-Box, mobile production capacity can be used to reallocate working opportunities in a cost-effective way when the customers demand this as a part of the business deal. The concept contains four parts: a technical solution, a logistics solution, a training solution for the local labour, and a methodology for how to move, install, and put the Factory-in-a-Box into production.

A fully developed Factory-in-a-Box enables a substantial reduction of the resources needed for sharing experience and knowledge compared with a conventional outsourcing strategy. The pre-study has resulted in two technical solutions for moving and housing production capacity abroad: a special container and a modular building solution see figure 5. The choice of final solution will be decided upon including a prognosis on usage frequency, production capacity needed, and the strategically value this solution gives Bombardier.

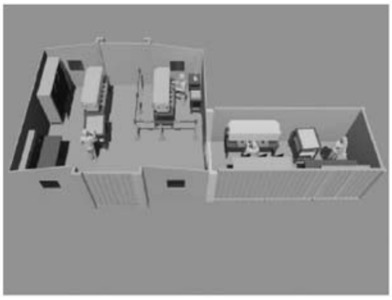


Figure 5. Illustration of a mobile workstation concept of demonstrator 5.

In order to exemplify a valid solution with a cost-analysis, the project has focused on a technical solution for the assembly of High-Voltage boxes demanding a capacity of 1-2 boxes a week. Even using this delimitation, the pre-study indicates that the Factory-in-a-Box easily can be modified to handle a much higher production volume. It also indicates that the same kind of solution can be useful for other Bombardier products. The pre-study has also resulted in a checklist to be used by the person responsible for the installation of the Factory-in-a-Box in order to shorten the start-up time at the customer.

Besides being used to win orders, the project has shown that the Factory-in-a-Box concept provides a substantial potential for cost-cut in the ongoing production at Bombardier.

5 DISCUSSION: FACTORY-IN-A-BOX AS A PRODUCT SERVICE SOLUTION?

The objective of this paper is to discuss the possibility of realizing a Product Service System using the results from the Factory-in-a-Box project. Roy describes the concept of sustainable product-service systems as the idea of cleaner production, eco-design and design for the environment [6]. He claims that the concept goes beyond the environmental optimization of products and processes and requires radical and creative thinking to reduce environmental impacts by a factor of between four and 20 times while maintaining an acceptable quality of service. Further, four main types of product service system that contribute to sustainability by reducing the total quantity of materials and energy required per unit of service are summarized and outlined in the paper [6];

- **Result services** - (sometimes called demand services or service products) aim to reduce the material intensity of existing systems by selling a ‘result’ instead of a product—for example selling a ‘clean clothes’ service rather than a washing machine. The service provider typically takes responsibility for supplying, maintaining, taking back and recycling all physical aspects of the system.
- **Shared utilization services** - (sometimes called product use services or community products) aim to increase utilization of the material parts of a system by sharing the products required.
- **Product-life extension services** - (sometimes called duration products) aim to substantially increase the useful life of products or materials through maintenance, repair, reuse and recycling, thus reducing the amount of energy and resources required to provide a given function.
- **Demand side management** - (sometimes called least-cost planning or integrated resource management) evolved into the idea of considering the end-use service that electricity buyers wanted - illumination, cooling, thermal comfort, etc. - and working out the least-cost method of supplying it.

The Factory-in-a-Box concept may be one innovative way of exploring and realizing at least three of the above mentioned types of PSS. Several of the above product service concepts have been explored in the Factory-in-a-Box research project. The research project have explored and analyzed the challenges and also in several cases developed solutions helping to realize a PSS concept. Support in implementing PSS can be found by analyzing the practical and industrial initiatives in the research project described in this paper.

One example of useful results and similar thinking between PSS and Factory-in-a-Box is the objective in the Factory-in-a-Box project to provide solutions for leasing or renting a production module. This idea of “functional sales” was explored in demonstrator 4 and has shown potential for the future, especially from a SME point of view which may lack capital and would be interested in paying e.g. per produced items instead of investing in equipment. A Factory-in-a-Box company providing a production module for rent or lease is a service provider taking responsibility for supplying, maintaining, taking back and recycling all physical aspects of the system described above as “result service” in a PSS.

Shared utilization service is another aspect that could be realized by a Factory-in-a-Box concept. A production module that can be transported and placed at a central location and then used by e.g. a community, needs to be mobile. The ability to move and reconfigure equipment has been explored in the Factory-in-a-Box project in demonstrators 2, 3, and 5. Especially the Bombardier demonstrator provide results e.g. technical solutions for moving and housing production capacity which can be used in order to realize a shared utilization product service system.

Finally, a product-life extension service might be realized by a mobile maintenance module. The concept “Maintenance-in-a-Box” is a visionary maintenance application that is mobile and possible to ship to any locations where maintenance needs to be performed [7]. This could be a Factory-in-a-Box purely dedicated for mobile assistance in maintenance for production systems, providing solutions for availability and mobility of flexible maintenance capacity, possibly increasing the useful life of products and/or production equipment.

6 CONCLUSION

The uncertainty in markets and rapid introduction of new products has created a growing need for flexible, reconfigurable, and responsive manufacturing systems. On top of these challenges, there is a gigantic need within the manufacturing industry for technology and strategies that will reduce CO2 emissions globally. One possible solution to the above scenario can be the concept of product service system (PSS) which can be described as “a system of products, services, supporting networks and infrastructure that is designed to be competitive, to satisfy customer needs and to have a lower environmental impact than traditional business models”.

In January 2005, the Swedish Foundation for Strategic Research started a research project called “Factory-in-a-Box”. The Factory-in-a-Box concept has the key characteristic of a modular production unit that is flexible, mobile and quick to ramp-up. The objective of this paper has been to discuss the possibility of realizing a PSS using the results from the Factory-in-a-Box project.

The discussion based on Roys summary and classification into four main types of PSS and the relation to the Factory-in-a-Box project shows much similarities between the two concepts. The conclusion is that the Factory-in-a-Box can help to realize and exemplify result services, shared utilization services, and product-life extension services.

7 ACKNOWLEDGMENTS

The Factory-in-a-Box research project has been funded by the Swedish Foundation for Strategic Research through the ProViking program.

8 REFERENCES

[1] Molinari, T. L., Pierpaoli, F., Urbani, A., and Jovane, F., New Frontiers for Manufacturing in Mass Customization, CIRP Journal of Manufacturing Systems, Vol. 33, No. 3, 2004.

[2] Jackson, M., An Analysis of Flexible and Reconfigurable Production Systems – An Approach to a Holistic Method for the Development of Flexibility and Reconfigurability, Linköping, Studies in Science and Technology, Dissertation No. 640, Linköping, Sweden, 2000.

[3] Mont, O., Clarifying the concept of product-service system, Journal of Cleaner Production 10 (2002) (3), pp. 237–245.

[4] M.J. Goedkoop, C.J.G. van Halen, H.R.M. te Riele and P.J.M. Rommens, Product service systems—Ecological and economic basics, Pi!MC, Stoorm C.S. and PRé Consultants, The Hague, Den Bosch and Amersfoort (1999).

[5] Yin, R. K. (1994). Case Study Research - Design and Methods. 2nd edition. USA, SAGE Publications.

[6] Roy, R., Sustainable product-service systems, Futures, Volume 32, Issues 3-4, April 2000

[7] Bengtsson, M., Elfving, S., and Jackson, M., (2006), The Factory-in-a-Box Concept and its Maintenance Application, In Proceedings of the 19th Congress of Condition Monitoring and Diagnostic Engineering Management (COMADEM), Ed. Kumar, U., Parida, A., and Rao, R., Luleå, Sweden

Electro-Physical and Chemical Processes

Planning the use of high-power excimer laser for psoriasis treatment

Hinnerk Hagenah[1], Michael Schmidt[2] and Florian Klämpfl[2]

[1] Chair of Manufacturing Technology, Department of Mechanical Engineering, School of Engineering, Friedrich-Alexander University of Erlangen-Nuremberg, Germany

[2] Bayerisches Laserzentrum GmbH, Erlangen, Germany

Abstract

American and European statistics have shown that 1-2% of the human population is affected by the skin disease psoriasis. Recent research reports promising treatment results when irradiating skin areas affected by psoriasis with high powered excimer lasers with a wavelength of 308 nm. In order to apply the necessary high energy dose without hurting healthy parts of the skin new approaches regarding the system technology must be considered. The aim of the current research project is the development of a sensor-based, automated laser treatment system for psoriasis.

In this paper the needed system technology to provide an automated approach to the irradiation will be presented. This consists of the laser, a sensor to detect infected areas a beam guiding system and a computer to plan the irradiation.

The focus of the paper will be on the algorithms used to cope with the difficulties of irradiating irregularly shaped areas on curved surfaces with a predefined energy level using a pulsed laser. As patients prefer the treatment to take as little time as possible the distribution of laser pulses on the surface to achieve the given energy level in every point of the surface has to be calculated within a limited time frame. We are using simulated annealing to optimize the irradiation and minimize the error. Within the paper the behavior of the optimization algorithm will be described thoroughly. Different influences on the planning result will be identified and discussed.

Keywords:

automated irradiation, planning, optimization

1 INTRODUCTION

Psoriasis is an inflammatory skin disease which is mainly caused by genetic factors. American and European statistics have shown that 1-2 % of the human population with white skin is affected by psoriasis [1]. According to current medical knowledge, psoriasis is neither curable nor infectious. However, since its symptoms - the inflamed skin itches and weeps - put severe strain on patients' lives, relief is highly desirable even if it is only temporary. It is already known that uv-radiation with a wavelength within the therapeutic window which ranges from 300 nm to 320 nm reduces the symptoms and heals the affected skin for some months [2]. Recent research shows very good treatment results and a longer time without recurrence of symptoms when irradiating the affected skin with a XeCl excimer laser with a wavelength of 308 nm [3, 4]. Using this laser has several advantages: it is industrially available and its wavelength is almost at the optimum of 311 nm. Furthermore, it is possible to irradiate only those areas of the skin which are affected by psoriasis which allows applying higher energy doses because the minimal erythema dose (MED) for skin affected by psoriasis is greater than for healthy skin [5].

However, the current approach of using a hand guided laser to apply the laser energy to the skin has some disadvantages. First, it is hard to deliver the laser energy exactly. As a consequence, only low energy doses can be applied to avoid hurting the patient. Using a low energy dose also requires the use of a laser with little power so the treatment sessions take a lot of time. This leads to the second disadvantage. Because it is labor-intensive the laser irradiation of the diseased skin with hand guided delivery causes considerable medical costs.

All these disadvantages can be overcome by using a sensor based and automated laser irradiation system to carry out the treatment [6]. It allows a much more exact application of the required energy dose than a hand guided system. Thus the energy can be increased and a laser with more power can be used. Automation also means that the required manpower and thus the medical costs for the treatment are reduced. The disadvantage of this approach is that the necessary system technology is more complicated.

Figure 1 shows an image of the system technology developed by the partners to carry out a sensor based and automated laser treatment. Using the data acquired by a sensor consisting of a camera and led light sources, the "Photometric Stereo Principle" [7] - also called "Shape from Shading" - is used to make a three dimensional model of that part of the patient's body which is to be treated. Furthermore, a color image of the treatment area is taken. Using a k-means clustering algorithm, the diseased skin areas are detected [8].

The beam of a XeCl excimer laser deflected by a scanner and controlled by an ordinary personal computer is used to apply the radiation to the affected skin areas meeting the medical requirements as well as possible. The average distance between the patient and the laser optics is 1000 mm. The maximum size of the treatment zone is 400 mm by 300 mm. The excimer laser being used in this project is a pulsed laser with a pulse frequency of 1 kHz and an average pulse energy of 8 mJ. The spot size of the laser on the surface of the human body will be in the range of 5*5 mm² to 10*10 mm².

Figure 1: System technology of the sensor based automated excimer laser treatment system for inflammatory skin diseases like psoriasis

One of the goals of the current project is to reach the highest energy dose permissible in one treatment to achieve good therapy results. Therefore, it is required that the applied energy meets the medical requirements for the treated skin area in the best possible way. This cannot be achieved by a hand guided laser system without errors regarding the applied energy dose. Even with an automated system as we are developing it, the irradiation is a challenging task. The skin area which must be treated is an irregularly bordered shape on an irregularly curved three dimensional surface. Furthermore, the energy profile of an excimer laser spot is not equally distributed even though a beam homogenizer based on micro lens arrays [9] is used to improve the beam quality. Our approach to solve the irradiation problem is to split it into two parts: pulse distribution planning on the one hand and beam track planning on the other. The main focus of this paper will be on pulse distribution planning.

2 ALGORITHMIC CHALLENGES

The planning of the laser pulse distribution on the surface can be considered to be an optimization problem. The coordinates of each pulse and its energy are the variables to optimize. When considering an optimization problem, the number of variables, i.e. the dimension of the solution space, is an important point. The maximum size of the area we consider is 300*400 mm². We assume that we have to irradiate 50 % of the potential treatment area. This is 600 cm². Furthermore, research with manual irradiation has shown that an average energy dose of 1 J/cm² is necessary for the treatment, so we have to set 75,000 pulses to apply the necessary energy in this case.

In theory, the coordinates and the energy of the pulses are the variables to optimize. This results in an optimization problem with 225,000 variables. Even the restriction of the system technology having a fixed energy for all pulses does not reduce the variable count significantly, there are still 150,000 laser pulse coordinates left to determine and to optimize. Optimization problems of this size require sophisticated heuristic algorithms to find a usable solution within reasonable time. The planning must be done during the treatment session of the patient. Furthermore, to keep costs of the irradiation system low, it is desirable to do the planning on a common PC.

3 PULSE DISTRIBUTION PLANNING

3.1 The optimization task

Before presenting the optimization algorithms in detail, a mathematical definition of the planning task will be done. First, we define G according to equation 1

$$G = \iint_A f\left(E_{target}(x,y), E_{Plan}\right) dA \qquad (1)$$

as a measurement for the quality of the current irradiation plan. $E_{target}(x,y)$ is the energy dose in the point (x, y) given by medical requirements. E_{Plan} is the currently planned energy at the point (x,y) which depends on the positions (x_n, y_n) of each laser pulse n and the energy of these pulses E_n at the point (x,y). Depending on E_{Plan} and E_{target} an error function f(...) is evaluted which is integrated over the whole considered surface A to get G as the measurement of the quality of the solution for the pulse distribution task. This function will be used later.

When working with numerical algorithms on machines with limited calculation precision, we cannot use continuous functions but have to discretize the integrals in equation 1 and change them into sums. This results in

$$G = \sum_x \sum_y f\left(E_{target}(x,y), E_{Plan}\right) \qquad (2)$$

In the remainder of this paper, we will assume that the energy of all laser pulses is constant. This is a technical restriction of the laser used. The pulse energy cannot be reduced below a limit of around 50-70 % of the maximum energy because of fundamental laser limitations.

The error function f has not been considered in detail so far. Usual definitions of an error function are the absolute difference:

$$f_1(E_{target}, E_{Plan}) = \left|E_{target} - E_{Plan}\right| \qquad (3)$$

or the square of the difference:

$$f_2(E_{target}, E_{Plan}) = (E_{target} - E_{Plan})^2 \qquad (4)$$

However, the error function is a possibility to take care of medical restrictions and requirements. It is possible to use the error function to implement a hard upper limit of the irradiation dose by defining an error function which returns infinity if E_{Plan} is greater than E_{target} plus a certain constant:

$$f_3(E_{target}, E_{Plan}) = \begin{cases} \left|E_{target} - E_{Plan}\right| & E_{target} + C \geq E_{Plan} \\ \infty & else \end{cases}$$

This error function ensures that the irradiation dose never exceeds the given limit in any point of the irradiated surface. The idea of introducing a lower limit as well by modifying the condition for the upper branch to

$$E_{target} + C_1 \geq E_{Plan} \geq E_{Target} - C_2$$

does not work. The optimization algorithms did not converge. Further medical research might result in the requirement of a different error function such as a bi-quadratic one that is shown in equation 5 and has been used for the remainder of the paper.

$$f_4(E_{target}, E_{Plan}) = \left(E_{target} - E_{Plan}\right)^4 \quad (5)$$

3.2 Calculation of an initial pulse distribution

When solving optimization problems, the results can be improved significantly regarding both the time needed for computation and the quality of the result in the case of heuristic algorithms by using a good starting point. Therefore much effort is spent in achieving a good starting point for the pulse distribution optimization.

The general idea of doing the initial planning is walking in a defined manner through the parameter space, in this case the surface, and checking if setting a pulse at a certain place improves the solution of the optimization problem. The simplest approach would be to walk across the surface line wise and column wise with a fixed step size. However, the initial distributions determined by this approach were rather bad.

Figure 2 shows the currently implemented approach. Depending on the size of a pulse and the average slope of the surface, an optimal distance d_{opt} is calculated. d_{opt} is divided by the estimated number of pulse layers n required to achieve the desired energy dose. This vector $d_{opt,n}$ is split into its two components $d_{opt,x}$ and $d_{opt,y}$. The algorithm then walks across the surface in multiple passes and tests pulses with a step size using the x- and y- components of d_{opt}. The starting point of each pass is moved by $d_{opt,n}$ against the previous pass. This procedure is repeated until a complete walk over the surface places no more pulses.

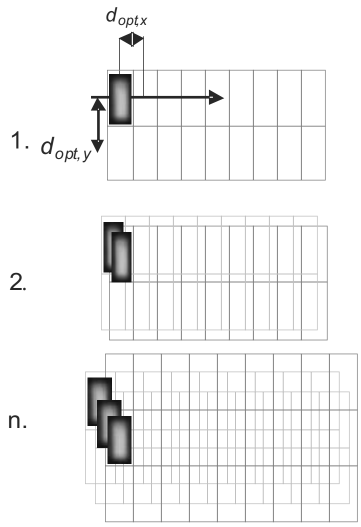


Figure 2: Current way of walking across the surface to do the initial pulse distribution planning}

3.3 Optimization of the pulse distribution

The initial pulse distribution already produces good results with respect to the desired energy dose distribution at the regions of interest. However, these results can be improved further to achieve better treatment results. From literature a lot of different optimization algorithms like hill climbing [10], genetic algorithms [11], swarm intelligence based ones [12] or simulated annealing [13] are known which could be used here. Choosing the right algorithm for a certain task requires first an analysis of the challenge to overcome.

Having a closer look at our optimization task, one quickly comes to the conclusion that its main property is what is called a good conditioning. When changing one variable of the solution, the result will not change much. Furthermore, we assume that the pulse distribution planning is an NP-equivalent problem because it is similar to NP-equivalent problems like vertex cover or bin-packing} [14]. We do expect that neither an exact algorithm nor an approximate algorithm for the pulse distribution planning exists. So a heuristic algorithm must be used which can benefit from the good conditioning of the optimization task and which can deal with the larger number of variables and with the probably existing local extremes of the planning problem. Simulated annealing [13] (SA) is an algorithm which fulfills these requirements and which has been already applied to tumor irradiation planning [15].

One important aspect of an SA implementation is the parameterization e.g. at which point is the cooling started, how fast is the cooling done and when the optimization algorithm is stopped. Currently, we do not change all variables in one cooling step, but only the variables, i.e. coordinates, of one pulse. Depending on the SA parameterization this change of the coordinates of a pulse is accepted or not. Our current implementation does 200 cooling steps per pulse. At every cooling iteration the parameters of all pulses are varied once. This approach proofed to produce good results within reasonable time.

4 FAST AND EFFICIENT OPTICAL SIMULATION

4.1 Optical simulation

The calculation of the function E_{Plan} used in equations 1 and 2 has not been considered in detail so far. As mentioned above, it returns the energy distribution on the surface depending on a given pulse distribution. Thus calculating E_{Plan} can be done by carrying out an optical simulation. Two well-known types of optical simulations [16, 17] are ray tracing and solving the Maxwell wave propagation equations. However, these approaches have a major disadvantage. They are very expensive regarding computational complexity so the simulation times are rather high. The pulse distribution planning requires, however, that thousands of different pulse distribution configurations can be tested within a time frame of a few seconds. The reason for this is that the pulse distribution planning must be done during the treatment of the patient. Thus a new and efficient approach for optical simulation is required to obtain E_{Plan} during the therapy process.

4.2 Fast bitmap-based empirical optical simulation

The idea which we implemented is based on modern graphics rendering. The simulation process works basically with bitmap transformations like distortion or scaling. It is assumed that each optical element does such a bitmap transformation. A bitmap from the object plane is transformed into a bitmap at the image plane. This bitmap is taken as the object plane bitmap of the next optical element. Thus the propagation of the beam profile through the optical system and its projection onto the treated surface can be considered as a series of bitmap transformations. To model changes in the optical path like different angles of mirrors or different working distances, the bitmap transformations are parameterizable. This is an empirical approach but it can be implemented very efficiently and allows a very fast calculation of the resulting energy distribution on the surface. Besides this, it permits a considerable amount of tuning and only those effects which are important for the current application need to be modeled. To match reality, the relevant effects need to be identified by experiments or simulations based on exact calculations of physical effects e.g. solving the Maxwell equations. To speed up the calculation of E_{Plan} even more, possible combinations of bitmap transformation parameters are collected in so-called \emph{parameter sets} and transformations with frequently used parameter sets are precalculated and stored for future use. Furthermore, transformations with a parameter set close to existing ones are not calculated. Instead, a simple interpolation between the closest available transformation results is carried out.

In this first step several simplifications were made. First of all, all wave propagation effects are ignored. This does not matter in our case because we are working with rather large optical elements. Further research will show how far this empirical method to calculate E_{Plan} fits reality and how we can improve the algorithms to arrive at improved results. Currently, the simulation algorithms and implementation are closely tied system technology and beam forming elements that were used here but it should be possible to extend this approach to other optical systems as well.

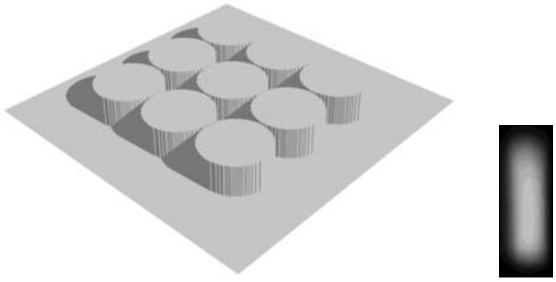
5 RESULTS

To test the pulse distribution planning algorithms, we use a set of test patterns which consist of real and artifical energy distributions to achieve. Figure 3 shows two examples of these test patterns with the desired energy dose at each point encoded with gray tones. Black means no irradiation, white irradiate with a normalized dose. The whole planning process works with normalized values because the energy dose depends on the skin of the patient. Using normalized energy doses is also what we do in the pictures visualizing energy distributions which are presented below.

(a) (b)

Figure 3 : (a) Artificial test pattern (b) Real test pattern.

The energy distributions can also be visualized in 3D as shown in figure 4 (a). The higher the drawn surface is, the higher the applied energy is at this point. The results presented below are created using the artifical test pattern. The used laser beam profile is shown in figure 4 (b).



(a) (b)

Figure 4: (a) 3D visualization of E_{target}. (b) Pulse profile used for irradiation planning (magnified)

As can be seen from Figure 5 the energy distribution after determining the initial distribution already looks quite good.

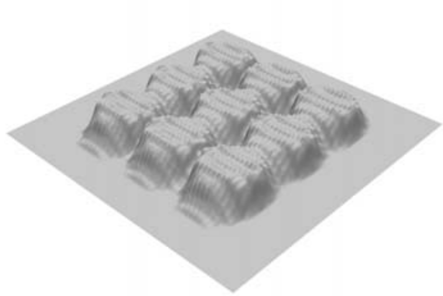


Figure 5: Energy distribution E_{Plan} for the artiicial test pattern using the initial solution.

The resulting planned energy distribution after the optimization is depicted in figure 6.

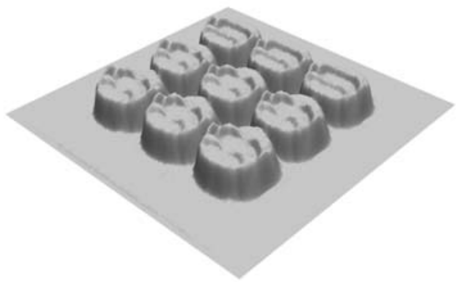


Figure 6 : Energy distribution E_{Plan} for the artificial test pattern after the optimization.

The diagrams in figures 7 and 8 show the development of the error while doing the planning. Figure 7 includes the evolution of the error and the total energy over the entire planning period, while figure 8 shows only the initial planning and the start of the optimization. As can be seen the initial distribution of the pulses is finished after 300 ms. No more pulses are added after that time, the total energy does not increase any more. The absolute error is 10 J at this point. Now the optimization through SA starts and the error decreases until it reaches 3 J after approximately 15 s. Taking into account that the total planned energy inserted by the initial distribution is too high and that there are no means of removing pulses this result is satisfying. Ongoing research will enable the removal and addition of pulses after the initial irradiation distribution has been determined by the simulated annealing algorithm.

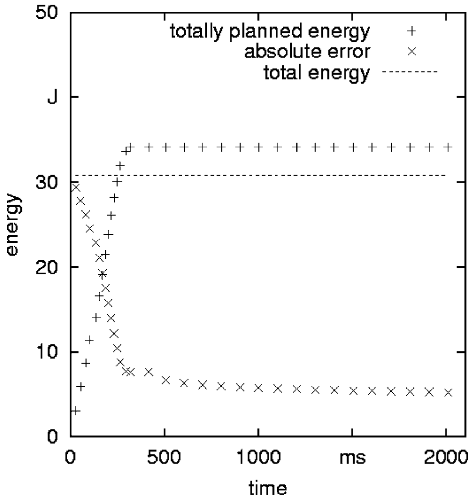


Figure 7: Development of the error during the optimization of the pulse distribution.

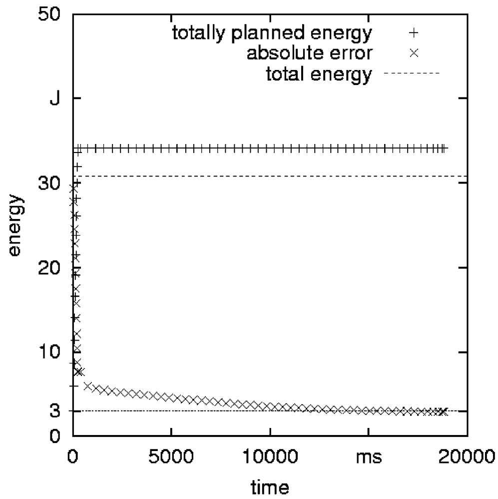


Figure 8: Development of the error during the calculation of the initial laser pulse distribution.

As mentioned above, we have considered the possibility of using different error functions when calculating G as defined in equation 2. Figures 9 and 10 show the resulting energy distribution when using equation 5 as quality measurement.

It can be seen that the algorithm converges and finds a solution. Nevertheless, the absolute error is higher than for the previous example. This can be explained, as the bi-quadratic difference was optimized in this case. However, the absolute error can be considered as a measurement of the convergence. Thus, it is possible to compare the results with the previous ones. One can see that a stable absolute error value is reached faster at a much higher level. This might be caused by the fact that this error function lowers the quality of the conditioning of the optimization task.

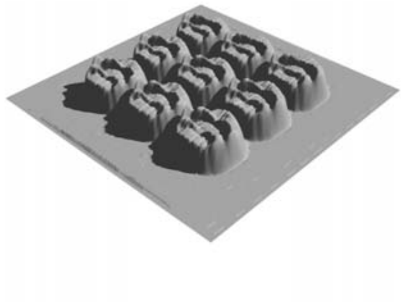


Figure 9: Energy distribution resulting from an optimization using the quality criterion given in equation 5.

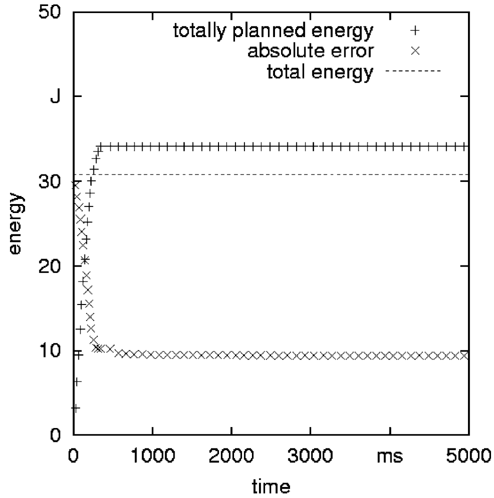


Figure 10: Error and total energy when using the quality criterion as given in equation 5.

6 CONCLUSION AND OUTLOOK

The results presented here show that it is possible to carry out an automated treatment of skin diseases like psoriasis with the system technology and computation power available in our days. It has been shown that the computation power of a standard PC is sufficient to detect which areas of the skin are affected by the disease, to carry out the irradiation planning and control the treatment.

Further medical research will show if the automated treatment of psoriasis with an uv-laser can improve the therapy results by applying the energy dose in a more controllable manner.

The automated irradiation and treatment may give new scientific insights because it allows a well defined irradiation of large skin areas and because it permits the study and analysis of different treatment patterns within one treatment session. Furthermore, the data acquired by the sensor can be stored in a database to enable easy access to the history of the disease of the patient. These data could be used to make it easier to find the optimum dose and irradiation strategy for the best therapy results.

The algorithms and software modules will be further improved. The current simulated annealing implementation does not add or remove pulses, the pulses are only moved. We believe that the current result can be further improved if after every step of simulated annealing another run is done which removes or adds new pulses to keep the global error given in equation 6 as small as possible.

$$e_{global} = f\left(\iint_A E_{t\arg et} dA, \iint_A E_{Plan} dA \right) \quad (6)$$

While the current research and development concentrates on psoriasis, we think that it is possible to apply the system technology presented here to the therapy of other skin diseases like vitiligo which is also known as leukoderma and can also be treated by uv light as emitted by an excimer laser with 308 nm wavelength. Our algorithms, hard- and software should be able to deal with this disease as well. The detection algorithms work without a priori knowledge so they could probably easily be parameterized to detect other kinds of skin diseases as well. The irradiation algorithms could also be easily adapted to the needs of irradiating other diseases by changing the energy dose, the error function of the optimization algorithm and the path planning constraints.

There are several other laser processes which might require the processing of a 2.5-D surface with predefined energy and power density. The algorithms and software presented here can be used in these cases as well. These applications can be laser marking or structuring of curved surfaces as well as automated and controlled cleaning of pictures or sculptures.

These applications might require an even more precise irradiation process than the one we have presented here, but as we mentioned above, there are still ways for further improvement of the pulse distribution planning software.

7 ACKNOWLEDGEMENTS

The authors would like to thank the Bavarian Research Foundation BFS for funding this research within the scope of project 542/02.

8 REFERENCES

[1] Grema, H., Grefe, B., Pfisterer, S., Raulin, C., 2002, Aktueller Stand der Psoriasistherapie, Derm PraktischeDermatologie 8

[2] Green, C., Ferguson, J., Lakshmipathi, T., Johnson, B., 1988, 311 nm UVB phototherapy - an effective treatment for psoriasis, Br J Dermatol/12, 691-696

[3] Farkas, A., Kemeny, L., 2006, Applications of the 308 nm excimer laser in dermatology, Laser Physics 16 (5), 876-883

[4] Gerber, W., Arheilger, B., Ha, T., Hermann, J., Ockenfels, H., 2003, Ultraviolet B 308-nm excimer laser treatment of psoriasis: a new phototherapeutic approach, Br J Dermatol. 149 (6), 1250-8

[5] Speight, E., Farr, P., 1994, Erythemal and therapeutic response of psoriasis to PUVA using high-dose UVA, Br J Dermatol , 667-672

[6] Wolfsgruber, F., Rühl, C., Kaminiski, J., Kraus, L., Häusler, G., Lampalzer, R., Häußler, E.-B., Kaudewitz, P., Klämpfl, F., Görtler, A., 2005, A simple and efficient optical 3D-Sensor based on Photometric Stereo (UV-Laser Therapy), 5. International Workshop on Automatic Processing of Fringe Patterns, 702-706

[7] Horn, B. K. P., Brooks, M. J., 1989, Shape from Shading

[8] Cucchiara, R., Grana, C., Piccardi, M., 2001, Iterative fuzzy clustering for detecting regions of interest in skin lesions, Settimo Congresso dell'Associazione Italiana per l'Intelligenza Artificiale, AI*IA Notizie, 31-38

[9] Stark, M., Völkl, R., Zimmermann, M., 2005, Strahlhomogenisierung auf der Basis von Mikrolinsenarrays, Tagungsband LEF 2005: Laser in der Elektronikproduktion & Feinwerktechnik 221-235

[10] Spall, J. C., 2003, Introduction to Stochastic Search and Optimization: Estimation, Simulation, and Control

[11] Koza, J. R., 1992, Genetic Programming: On the Programming of Computers by Means of Natural Selection

[12] Beni, G., Wang, J., 1989, Swarm intelligence in cellular robotics systems, Proceeding of NATO Advanced Workshop on Robots and Biological System

[13] Kirkpatrick, S., Gelatt, C. D., Vecchi, M. P., 1983, Optimization by Simulated Annealing, Science, Number 4598/220, 671-680

[14] Rao, R. L., Iyengar, S. S., 1994, A stochastic approach to the bin-packing problem, SAC '94: Proceedings of the 1994 ACM symposium on Applied computing, 261-265

[15] Aubry, J.-F., Beaulieu, F., Sévigny, C., Beaulieu, L., Tremblay, D, 2006, Multiobjective optimization with a modified simulated annealing algorithm for external beam radiotherapy treatment planning Medical Physics 33 (12), 4455-4773

[16] N. N., 2001, ZEMAX, Optical Design Program, User's Guide

[17] Litfin, G., 2001, Technische Optik in der Praxis

Optimization Strategies of Laser Hardening of Hypo-eutectoid Steel

Giovanni Tani[1], Leonardo Orazi[2], Alessandro Fortunato[1], Giampaolo Campana[1], Alessandro Ascari[1] and Gabriele Cuccolini[2]

[1]Department of Mechanical Construction Engineering – University of Bologna, Bologna, Italy

[2]Department of Science and Method for Engineering– University of Modena and Reggio Emilia, Reggio Emilia, Italy

Abstract

The interest towards LASER hardening of steels has been increasing since the last few years due to its undoubted advantages. The main drawback affecting this manufacturing technology is the tempering effect induced when multiple passes on the same surface must be carried out. In order to minimize the softening effect due to tempering and to speed up the process a numerical model for the simulation of the treatment is proposed. This model is able to detect the optimal LASER path trajectory according to the source parameters and the scanning velocity, and it is able to predict the resulting microstructures and the relating hardness. Some examples on an hypo-eutectoid steel are presented together with validation tests.

Keywords:
Laser surface hardening, simulation, numerical model, tempering

1 INTRODUCTION

Laser surface hardening treatment is a manufacturing technology which gained a great industrial interest in the last years [1]. The possibility of integrating the heating source directly on the production line, together with the absence of the quenching medium meet, in fact, the production needs of modern industries. Besides, the versatility of this technology is high if compared to the more traditional surface treatments such as induction or flame hardening, and makes possible to deal with very complex 3D shapes and with very small and confined surfaces. These characteristics are absolutely peculiar of laser surface hardening and they determine the growing exploitation of this technology at the expense of the more traditional ones. The main field of application of laser hardening concerns carbon steels on which the average industrial requests deal with 1 mm penetration depth of martensite. On the other hand when large surfaces have to be treated the laser beam trajectory must be organized in order to cover the whole surface with sequential passes inevitably interact each other causing the tempering of the previously hardened material. According to this, the present paper deals with optimization techniques which make possible to calculate the best laser beam path strategy in order to achieve an evenly distributed martensitic layer of a desired depth on surfaces larger than the laser beam spot. These techniques are based on the exploitation of an original laser hardening simulation software developed by the authors, named Laser Hardening Simulation LHS [2], which also takes into consideration the tempering effects due to a beam pass on the previous ones [3].

The latest developments of the simulation code presented in this paper makes it possible to take into account the effects of the hysteresis for martensite-austenite transformations during the re-heating and the martensite tempering during multiple passes. Moreover, a physical parameter is proposed for the laser trajectory optimization, allowing the optimum scanning trajectories to be predicted in terms of distance between the passes, time occurring between the passes and drawing sequence of the passes. Several experimental tests are proposed on an AISI 1040 (UNI C40) carbon steel plate with a CO_2 laser source in order to show the accuracy of the model.

2 NUMERICAL MODEL

2.1 Heat transfer and microstructures transformations

The first problem that must be faced is the evaluation of the time-dependent temperature distribution *T(t,x,y,z)* through the work-piece due to the heat flux into the material. The solution of this problem is obtained by solving Equation 1 where the coefficient c_p is the specific heat in *J/kgK*, ρ is the material density in *kg/m*^{3}, *k* is the target thermal conductivity in *W/mK* and *q* is the laser heat flux into the material in *W/m*^{3}.

$$c_p \rho \frac{\partial T}{\partial t} = \frac{\partial}{\partial x}\left(k \frac{\partial T}{\partial x}\right) + \frac{\partial}{\partial y}\left(k \frac{\partial T}{\partial y}\right) + \frac{\partial}{\partial z}\left(k \frac{\partial T}{\partial z}\right) + q \quad (1)$$

Equation 2 determines the time-dependent solute distribution through the work-piece being C_υ the solute concentration in the phase υ and D_υ the solute diffusivity in the phase υ. This formulation was already proposed by [4] [5] [6] and most of their notations are here adopted.

$$\frac{\partial C_v}{\partial t} = \frac{\partial}{\partial x}\left(D_v \frac{\partial C_v}{dx}\right) + \frac{\partial}{\partial y}\left(D_v \frac{\partial C_v}{\partial y}\right) + \frac{\partial}{\partial z}\left(D_v \frac{\partial C_v}{\partial z}\right) \quad (2)$$

According to Figure 1, the initial microstructures are taken into account by means of a digitized photomicrograph of the workpiece material by assigning a letter (p) for the pearlite phase, (f) for the ferrite and the asterisk (*) for the intermediate phase pearlite-ferrite.

Carbon homogenization, obtained by means of Equation 2, is performed solely over the pearlite (p) and intermediate pearlite-ferrite (*) cells as proposed in [7].

The ferrite (f) is transformed into austenite when the A_{c3} temperature is reached.

$$L\lambda \le \int_{2}^{t_1} D_0 \, exp\left(-\frac{Q}{RT(t)}\right)dt \tag{3}$$

Moreover, the coarsening of the structures and heating rate is considered in the model by means of the Equation 3.

In Equation 3 L represents the average diameter of the pearlite colonies, λ the average spacing of the lamellae, *Q* the activation energy, D_0 is the diffusion constant, t_1 the time when the eutectoid temperature is reached and t_2 the current time. *R* is the gas constant and *T(t)* the time-dependent temperature.

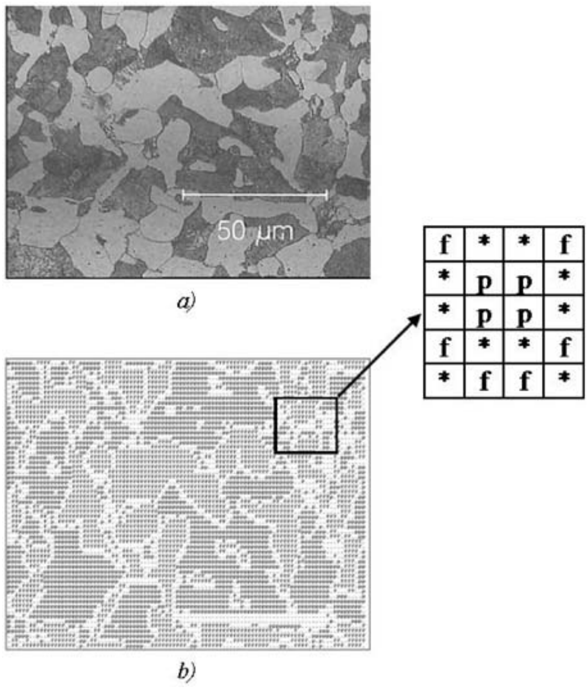


Figure 1: Example a) Digitized photomicrograph, b) Matrix of characters representing the initial microstructure.

Each pearlite or intermediate pearlite-ferrite cell can be partially or completely transformed into austenite according to the heat cycle which is subjected to; the pearlite and the intermediate pearlite-ferrite phases not transformed into austenite will influence the hardness of the cell because during the quenching they doesn't participate to the hardening process. In fact, in the subsequent quenching, only the austenite cells will be transformed into martensite whose hardness depends on the carbon content calculated by means of Equation 2. The hardness obtained by the process can be calculated by means of Equation 4 knowing the volume of martensite and pearlite resulting after quenching.

$$H = f_m H_m + (1 - f_m) H_p \tag{4}$$

f_m being the volume fraction of the martensite and pearlite and H_m, and H_p the hardness of martensite and pearlite respectively.

2.2 The modelling of the tempering

As described in the previous paragraph, LHS is able to predict the resulting micro-structures after quenching, the extension of the treated area and the hardness according to the workpiece material, the laser source parameters, the laser scanning trajectories, the initial type of the micro-structures and their coarsening, while no information can be obtained for the prediction of the hardness of the tempered area when multiple passes have to be performed.

In order to solve these problems, without predicting the resulting structures due to the martensite softening transformation as in [7] and [8] which is very time consuming, two topics have to be faced:

- The re-austenitization of the martensite.
- The non-constant tempering temperature.

The re-austenitization of the martensite is a typical effect which occurs when multiple laser paths have to be used: subsequent laser paths interact with the previous ones causing a partial re-austenitization of the pre-treated structures. This consideration leads to different situations in the re-heated area: areas simply tempered whose hardness depends on the heat cycle and areas re-austenitized and subsequently quenched whose hardness is calculated as described in the previous paragraph adopting Equation 4. A physical parameter, which controls this transition, is proposed and presented as an integral energy force named I_{ma} as in Equation 5:

$$I_{ma} = \int_{t_1}^{t_2} \exp - \frac{Q_{ma}}{RT(t)} dt \tag{5}$$

Q_{ma} being the activation energy for the martensite to austenite transformation, t_1 being the time when the A_{c1} temperature is reached and t_2 the current time.

In particular, a threshold value I_{maT} for this parameter is determined which can be assimilated to a material property: for medium carbon steel this value is set equal to I_{maT}=0.5 *s* for Q_{ma}= 2000 *J/mol*. The previous values for I_{maT} and Q_{ma} are evaluated by means of comparison between the experimental and predicted hardness profile in the case of two laser path trajectories with 50% overlapping spot, 1200 W laser power and 400 *mm/min* scanning velocity.

Equation 5 is also used to take into account the hysteresis phenomena during tempering.

The non-constant tempering temperature leads to a reformulation of the Equation 6 typical for furnace tempering of martensite [9]:

$$\begin{aligned} HV_M &= -74 - 434C - 368Si - 25Mn + 37Ni \\ &-335Mo - 2235V + \frac{10^3}{P_c}(260 + 616C + 321Si \\ &-21Mn - 35Ni - 11Cr + 352Mo + 2354V) \end{aligned} \tag{6}$$

where P_c is expressed in Equation 7. H_a being the activation enthalpy of the micro-structural transformation during tempering.

$$P_c = \left[\frac{1}{T} - \frac{nR}{H_a} \cdot \log\left(\frac{t}{t_0} \right) \right]^{-1} \tag{7}$$

Equations 7-8 can be applied to the tempering processes carried out in furnace where the process temperatures T are constant for the whole process time t, but it cannot be applied in laser hardening. In order to solve the problem a tempering factor parameter τ is proposed that operates when the martensite start temperature M_s is reached. The τ parameter allows the softening of the micro-structures to be calculated as a function of time as presented in Equation 8.

$$\tau_{k+1} = 1 - \frac{\left(Hv^k - Hv^{k+1} \right)}{Hv^k} \tag{8}$$

Hv^k and Hv^{k+1} are the hardness evaluated by means of Equations 6 in correspondence of the two subsequent instant t_k and t_{k+1} during tempering.

The tempering factors τ_{k+1} makes it possible to calculate the hardness Hv_j^k of a generic phase j at the instant t_k starting from the initial hardness Hv_j^0 as presented in Equation 9.

$$Hv^k = Hv_j^0 \prod_{j=0}^{k} \tau_j \tag{9}$$

3 EXPERIMENTAL

3.1 Experimental Setup

All the experimental trials were made exploiting a 3 kW El. En. FAF CO_2 TEM_{01^*} laser source. The varied parameters, together with their respective levels, are summarized in Table1.

LASER Power	1200 *W*, 1750 *W*
LASER beam speed	0.4 *m/min*, 1 *m/min*
Overlapping	8%, 20%, 45%, 60%, 65%
Pass sequence	1-2-3, 1-3-2

Table 1: Process parameters.

The test specimens were obtained on the axis of 65x70x15 *mm* AISI 1040 steel plates which were previously coated with spray graphite in order to maximize the laser power absorption coefficient. The ratio between the extension of the plates and the extension of the hardened zone is high enough in order to consider the heat exchange towards the environment negligible. These trials were made with the aim of validating the predictive capabilities of the simulating software in terms of hardness profile, hardened depth and tempering effect due to overlapping laser beam trajectories. Considering the fact that a hardened layer at least 1 mm thick had to be achieved the matching between laser beam power and scanning speed was set as the following: 1200 *W* - 0.4 *m/min* and 1750 *W* - 1 *m/min*. Figure 2 reports the dimension of the specimen, the position of the laser path into the specimen and the laser scan strategy adopted in Figures 3-8. Figure 8 shows the hardness profile measured in the vertical direction of the third laser pass for a laser power of 1750 *W* and a laser beam velocity of 1 *m/min*. According to that graph the aim of achieving a 1 *mm* thick hardened layer is accomplished for this set of laser parameters. Figures 10-11 show the hardness profiles in the laser beam pass drawing sequence 1-2-3 as outlined in Figure 9. It is evident that the influences can be high with this laser path strategy and that the third laser path also influences the first one if the overlapping is higher than 60%.

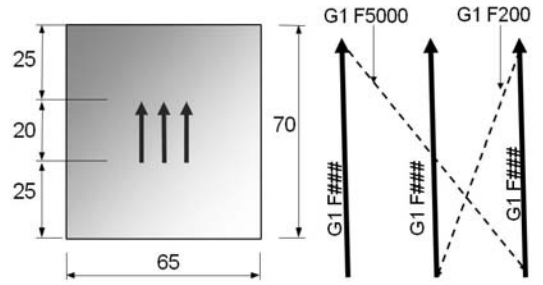


Figure 2: The workpiece dimensions and the part program.

3.2 Experimental results

The experimental results are shown in Figures 3-11. Any hardness profile was measured on a cross section of the specimen normal to the laser beam direction. The measurement line is parallel to the workpiece treated surface at a 150 μm distance. Figures 3-7 show the hardness profiles obtained with a laser beam pass drawing sequence 1-3-2 as proposed in Figure 2.

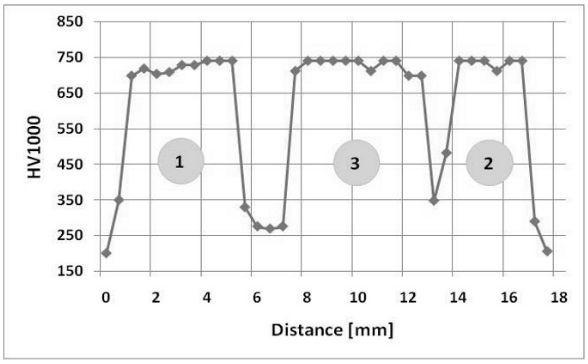


Figure 3: 1750 *W*, 1 *m/min*, 8% overlapping.

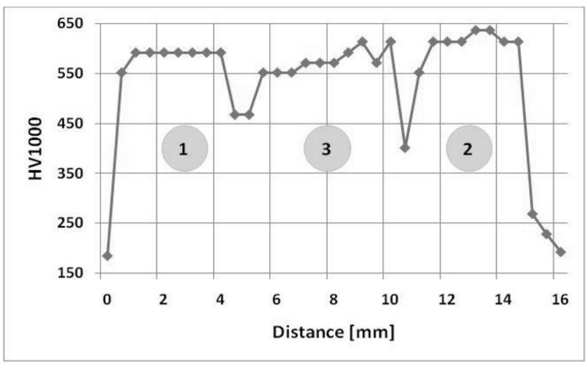


Figure 4: 1750 *W*, 1 *m/min*, 20% overlapping.

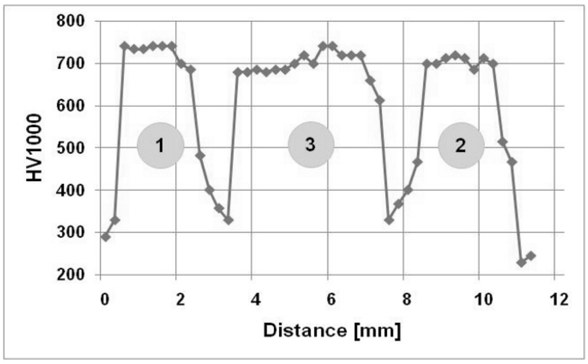


Figure 5: 1750 *W*, 1 *m/min*, 45% overlapping.

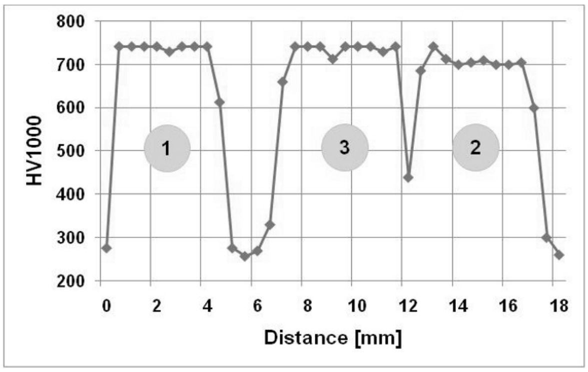


Figure 6: 1200 *W*, 0.4 *m/min*, 8% overlapping.

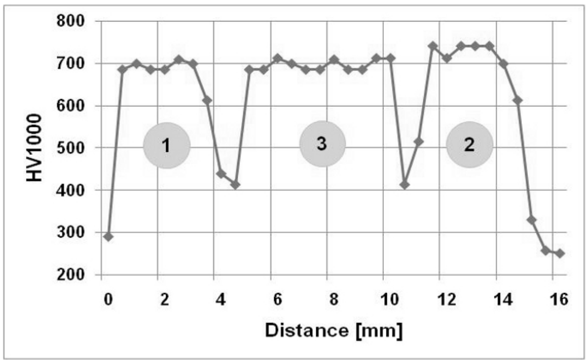


Figure 7: 1200 *W*, 0.4 *m/min*, 20% overlapping.

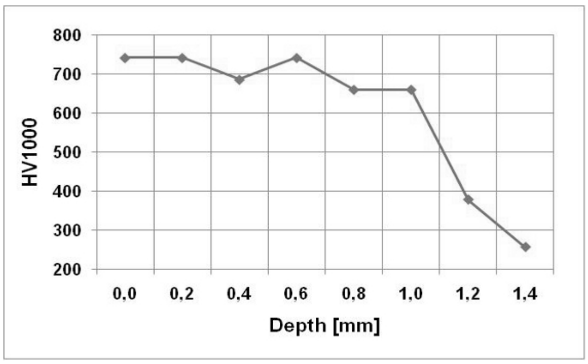


Figure 8: 1750 *W*, 1 *m/min*, 8% overlapping, hardened depth.

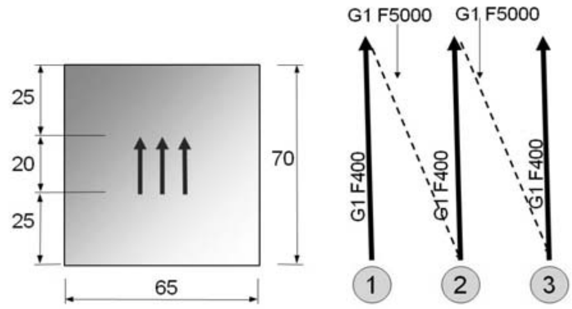


Figure 9: Scanning strategy in three subsequent laser path.

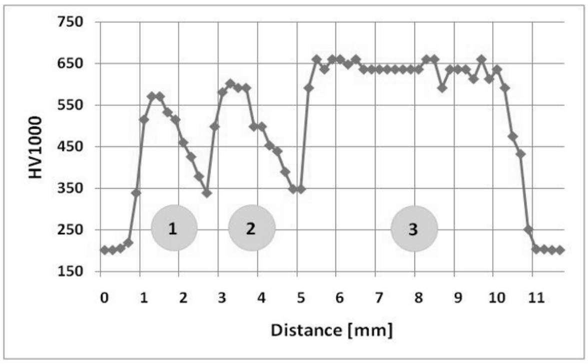


Figure 10: 1200 *W*, 0.4 *m/min*, 60% overlapping.

4 PROCESS OPTIMIZATION BY MEANS OF SIMULATION

In this paragraph, the experimental hardness results are compared to the simulations. The aim is to show the accuracy of the model and, at the same time, to investigate the existence of a physical parameter which drives the choice of the optimal process parameters. Figures 12-15 reports the comparison for a laser power of 1200 *W,* a laser beam velocity of 400 *mm/min* and exploiting the scanning strategy presented in Figure 2. Figure 12 shows the predicted hardness in the first path (dashed line) when the laser is executing the second trajectory after 4.388 *s*.

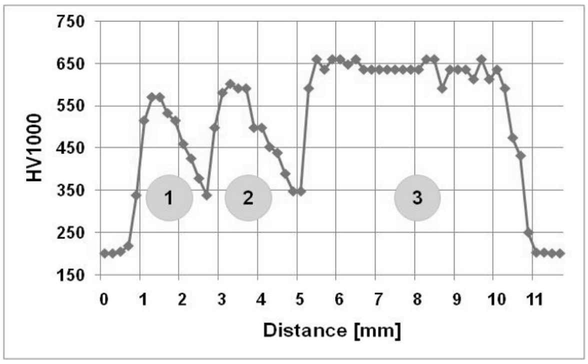


Figure 11: 1200 *W*, 0.4 *m/min*, 65% overlapping.

The temperature in the second path increases due to the interaction with the laser beam (continuous dotted line) and the hardness is that of the base material. After 11.263 *s*, before the third laser path, the quenching leads to martensite formation in the two previous areas, the interaction between the first and the second path is absent and no tempering effect is predicted. The effect of the third trajectory on the previous treated zones is presented in Figure 14 after 15.638 *s* where the tempering, caused by the increasing of the temperature, causes a reduction of the extension of the martensite. After 25.638 *s* the temperature into the bulk is approximately 300 *K* and the process is completed. At this time the results of the simulation can be compared to the experimental ones.

A criterion for the laser trajectory optimization could be obtained by controlling the temperatures differences of two adjacent paths calculated in correspondence of the maximum temperature reached in the last trajectory. For the simulation in Figure 15 in correspondence of the maximum temperature in the third pathT_{M3}=1730 *K* the temperature in the first and the second are respectively T_{M1}=500 *K* and T_{M2}=625 *K*.

Figure 16 is obtained with 1750 *W*, 1 m/min and 20% overlapping; the temperatures in the laser paths 1, 2 and 3 are quite similar to the temperatures calculated in Figure 15 but in the case reported in Figure 16 the temperatures are practically the same. By having an uniform temperature, the experimental trial shows a reduced softening in the overlapped paths. This result is more evident if compared to the results in Figure 17 where the laser scan strategy leads to a very different temperatures in the paths with a subsequent non homogeneous hardness profile.

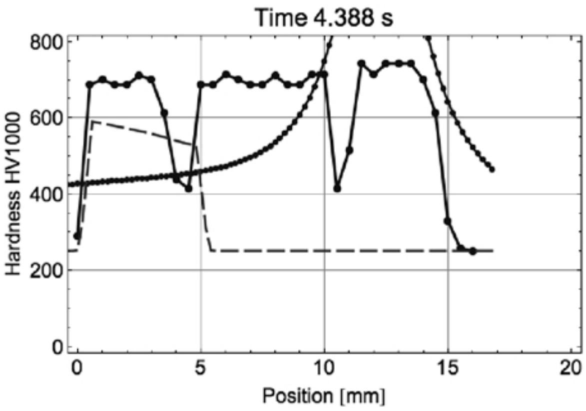


Figure 12: 1200 *W*, 0.4 *m/min*, 20% overlapping.

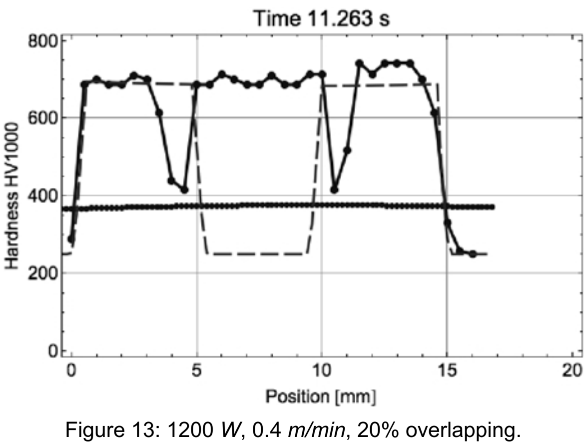


Figure 13: 1200 *W*, 0.4 *m/min*, 20% overlapping.

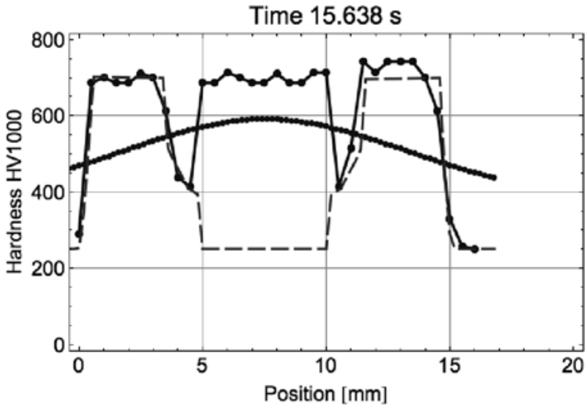


Figure 14: 1200 *W*, 0.4 *m/min*, 20% overlapping.

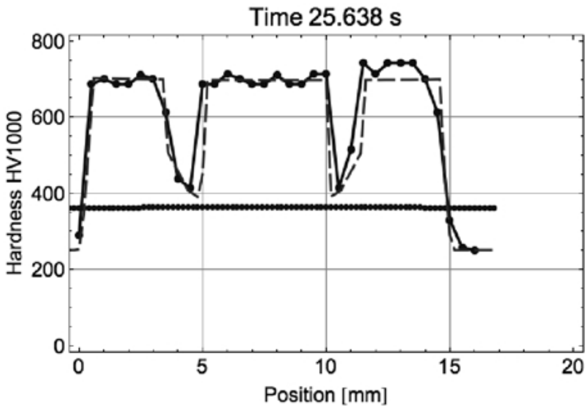


Figure 15: 1200 *W*, 0.4 *m/min*, 20% overlapping; T_{M3}=1730 *K*; T_{M1}=500 *K* and T_{M2}=625 *K*.

The above mentioned temperature difference parameter calculated in two adjacent paths is not the only physical criterion which can be used in the optimization of laser surface hardening. This paper represents only a preliminary work in this topic and further studies will be conducted by the authors in order to increase the accuracy of the proposed criterion.

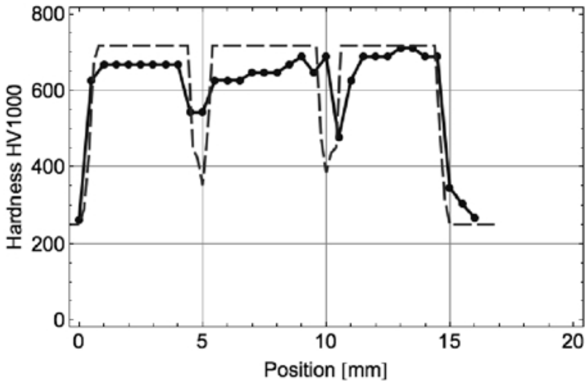


Figure 16: 1750 *W*, 1 *m/min*, 20% overlapping; T_{M3}=1743 *K*; T_{M1}=518 *K* and T_{M2}=515 *K*.

Figure 17 reports the comparison of the experimental trial presented in Figure 10.

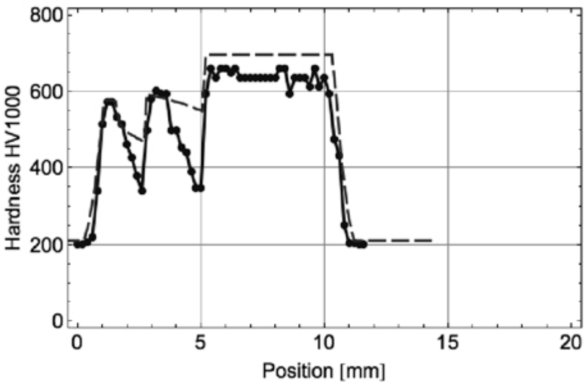


Figure 17: 1200 *W*, 0.4 *m/min*, 60% overlapping. T_{M3}= 661 *K*; T_{M1}=1314 *K* and T_{M2}=1547 *K*.

5 CONCLUSION

In this paper a numerical model for the laser hardening optimization is presented. The model is able to predict the resulting micro-structures after quenching due to single and multiple passes and the hardness of the micro-structures after tempering. The model is a useful tool in laser surface hardening when multiple pass are involved because it makes possible to predict the best laser path sequence strategy in order to obtain an optimal hardness distribution with the maximum process speed according to the required hardness depth. Three overlapped laser paths are used to prove the accuracy of the predicted results on a surface hardening of a AISI 1040 steel plates where the maximum surface temperatures calculated in the laser paths were selected as the physical parameter which can drive in the choice of the optimal laser scanning strategy.

6 REFERENCES

[1] Kennedy, E., Byrne, G., Collins, D.N., 2004, A Review of the Use of High Power Diode Laser in Surface Hardening, Int. J. of Material Processing Technology 156/10:1855-1860.

[2] Tani, G., Orazi, L., Fortunato, A., Campana, G., Cuccolini, G., 2007, Laser hardening process simulation for mechanical parts, Proc. of SPIE: High Energy/Average Power Lasers and Intense Beam Applications 6454.

[3] Tani, G., Orazi, L., Fortunato, A., Campana, G., Cuccolini, G., 2008, 3D modelling of LASER hardening and tempering of hypo-eutectoid steels, JLMN-Journal of Laser Micro/Nanoengineering, in press.

[4] Ashby, M.F., Easterling, K.E., 1984, The Transformation Hardening of Steel Surface by Laser Beams-I Hypo--Eutectoid Steel, Acta metall. 32/11:1935-1948.

[5] Ohmura, E., Inoue, K., Haruta, K., 1989, Computer Simulation on Structural Changes of Hypoeutectoid Steel in Laser Transformation Hardening Process, JSME International Journal 32/1:45-53.

[6] Skvarenina, S., Shin, Y.C., 2006, Predictive modeling and experimental results for laser hardening of AISI 1536 steel with complex geometric features by a high power diode laser, Surface and Coating Technology 201/6 :2256-2269

[7] Jacot, A., Rappaz, M., Reed, C., 1998, Modeling of Reaustenitization from the Pearlite Structure in Steel, Acta mater. 46711:3949-3962.

[8] Wang, Y., Denis, B., Appolaire, B., Archambault, P., 2004, Modelling of Precipitation of Carbides during Tempering of Martensite, J. of Phys. IV France 120:103-110.

[9] American Society of Metals, 1991, ASM Handbook 4: Heat Treating.

Fabrication of Wire Saw with Patterned Hard Bumps by Electrical Discharge Machining with Powder Suspended in Working Oil

Katsushi Furutani and Yasuo Murase
Department of Advanced Science and Technology, Faculty of Engineering
Toyota Technological Institute, Nagoya, Japan

Abstract
This paper deals with a fabrication method of a wire saw by electrical discharge machining (EDM). Hard bumps were discretely patterned on the surface of a thin stainless steel wire by EDM with titanium powder suspended in working oil. With an increase of the machining time, the average of the bump height was increased and no removal was observed. The bumps were deposited from four directions in an orthogonal sequence by manually turning the wire. A soda lime glass rod could be cut with the wire saw. The bumps tightly adhered on the wire.

Keywords:
Slicing; Deposition; Titanium carbide

1 INTRODUCTION

Today, wire saws are increasingly used because of a low kerf loss and applicable to cutting a large silicon ingot [1-4]. A fixed-abrasive wire saw is preferable to improve the machining speed. The wire saw is frequently used for cutting complex contours because it has an omnidirectional cutting edge [2]. Diamond grits are uniformly coated with electrodeposited nickel or resin on a core wire [3, 4]. Chips and sludge are often loaded on the abrasive layer so that the cutting performance will drop [5]. Therefore, an uneven or unconnected abrasive layer is sometimes preferable to scrape out chips and sludge. A wire saw with spiral abrasive bumps and twisted one have been proposed [6-8]. Because the deposited layer removed by grinding [6] or a thin wire should be covered with a fluororesin tape in advance of electrodeposition [7], it takes much labour. The section of the twisted wire saw is not round [8].

On the other hand, surface modification and deposition methods by electrical discharge machining (EDM) have been studied. In a deposition process using a green compact or a semi-sintered electrode made of titanium alloy or tungsten carbide [9], a surface of the electrode is crumbled by discharge and powder is always supplied on the workpiece after the powder reacts on the cracked carbon that comes from working oil by the heat caused by discharge. Then a thick, hard layer made of TiC or WC can be obtained by the adhesion of the powder. Because the mechanical strength of the green compact electrode is low, it cannot be used for patterning fine bumps. EDM with powder suspended in the working oil (hereafter the powder-suspended EDM) is a process to finish a workpiece surface or to harden it. The authors have proposed a deposition process by the powder-suspended EDM with a thin electrode [10]. Compared with the deposition method with a green compact electrode, this method saves a production time of the electrode and, in addition, can selectively deposit a hard layer on a small area.

The authors have deposited an abrasive layer by EDM with the green compact electrode and ground steel [11] and proposed to apply this method to the fabrication of the wire saw [12, 13]. This wire saw also had a uniform abrasive layer.

In this paper, a fabrication method of a wire saw with patterned hard bumps is proposed by using the powder-suspended EDM. At first, the deposition method by the powder-suspended EDM is introduced. Then the deposition of titanium carbide (TiC) bumps is demonstrated and the performance of the wire saw is evaluated through a cutting experiment.

2 DEPOSITION PRINCIPLE

Figure 1 illustrates the deposition principle [10]. Ti powder is suspended in working oil. The gap between a thin electrode and a workpiece soaked in the working oil is several tens up to hundreds micrometers. The polarity of the electrode is negative. The electrode moves reciprocally to stabilize a machining process. With a progress of machining, carbon is

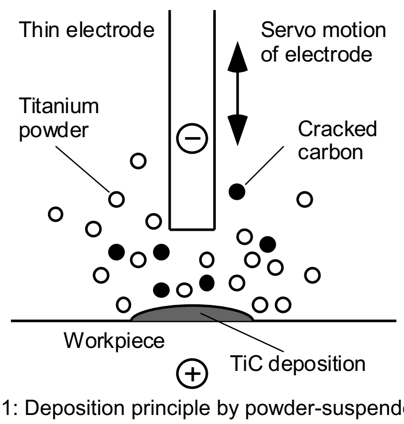


Figure 1: Deposition principle by powder-suspended EDM.

The 41st CIRP Conference on Manufacturing Systems, 2008

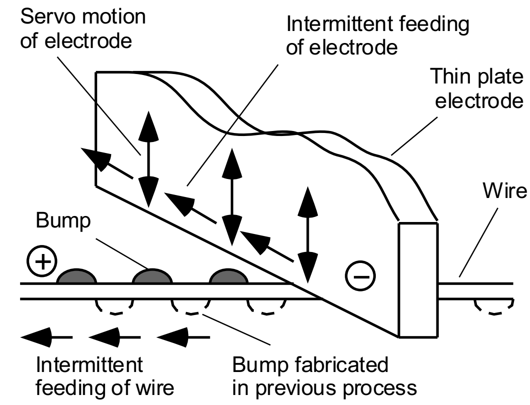


Figure 2: Fabrication of hard bumps by powder-suspended EDM.

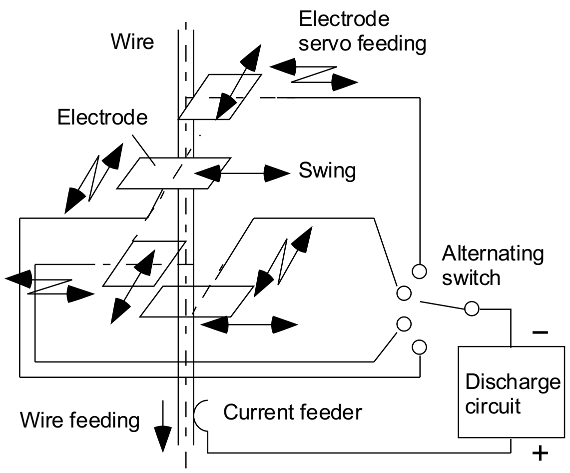
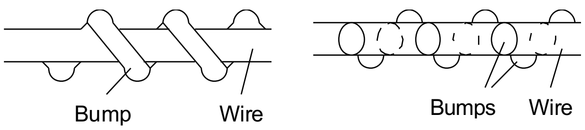


Figure 3: Concept of wire saw coater.



(a) Spiral. (b) Checked.

Figure 4: Examples of bump patterns.

cracked from the working oil. Ti powder is deposited on the workpiece as TiC after the reaction of the cracked carbon. The powder should be always provided into the small gap to keep the powder concentration high. Because a thin electrode seldom thrusts the working oil, it is useful for accretion process by the powder-suspended EDM.

Figure 2 shows an arrangement of an electrode and wire. A thin copper plate is used as an electrode to make the electrode alignment to a wire easy. A bump is deposited only in the faced part of the wire to the electrode. The electrode is shifted by a certain pitch to avoid heavy wear of the electrode when the wire is fed by one pitch.

Figure 3 shows a concept of a wire saw coater. The basic configuration of the unit is the same as one with the green compact electrode [14]. TiC layer is deposited from 4 directions. Each electrode is fed with a linear actuator and swung with a motor. The electrodes are switched to supply the current from a discharge circuit. Various bump patterns is fabricated by a combination of deposition direction with the wire feeding as shown in Figure 4. If thin rods are used as the electrode, the swing devices do not need.

Table 1 Conditions for TiC deposition process.

Electrode	Copper plate, t0.5 mm
Wire	Stainless steel, ϕ0.2 mm
Polarity of electrode	Negative
Open gap voltage	180 V
Discharge current	4 A
Pulse duration	8 μs
Pulse interval	1024 μs
Machining time/bump	30 s
Powder	Titanium < 24 μm
Working oil	EDF-K (Nippon Oil)
Powder concentration	50 kg/m^3
Circulation of working oil	19 l/min
Feeding pitch of wire	1 mm
Feeding pitch of electrode	0.5 mm
Jump height	1.8 mm
Jump speed	300 mm/min (5 mm/s)
Jump time	1 s

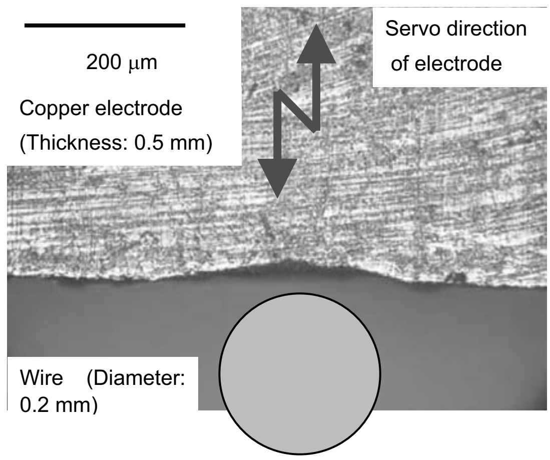


Figure 5: Appearance of electrode wear after deposition.

3 DEPOSITION EXPERIMENT

A normal die-sinking electrical discharge machine was used in the experiments. Table 1 shows the conditions for TiC deposition process. The electrical conditions were decided according to ones for the previous study [10]. Because the piano wire was inappropriate for this process due to the embrittlement by the cementation, an ASTM S30400 (JIS SUS304) stainless wire with a diameter of 0.2 mm was used as a core. The working oil was circulated with a pump at a flow rate of 19 l/min to avoid the sedimentation of powder. Tungsten carbide (WC) powder with an average diameter of 3.3 μm was mixed in a preliminary experiment. It was hard to deposit because of heavy sedimentation. Consequently, Ti

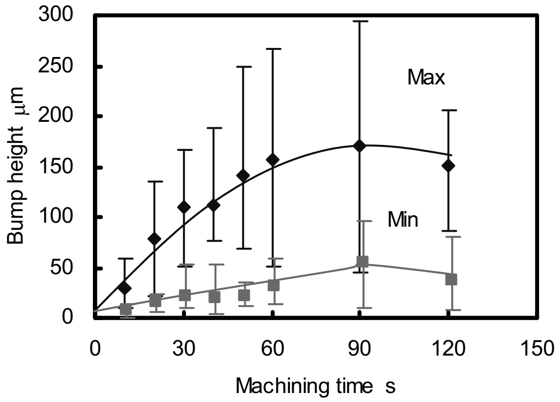
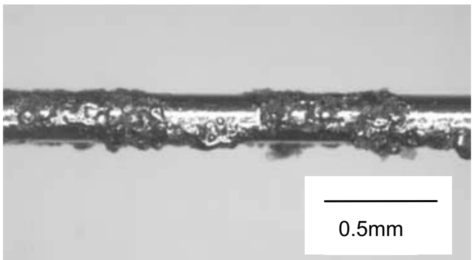


Figure 6: Relationship between bump height and machining time.

(a) Before cutting.



(b) After cutting.

Figure 7: Appearances of wire saw.

Table 2: Cutting conditions.

Workpiece	ϕ5 mm, soda lime glass rod
Cutting load	4.5 N
Stroke	60 mm
Average cutting speed	8 m/min
Working fluid	Soluble oil

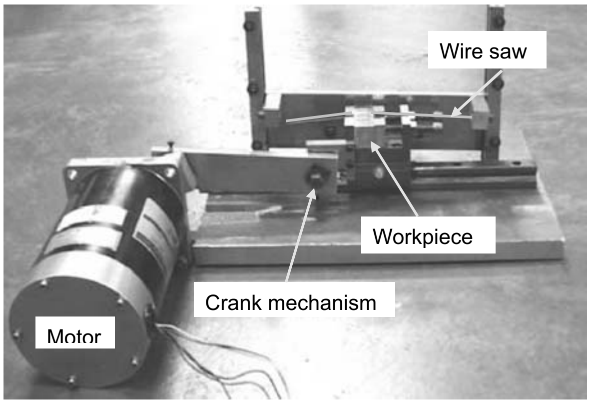


Figure 8: Appearance of cutting machine.

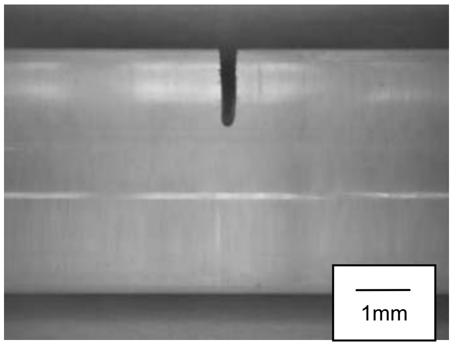


Figure: 9 Result of cutting.

powder was used. A working tank measured 130 mm × 130 mm × 100 mm.

Figure 5 shows an appearance of the electrode after a 30-s deposition. It wore 25μm in depth and 300 μm in width. Therefore, the electrode was shifted 0.5 mm, which is enough of the wire diameter, in the perpendicular direction to the wire after one process as shown in Figure 2.

Figure 6 shows the relationship between the height of 10 bumps and the machining time. The height was measured with a video microscope from the side. Although the height was dispersed very much, the wire surface was not removed. The average height of the deposition was increased with the increase of the machining time. Because the whole surface of a deposited portion was covered with the bump over 30 s, the machining time was determined to 30 s.

The bumps were deposited from 4 directions by the diagonal sequence to avoid the influence of the adjacent bump [14]. The wire was strained on a 110-mm frame and turned manually. Figure 7 (a) shows an appearance of the bumps with an interval of 0.5 mm. The average of maximum bump height was 70 μm. The bump width is the same as the electrode thickness.

4 CUTTING PERFORMANCE

Table 2 shows cutting conditions. A soda lime glass rod with a diameter of 5 mm was used as a workpiece. It was fed with a crank mechanism driven with a stepping motor shown in Figure 8. The average feeding speed and the cutting load were set to 8 m/min and 4.5 N, respectively. Because the stroke is short and the same portion is rubbed in a short time, this device is good for a durability test of the wire saw.

Figure 9 shows a result of cutting and Figure 10 shows a progress of cutting. The groove depth was measured every 500 reciprocating motions. Then the section was calculated. The cutting speed was constant after 500 reciprocating motions though it was smaller than the initial one. The wire saw broke after 2980 reciprocating motions. The kerf loss was 290 μm.

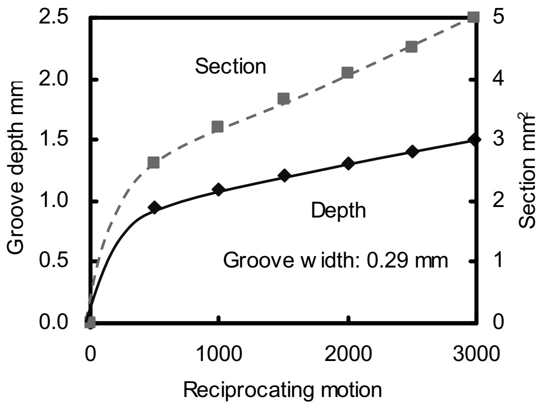


Figure 10: Progress of machining.

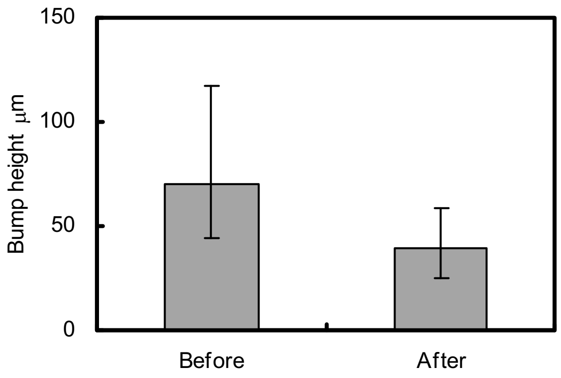


Figure 11: Bump height before and after machining.

Figure 7 (b) shows an appearance of the wire saw after cutting. The bumps neither scale off nor chip but wore off. Stress concentration at the boundary of a bump caused the break. Figure 11 shows the bump height before and after cutting. Ten bumps were measured with the video microscope. Because higher parts wore earlier, the distribution of the bump height became smaller after cutting. Rubbing with the core wire caused the break in the case of a bump interval of 1.5 mm. An appropriate interval should be investigated in future.

5 CONCLUSIONS

The patterning method of the hard bumps on a wire saw was proposed in this paper. Conclusions can be drawn as follows.

(1) The bumps can be patterned on a thin wire by using the powder-suspended EDM.

(2) The fabricated wire saw with TiC bumps can be used for cutting a glass rod.

The authors will build a wire saw coater to accelerate the fabrication speed and to fabricate various patterns automatically.

6 ACKNOWLEDGMENTS

The authors would like to thank Prof. Kazuyoshi Kondo of Toyota Technological Institute and Dr. Masayuki Suzuki of Sansyu Engineering Service for their valuable advice, and Mr. Toshiyuki Yamauchi and Mr. Hiroyuki Ohue of Tokusen Kogyo for their helpful cooperation. This study was supported by Industrial Technology Research Grant Program in 2005 from New Energy and Industrial Technology Development Organization (NEDO) of Japan.

7 REFERENCES

[1] Sreejith, P. S., Udupa, G., Noor, Y. B. M., Ngoi, B. K. A., 2001, Recent Advances in Machining of Silicon Wafers for Semiconductor Application, Int. J. Adv. Manuf. Technol., 17/3:157-162.

[2] Clark, W. I., Hardin, C. W., Lemaster, R. L., Shih, A. J., Mcspadden, S. B., 2003, Fixed abrasive diamond wire machining. Part I. Process monitoring and wire tension force, Int. J. Mach. Tools Manuf., 43/5:523-532.

[3] Chiba, Y., Tani, Y., Enomoto, T., Sato, H., 2003, Development of a High-Speed Manufacturing Method for Electroplated Diamond Wire Tools, CIRP Ann., 52/1:281-284.

[4] Enomoto, T., Shimazaki, Y., Tani, Y., Suzuki, M., Kanda,, 1999, Development of a Resinoid Diamond Wire Containing Metal Powder for Slicing a Slicing Ingot, CIRP Ann., 48/1:273-276.

[5] Sugawara, J., Hara, H., Mizoguchi, A., 2004, Development of Fixed-Abrasive-Grain Wire Saw with Less Cutting Loss, SEI Tech. Rev., 58:7-11.

[6] Ishikawa, K., Suwabe, H., Makino, M., Yamasaka, S., 1996, Development of Spiral Chip-pocket Wire Tool Electrodeposited Diamond Grains, J. Jpn. Soc. Precis. Eng., 62/2:242-246.

[7] Suwabe, H., Ishikawa, K., Yamasaka, S., 1999, Development of Spiral Wire Tool Electrodeposited Diamond Grains Used Teflon-Coating, J. Jpn. Soc. Abrasive Technol., 43/10:431-436.

[8] Ishikawa, K., Suwabe, H., Ohta, K., 2003, Development of a Diamond Wire Tool Using Twisted Core Wire, J. Jpn. Soc. Abrasive Technol., 47/9:495-500.

[9] Mohri, N., Saito, N., Tsunekawa, Y., 1993, Metal surface modification by electrical discharge machining with composite electrode, CIRP Ann., Vol42, no1, 219-222.

[10] Furutani, K., Saneto, A., Takezawa, H., Mohri, N., Miyake, H., 2001, Accretion of Titanium Carbide by Electrical Discharge Machining with Powder Suspended in Working Fluid, Precis. Eng., 25/2:138-144.

[11] Furutani, K., Sunada, H., 2004, Fabrication of Abrasive Layer Using Dispersion of Hard Powder by Electrical Discharge Machining (1st Report) Dispersion of Insulating Powder into Deposit, Int. J. Electr. Mach., Vol9, 15-20.

[12] Furutani, K., Kanai, M., Mieda, Y., Suzuki, M., 2006, Proposal of Abrasive Layer Fabrication on Thin Wire by Electrical Discharge Machining, Proc. Int. Conf. Adv. Mater. Process. Technol., Las Vegas, NV, USA, NT-II-6/1-4.

[13] Furutani, K., Suzuki, K., 2007, A Prototype of Electrical Discharge Coater for Wire Saw, Proc. 40th CIRP Int. Seminar Manuf. Syst., Liverpool, UK, PT-7/1-6.

Creation of Cross-Section Changing Hole with a Hemisphere by Means of Electrical Discharge Machining

T. Ishida[1], H. Nagasawa[1], M. Kita[2], Y. Takeuchi[1]
[1]Department of Mechanical Engineering, Graduate School of Engineering, Osaka University, Japan
[2]Department of Machinery System Production Technology, Kinki Polytechnic College, Japan

Abstract
In general, machined holes have straight axes and their cross-sections are constant circles since holes are machined by drills. In other words, shapes of machined holes do not have so much variety. To solve the problem, the device was developed which can machine a certain shape on an inside wall of a straight hole by means of electrical discharge machining. This results in fabrication of holes whose cross-sections change variously. In the study, the holes are called cross-section changing holes. However, the device had only ability to make two-dimensional shapes. Therefore, the device has been improved so that three-dimensional shapes can be machined inside holes. From the result of the machining experiment, it is found that the improved device can machine a cross-section changing hole with a hemisphere.

Keywords:
Electrical Discharge Machining, Cross-Section Changing Hole, Master-Slave System

1 INTRODUCTION

In the field of machining, general drilling technique by means of rotational cutting tools such as drills has been established for many years. Therefore, most of mechanical engineers have taken it for granted that drilling is to machine holes which have straight axes and constant cross-sections. Although other machining methods have been established and put into practical use, achievable shapes of their machining holes are generally limited to straight axes and constant cross-sections as long as the practical machining methods are employed in their usual manners. In other words, range of hole shape which the machining methods can realize is restricted. This means that degree of freedom of mechanical design is restrained. To overcome the restriction, it is necessary to diversify the realizable hole shape.

To meet the requirement, the authors have developed new methods to machine a hole whose cross-section changes complicatedly, which is called a cross-section changing hole. The hole shape can be machined by wire electrical discharge machining and multi-axis control machining [1]. However, the achievable hole shapes by these machining methods are also limited since the former cannot fabricate the shapes which cannot be realized by wire scanning and the latter cannot do the shapes which cannot be reached by cutting tools. Accordingly, the study has aimed at development of a new machining method for the cross-section changing holes which these machining methods cannot fabricate. Basic strategy of the study is to develop a unique device to make the most complicated shape possible against an inside wall of a beforehand drilled straight hole by means of electrical discharge machining. Actually, prototypical devices have been developed which can do such machinings and some cross-section changing holes can be fabricated by using the devices [2, 3]. However, the shapes of the holes which can be realized by the devices have been restricted to two-dimensional ones.

To solve the problem, in the study, the last device is improved so as to have the ability to fabricate three-dimensional complicated cross-section changing holes by introducing a rotational servomotor. As the first step to confirm the fundamental performance of the improved device, in the paper, the improved device is subjected to experiments in the condition that the target shape to be machined is set to a relatively simple shape, i.e., a cross-section changing hole with a solid of revolution.

2 DEVICE AFTER THE IMPROVEMENT

2.1 Structure and motion of the device

Figure 1(a) illustrates a schematic view of the improved device developed in the study. The device is installed on a die-sinking electrical discharge machine (EDM) and is used with the EDM.

As shown in the figure, the device has a rotational servomotor and two linear actuators. The linear actuators have identical components, that is, each linear actuator consists of a servomotor and a ball screw with a guide and a nut and the components are identical. The rotational servomotor is installed on a gate-shaped jig fastened to the machine table, i.e., the bottom of the working tank of the EDM. On the rotational servomotor, besides, the linear actuators are mounted. To the nuts of the linear actuators, moreover, shafts are respectively attached at their upper ends so as to be through holes in the center of the rotational servomotor and the gate-shaped jig. At their lower ends, furthermore, one shaft is joined to a tool electrode for electrical discharge machining through a link and the other shaft is done to the electrode directly. The former shaft and the latter shaft are respectively labeled ‘Shaft 1’ and ‘Shaft 2’. Additionally, the respective linear actuators to which the former and the latter are attached are called ‘Actuator 1’ and ‘Actuator 2’. Shaft 1, the link, the electrode and Shaft 2 are each other assembled through a tiny bearing with a tiny rod. In the respective connecting points between them, they can rotate each other. As a consequence, every connection of them is rotation kinematic pair and they constitute a slider crank chain. On the other hand, a linear scale is installed on a wall of

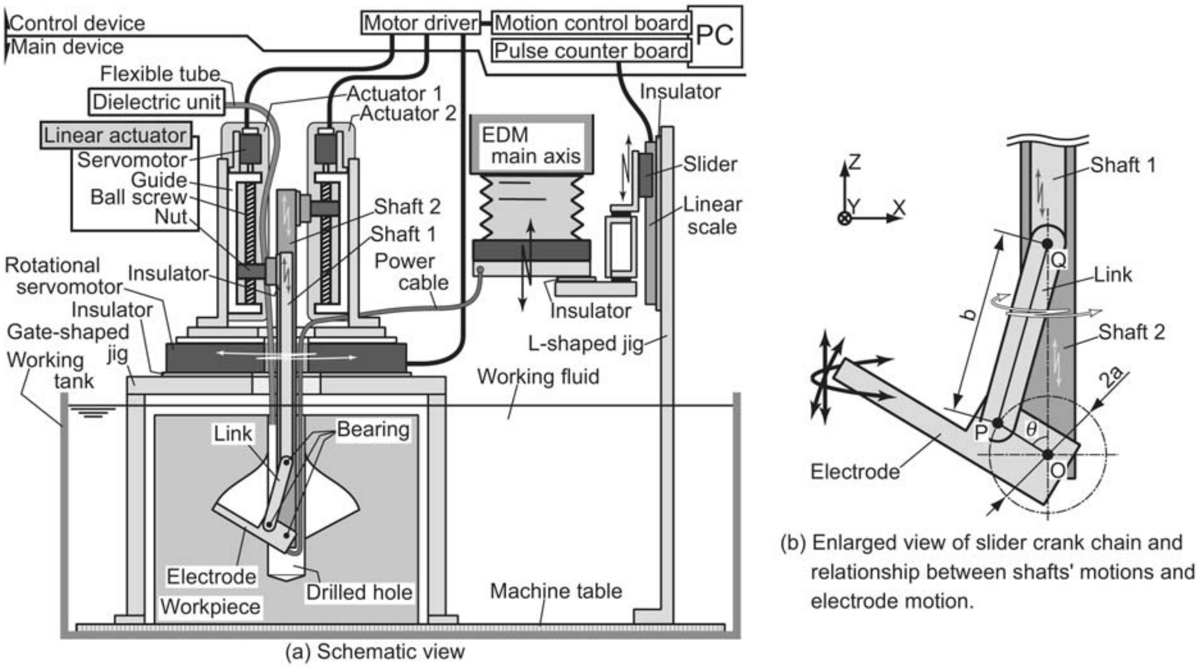


Figure 1: Schematic view of improved device, enlarged view of slider crank chain and relationship between the motions.

an L-shaped jig, which is clamped on the machine table of the EDM, so as to be parallel with the perpendicular motion of the main axis of the EDM. The slider of the linear scale is attached to the EDM main axis.

The discharge current to machine a workpiece is supplied from the EDM main axis to the electrode through a power cable connected between them. The use of the power cable can prevent all of the discharge current from passing through the bearings, which has possibility of damaging the bearings since the electrical discharge machining is brought about between the outer ring and the inner ring of the bearing. After machining, the discharge current passes from the workpiece to the machine table. The other discharge current paths are insulated. In addition, the rotational servomotor, the linear actuators and the linear scale are insulated from the gate-shaped jig, the shafts, the L-shaped jig and the EDM main axis by inserting insulators between them. Besides, working fluid can be provided from the dielectric unit of the EDM through a flexible tube to remove debris occurring in electrical discharge machining.

Near the EDM, a personal computer (PC) is placed. The PC is equipped with a pulse counter board and a motion control board. The former and the latter are respectively connected with the linear scale and three motor drivers which are connected to the rotational servomotor and the servomotors of Actuator 1 and 2. Therefore, the PC can get the position of the EDM main axis and can control the respective servomotors, namely, the rotational angle and the perpendicular positions of Shaft 1 and 2, simultaneously. In other words, the rotational motion of the whole of Shaft 1 and 2 and Actuator 1 and 2 and the respective perpendicular motions of Shaft 1 and 2 can be independently and arbitrarily controlled according to the feed of the EDM main axis. That is to say, the device constitutes the master-slave system that the feed of the main axis is the master and the rotational and perpendicular motions of Shaft 1 and 2 are the slaves. In addition, the relationship between the master and the slaves can be defined by means of software in the PC. This means that the relationship can be easily changed and be made to have high degree of freedom.

Figure 1(b) illustrates an enlarged view of the slider crank chain and the relationship between the motions of Shaft 1 and 2 and the motion of the electrode. A coordinate system is defined as shown in the figure.

When Shaft 1 and 2 have a relative motion in the perpendicular direction, i.e., in the direction of Z axis, as seen from the figure, Shaft 1, the link, the electrode and Shaft 2 constitute a reciprocating block slider crank mechanism that Shaft 1 plays a role of the reciprocating block slider. This results in the rotation of the electrode around the direction of Y axis. In letting the relative distance between Shaft 1 and 2 be h and the rotational angle of the electrode be θ, the relationship between h and θ can be approximately expressed by the following equation:

$$h = a\left(1 - \cos\theta + \frac{a}{2b}\sin^2\theta\right) \quad (1)$$

, where a is the crank length OP, and b is the link length PQ, as shown in the figure. Solving Equation (1) for θ, Equation (2) is obtained [2]:

$$\theta = \cos^{-1}\left(\frac{-b + \sqrt{(a+b)^2 - 2bh}}{a}\right) \quad (2)$$

Besides, when the whole of Shaft 1 and 2 travels in the perpendicular direction, i.e., in the direction of Z axis, the electrode has same motion. Additionally, when the rotational servomotor has a motion, the electrode rotates around the direction of Z axis.

As a result, the electrode can make the complicated motion that three degrees of freedom, namely, the translation in the

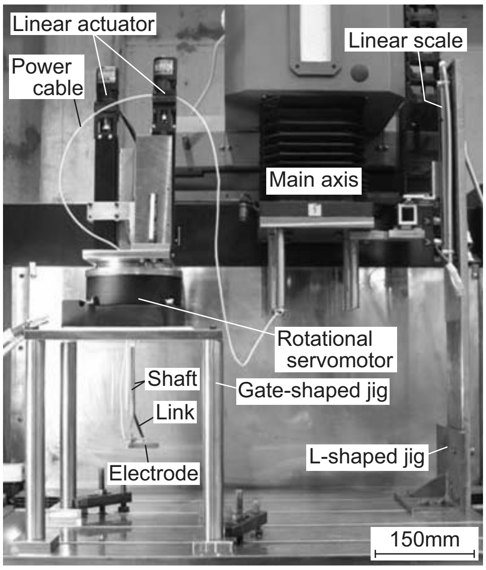


Figure 2: Actual view of improved device around the EDM main axis.

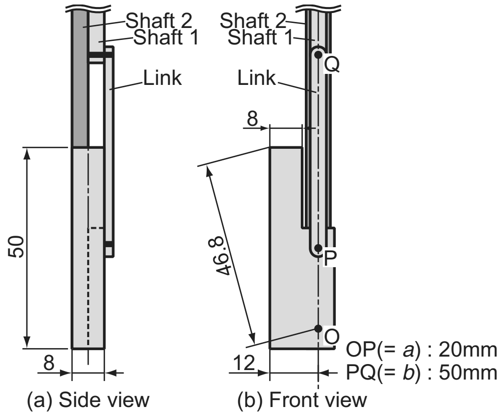


Figure 3: Dimension of employed slider crank chain and electrode.

direction of Z axis and the rotations around the directions of Y axis and Z axis, are combined. Moreover, discharge gap control by the EDM is realized at the electrode since the complicated electrode motions are derived from the perpendicular motion of the EDM main axis, which is the master of the master-slave system. This means that electrical discharge machining with the electrode is achieved, that is to say, an envelope of the electrode moving locus can be machined.

2.2 Actual device

Figure 2 depicts an actual view of the improved device around the EDM main axis. The employed motor driver, servomotor and ball screw with guide and nut, which constitute the linear actuator, are AC servo driver (model: TA8110N), AC servomotor (model: TS4502), which are made by Tamagawa Seiki and LM guide actuator (model: KR2001A) made by THK, respectively. The utilized rotational servomotor and motor driver is Megatorque Motor (model: M-YSB2020) and Driver Unit (model: M-ESB-YSB2020), which are made by NSK. As the linear scale, feedback scale (model: SD-608RA) made by Sony Precision Technology is used. The power cable is a coated wire of 2.0mm in diameter including coating. Its current tolerance is 20A. The EDM employed in the study is numerically controlled die-sinking electrical discharge machine (model: EDNC65) made by Makino Milling Machine.

The PC utilized in the study is AT/AT compatible whose main memory capacity is 1.0GB. Its CPU is Pentium 4 processor made by Intel, whose clock rate is 3.40GHz. The pulse counter board and the motion control board which are installed on the PC are counter/timer (model: PCI-6602) and servo motion controller (model: PCI-7344) made by National Instruments, respectively. Used OS is Windows 2000 SP4 made by Microsoft and the software to control the device is developed by means of LabView made by National Instruments.

Figure 3 illustrates the dimensions of the actually employed slider crank chain and electrode. The crank length *a* and the link length *b* are set to 20mm and 50mm, respectively. The electrode is made of oxygen-free copper and its shape is a rectangular solid of 8mm in length, 8mm in width and 50mm in height. The rotational radius around the direction of Y axis is 46.8mm.

3 MOTION AND MACHINING EXPERIMENTS

3.1 Target electrode moving locus

The electrode moving locus, i.e., the machined shape, which can be realized by the improved device exists variously since

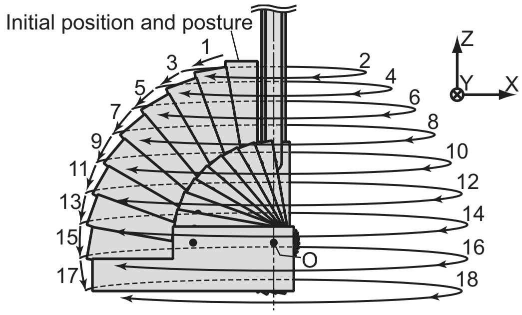


Figure 4: Target electrode motion.

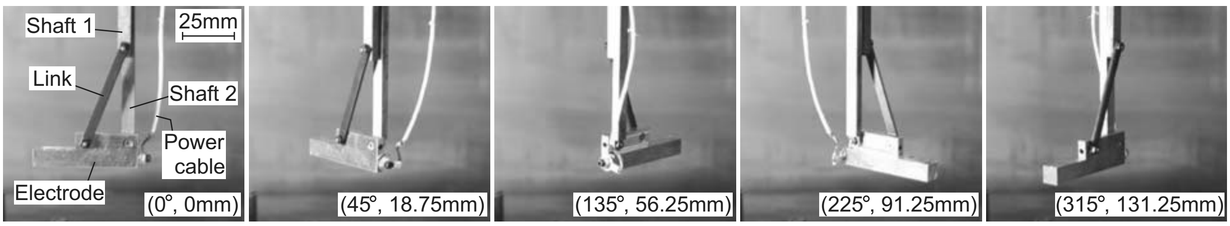


Figure 5: Actual motion of electrode in feeding the EDM main axis when rotational angle around Y axis is 90°.

the device has the ability to theoretically achieve the electrode motion of infinity. This makes it difficult to choose the electrode motion which is subjected to motion and machining experiments. To confirm the effectiveness of the improved device, in the study, the electrode moving locus should be selected which can be achieved only by the improvement and which is not very complex.

Figure 4 illustrates the target electrode motion. The electrode in the initial position and posture rotates 10º around the direction of Y axis first, and then 360º around the direction of Z axis. After that, these processes are repeated until the rotational angle around the direction of Y axis becomes 90º. The electrode motion is so simple that the improved device always controls just one servomotor. However, the electrode moving locus cannot be realized by the device before the improvement.

3.2 Motion experiment

Motion experiment is conducted by making the program to control the electrode motion shown in Figure 4 so as to realize the electrode moving locus. In the rotation around the direction of Y axis, the main axis feed is found from Equation (1) so that the rotation of 10º is achieved. In the rotation around the direction of Z axis, the rotation of 360º is uniformly assigned to the main axis feed of 150mm.

Figure 5 depicts change in posture of the electrode which rotates 360º around the direction of Z axis when the rotational angle around the direction of Y axis is 90º. The angle and length indicated in each figure are the angle around the direction of Z axis and the main axis feed, respectively. As can be seen from the figure, the electrode has the expected motion.

3.3 Machining experiment

Machining experiment is done by using the electrode motion shown in Figure 4. When the electrical discharge machining that employed the electrode motion starts in a straight hole, it is expected that a hemisphere shape is machined in the straight hole. In other words, a cross-section changing hole with a hemisphere will be fabricated.

Machining conditions are set as follows. Discharge current, pulse duration and duty factor are respectively set to 10A, 155μm and 80% under the polarity of electrode (+) / workpiece (-). Jump and orbital motion of the EDM main axis are not used. Workpiece is made of aluminum alloy (A5052) and its shape is a rectangular solid with a straight hole of 26mm in diameter and 100mm in depth. Working fluid is oil and is poured into the straight hole through a flexible tube.

Figure 6 depicts a sectional view of the machined workpiece. The machining time is 2676min. and the machining rate is 72.5mm^3/min. The electrical discharge machining in the experiment is stable. From the figure, it can be seen that the expected cross-section changing hole is machined. From the comparison between Figures 4 and 6, in addition, it is found that the envelope of the electrode moving locus is almost same as the sectional shape of the machined cross-section changing hole. These prove that the improved device has the ability to fabricate a three-dimensional cross-section changing hole.

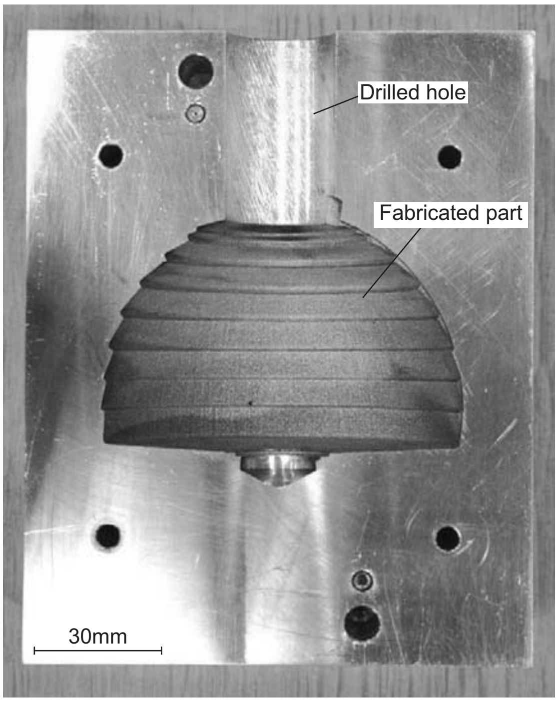


Figure 6: Sectional view of machined workpiece.

4 CONCLUSIONS

Aiming at shape diversification of holes which can be machined, the devices have developed which can machine some shape against an inside wall of a straight hole by means of electrical discharge machining. By using the device, cross-section changing holes can be machined. However, these holes are two-dimensional. To overcome the restriction, in the study, the last device has been improved so that three-dimensional cross-section changing holes can be fabricated by introducing a rotational servomotor. From the results of the motion and machining experiments, it is found that the improved device can fabricate a cross-section changing hole with a hemisphere. This means that the improved device has the ability to fabricate various three-dimensional cross-section changing holes.

5 ACKNOWLEDGEMENTS

The study is partly supported by the Ministry of Education, Culture, Sports, Science and Technology, Grant-in-Aid for Young Scientists, (B)16760091, and by Mitutoyo Association for Science and Technology (MAST).

6 REFERENCES

[1] Morishige, K. and Takeuchi, Y., 2000, Strategic Tool Attitude Determination for Five-Axis Control Machining Based on Configuration Space, CIRP J. of Manufacturing Systems, 31/3:247-252.

[2] Ishida, T., Mochizuki, Y., and Takeuchi, Y., 2002, Fin Fabrication inside a Hole by Means of Electrical Discharge Machining and Reciprocating Block Slider Crank Mechanism, Proc. of 2002 Japan-USA Symposium on Flexible Automation (2002JUSFA), 349-355.

[3] Ishida, T., and Takeuchi, Y., 2005, Development of Hole-Inside shape Fabricating Method ---Machined Shape Control by Means of Software---, Proc. of 5th Manufacturing & Machine Tool Conference, Japan Society of Mechanical Engineers (JSPE), 293-294 (in Japanese).

Cutting

Study of Cutting Temperature in Cold-Air Milling of Ti6Al4V Alloy

Feng Jiang[1, 2], Jianfeng Li[1], Yiming (Kevin) Rong[2], Jie Sun[1], Song Zhang[1]

[1] School of Mechanical Engineering, Shandong University, China

[2] Department of Mechanical Engineering, Worcester Polytechnic Institute, United State

Abstract

The fast tool wear and poor surface integrity resulted from high temperature are concerned in the milling titanium alloy Ti6Al4V. Using cutting fluid is the most common method to control the temperature but it may be harmful to the environment and personal healthy. Cold-air with oil mist is an effective substitute. In this study, the friction factors and temperature-dependent heat transfer coefficient in dry cutting, cold-air cutting with different air pressure, and cold-air+oil cutting, are studied with theoretical analysis and experimental data. Then the temperature distribution is predicted by using the finite element analysis method.

Keywords:

Cold-air milling; Cutting temperature; Ti6Al4V

1 INTRODUCTION

Titanium alloy Ti6Al4V is widely used in aircraft industry, marine and commercial applications due to its high specific strength and excellent corrosion resistance. However, Ti6Al4V is a typical difficult-to-machine material. High cutting temperature in machining Ti6Al4V usually leads to fast tool wear and poor surface quality. Using cutting fluid is the most common strategy to control the cutting temperature, but it also brings in the environmental and cost concerns. It has been pointed out that cutting fluid, especially water soluble coolant, may lead to hydrogen and oxidation of workpiece and cutting tool materials, which result in excessive chipping at the cutting edge and fracture on the rake face of the cutting tool [1]. Cold-air cutting method is an effective substitute of conventional cutting fluid. This method uses the cutting fluid of only a small amount, typically less than 50ml/h, which is typically jetted into the cutting zone with a flow of compressed cold-air. It is a Minimal Quantity Lubrication (MQL) technology.

The research of cold air cutting was pioneered in 1996 [2]. The technology was developed and applied in turning and grinding. The cold air turning of Ti6Al4V was studied and it was proposed that the tool life using cold air was equal to the tool life using minimal MQL [3]. The influence of oil mist on the tool temperature was studied experimentally in the turning, intermittent turning and end milling. The effect of oil mist was found greater in intermittent cutting, such as end milling, than in continuous cutting such as turning [4]. The cutting experiments were performed with air blow, air blow+water, air blow+oil conditions. It was found that the cutting forces and oscillation amplitude decreased, especially noticeable with low cutting speeds and high feed rate [5]. The tribological action and cutting performance of MQL media was studied in machining Aluminium Alloy. It was found that the adsorption behaviour onto metal surfaces was in close correlation with the cutting performance of lubricant and carrier gases in practical MQL machining of aluminium [6].

Many researches can be found in this field but most of them were experimental researches, focusing on tool wear, process optimization, and lubricant performance and. The prediction of temperature field is only mentioned in few researches. Cutting temperature is important for prediction of tool wear and surface integrity. Therefore this study focuses on modelling thermal boundary conditions and the prediction of cutting temperature with an FEA model. By using experimental data and theoretical analysis, the friction factor in dry cutting, cold air cutting without oil mist and cold air cutting with oil mist are evaluated. The temperature-dependent heat transfer coefficients in natural air cooling and forced air cooling with different air pressure are calculated. The temperature field in different cooling and lubrication conditions are simulated using the finite element software AdvantEdge. Finally the temperatures in different cooling and lubrication conditions are measured by NEC infra-red (IR) pyrometer to validate the simulation results.

2 PREDICTION OF FRICTION EFFECTS

2.1 Prediction methods

There are experimental and theoretical methods for the prediction of static friction effects. Pin-on-disk (ring-on-block) test is designed for low velocity friction experiment, where the friction coefficients are predicted only in the steady state condition [7]. This is very helpful in understanding the mechanisms involving degradation of material surface in contact. But the results may not valid in the high velocity, temperature and pressure in the high speed cutting. Theoretical method is based on the elementary cutting principle and experimental data, basically the cutting forces. The friction coefficient was predicted in dry cutting and cryogenic cutting Ti6Al4V with experiments [8]. Experimental precision may influence the prediction of friction coefficient greatly. The theoretical method is employed in this paper and cutting force experiments are performed.

2.2 Theoretical model

Assuming the cutting tool is absolutely sharp and no helix angle, the cutting forces mainly result from shearing on the major/minor shear plane and the friction on the tool-chip interface [9], as shown in Figure 1(a). The friction force F_f can be approximately expressed as a function of friction coefficient and the normal pressure on the tool-chip interface, which is the component of shearing force, as shown in Figure 1(b).

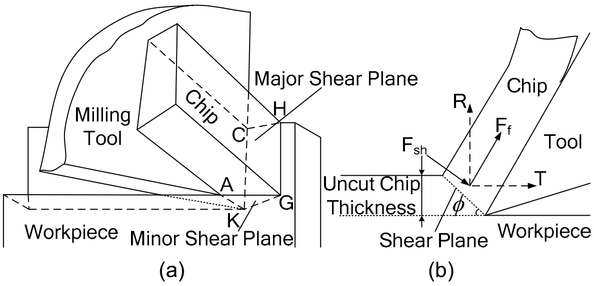


Figure 1: Double-edge cutting principle.

It is well known that the shear forces F_{sh} is in a direct proportion to the area A of shear plane [10]:

$$F_{sh} = \tau_s \cdot A$$

where τ_s is the shear stress of material, which depends on temperature, strain, and strain rate.

Since the shearing force F_{sh} in the main shear plane CKGH, friction force F_f can be respectively expressed as:

$$F_{sh} = \tau_s \cdot h(t) \cdot a_p / \sin\phi$$

$$F_f = \tan\beta_a \cdot \tau_s \cdot h(t) \cdot a_p \cdot \cos(\phi - \alpha) / \sin\phi$$

where a_p is the depth of cut and $h(t)$ is the instantaneous uncut chip thickness, which is approximated as [11]:

$$h(t) = f_z \cdot \sin(\gamma - wt)$$

where γ is the entry angle and it is expressed as:

$$\gamma = \arccos(1 - \frac{a_e}{R})$$

where a_e is the width of cut and R is the cutter radius.

Applying the minimum energy principle in predicting the shear angle [10], the expression of shear angle becomes,

$$\phi = \frac{\pi}{4} - \frac{\beta_a - \alpha}{2}$$

where α is rake angel of cutting tool and β_a is friction angle. The friction coefficient can be expressed as:

$$\mu = \tan\beta_a$$

The force components in the direction R and T are :

$$F_T = F_{sh} \cdot \cos\phi - F_f \cdot \sin\alpha + F'_{sh}$$

$$F_R = F_{sh} \cdot \sin\phi + F_f \cdot \cos\alpha$$

where F_{sh}' is the shearing force in the minor shear plane AKG and it is expressed as [9]:

$$F'_{sh} = \frac{1}{2} \cdot \frac{h(t)}{a_p} (\mathrm{ctan}\phi + \tan\alpha) \cdot \sin\phi$$

In the global coordinate system, the forces components in the direction X and Y are expressed as [11]:

$$\begin{bmatrix} F_X \\ F_Y \end{bmatrix} = \begin{bmatrix} \cos(\gamma - wt) & -\sin(\gamma - wt) \\ \sin(\gamma - wt) & \cos(\gamma - wt) \end{bmatrix} \begin{bmatrix} F_T \\ F_R \end{bmatrix}$$

Ratio k can be used to define as F_X/F_Y and shear stress τ_s is eliminated. Ratio k can be also obtained from cutting experiments. Then the friction angle can be estimated.

2.3 Experimental condition

The cutting tool selected is SECO plunging cutter of 25mm diameter with uncoated carbide inserts XOMX (ISO K40 grade, axial rake angle=+5°, helix angle=0°, edge radius=5μm).

The machine tool is DAEWOO high speed machining center. The cutting force is measured using a three-component piezoelectric dynamometer (Kistler 9275A). The experimental system is shown in Figure 2.

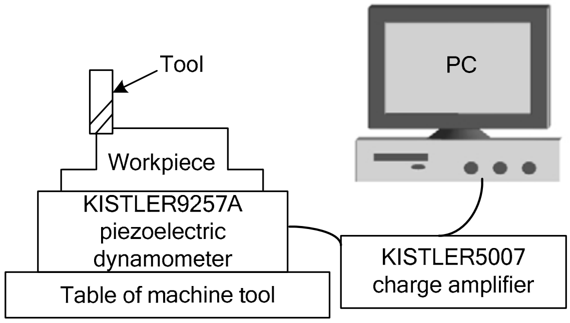


Figure 2: Sketch of measurement and machining system.

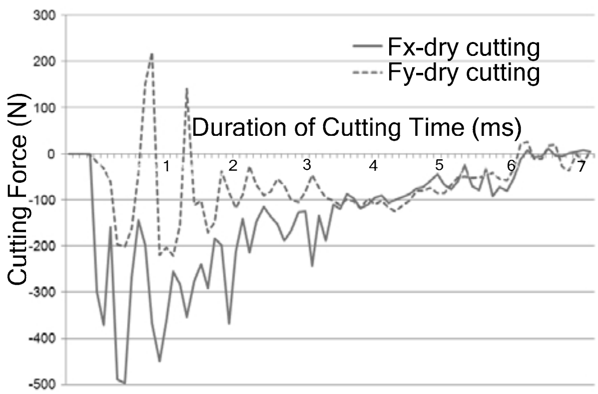
In the experiments, only one insert was installed in the plunging cutter to avoid the interference with others. Two cold-air nozzles were used, aligning to the rake and relief surfaces of the insert. The machining parameters are listed in Table 1.

Table 1: Machining parameters.

Cutting Mode	down milling
Cutting Speed	100m/min
Feed per Tooth	0.1mm
Depth of Cut	1mm
Width of Cut	5mm
Cutting conditions	dry/cold air only/cold air+oil mist

2.4 Results and discussions

The cutting forces measured and friction coefficients estimated are shown in Figures 3~5.



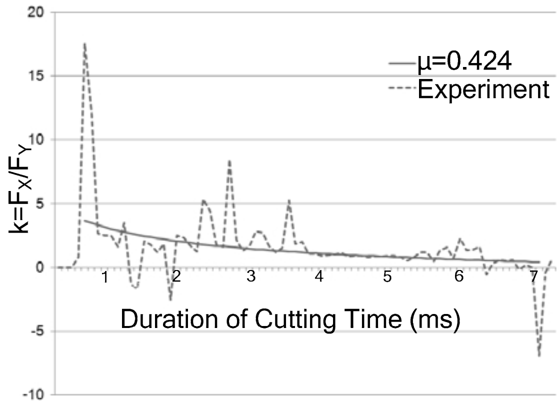


Figure 3: Cutting forces and friction coefficient in dry cutting.

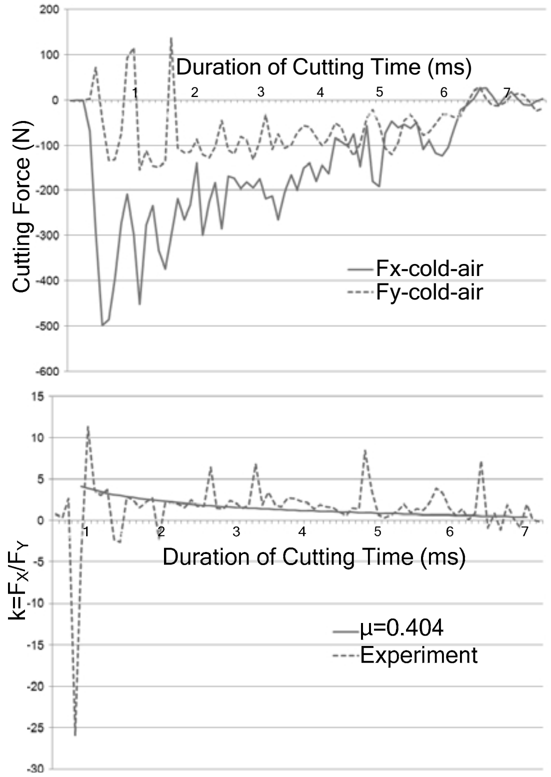
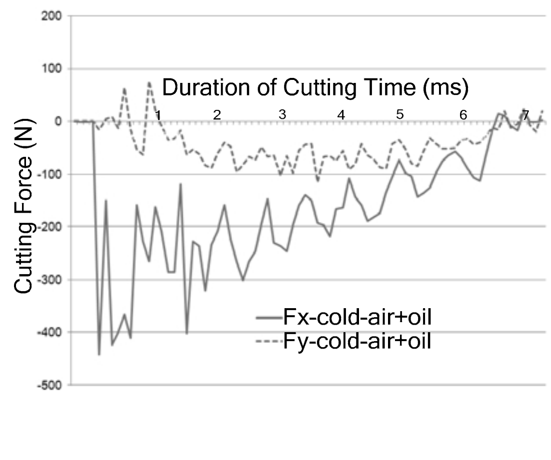


Figure 4: Cutting forces and friction coefficient in cold-air cutting without oil mist.



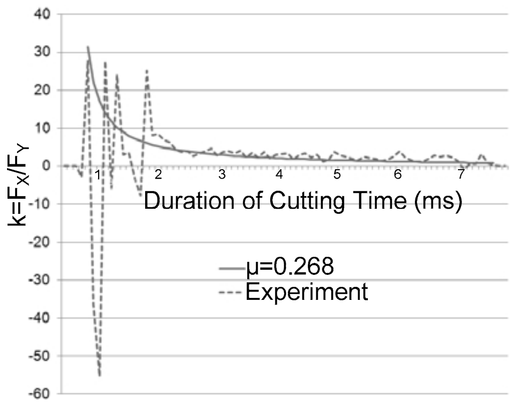


Figure 5: Cutting forces and friction coefficient in cold-air cutting with oil mist.

The X-direction cutting force decreases by about 100 N with oil mist, but little change with cold-air only. The Y-direction cutting force has little change in different cooling conditions, but it seems smoother when the oil mist is applied. The friction coefficient in dry cutting is about 0.424, which agrees with the results in the literature [8]. In cold-air cutting without oil mist it is about 0.404 and 0.268 in cold-air cutting with oil mist. The friction coefficient in cold-air cutting without oil mist decreases to some extent, compared with dry cutting. The main reason is that vapor in the cutting zone condenses to water, which acts as the coolant, as well as cooling itself can reduce the friction [12].

3 ESTIMATION OF HEAT TRANSFER COEFFICIENTS

3.1 Heat generated and heat lost

In metal cutting, three heat sources are assumed to exist. The first heat source is the shear zone where about 90% shear energy is changed to heat [13]. The second one is the friction zone between chip and race face of the cutting tool. The last one is the friction zone between machined workpiece surface and the relief face of the cutting tool. Heat lose includes the heat transferred away from the cutting zone through the convection and radiation besides the heat conducted into the workpiece and the cutting tool. The heat lost through radiation is small so that it can be ignored. Convection is the primary way of heat lost and is can be estimated by,

$$Q_c = h \cdot (T_s - T_{ambient})$$

where h is heat transfer coefficient, T_S is the surface temperature and $T_{ambient}$ is the ambient temperature. Heat transfer coefficient h can be estimated from the experimental data of temperature measurement.

3.2 Temperature measurement

Nowadays there are several experimental techniques available for the temperature measurement, such as using thermo-couples, IR pyrometers, and metallurgical inspection [14]. Thermo-couple method can be used to measure the average temperature or the peak temperature in a certain point. Metallurgical inspection method can only evaluate the temperature scope though the phase diagram. IR pyrometer method is simple and relatively accurate, but radiation from the measured surface is easy to be blocked in the experimental environment. In this experiment, the temperature distribution is required and the cold-air may not block the radiation from the surface measured, unlike cutting fluid. Therefore the IR pyrometer method is the best choice.

In the experiments, the uncoated carbide inserts SECO XOMX are selected. They are heated in the resistance furnace to about 900°C and then put on a bracket with asbestos net under it. Different cooling methods are applied with a NEC IR pyrometer recording the thermal images following the temperature goes down. The cooling methods include:

- Natural air;
- Cold-air with pressure 0.05MPa, temperature -20°C;
- Cold-air with pressure 0.1MPa, temperature -20°C;
- Cold-air with pressure 0.15MPa, temperature -20°C;

3.3 Calculation of heat transfer coefficients

The insert size is small and the steady-state one-dimensional heat transfer equation is employed to approximately estimate the temperature gradient from the equation [15]:

$$\frac{\partial^2 T}{\partial x^2} = 0, \quad T(x) = kx + T_0$$

where k is the steady-state one-dimension temperature gradient and T_0 is the temperature in the centre of temperature field, namely point O in Figure 6.

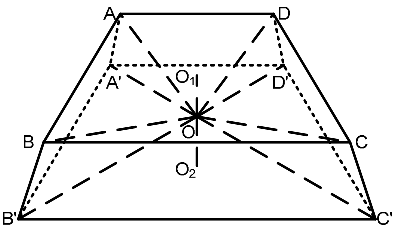
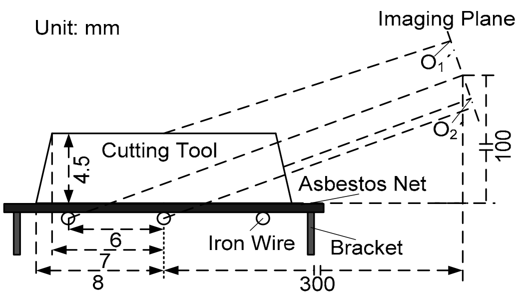


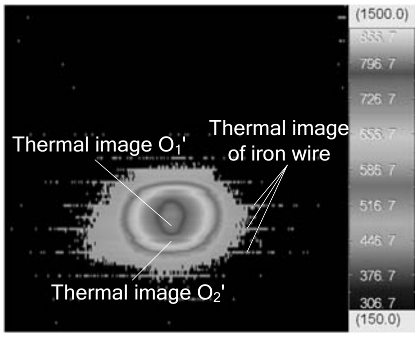
Figure 6: Insert model.

The insert is divided into six parts, polyhedron O-ABCD, O-A´B´C´D, O-ABB´A´, O-BCC´B´, O-CDD´C´, O-ADD´A´, as shown in Figure 6. Point O is the centre of insert and the temperature at point O is considered to be equal to T_0. O_1 is the projection of O on plane ABCD and O_2 is the projection of O on plane BCC´B´. The temperatures of points O_1 and O_2 are recorded as T_{O1} and T_{O2} to calculate the temperature gradients.

As shown in Figure 7(a), the thermal images of points O_1 and O_2 are at O_1' and O_2' respectively. It can be considered that image O_1' is the point with the highest temperature in the thermal image and the position of image O_2' can be defined by image distance $\overline{O_1'O_2'}$. Iron wires and their thermal images are used to mark the relationship between actual distance and image distance, as shown in Figure 7(b).



(a)



(b)

Figure 7: Experimental apparatus and thermal image.

Therefore the heat energy of the six parts can be calculated:

- The heat energy in O-ABCD is defined as:

$$Q_1(t_i) = \int_0^{\overline{OO_1}} c\rho V_1(x) T(x, t_i) dx$$

- The heat energy in O-ABB´A´, O-BCC´B´, O-CDD´C´ and O-ADD´A´ is same and defined as:

$$Q_2(t_i) = \int_0^{\overline{OO_2}} c\rho V_2(x) T(x, t_i) dx$$

- That no heat transfer in the surface A´B´C´D, polyhedron O-A´B´C´D is considered as an isothermal body and the temperature is the same as T_0. The heat energy in O-A´B´C´D can be defined as:

$$Q_3(t_i) = \int_0^{\overline{OO_1}} c\rho V_2(x) T_0(t_i) dx$$

where V_1 are the volume of polyhedron O-ABB´A´, O-BCC´B´, O-CDD´C´, O-ADD´A´, V_2 is the volume of O-ABCD and O-A´B´C´D, t_i is the cooling time, c is the specific heat and ρ is the density of the insert material.

The total heat energy is:

$$Q(t_i) = Q_1(t_i) + Q_2(t_i) + Q_3(t_i)$$

The heat transfer coefficient between cooling time t_1 and t_2 can be estimated as:

$$h\left(\frac{T_1 + T_2}{2}\right) = \frac{Q(t_1) - Q(t_2)}{A \cdot (t_2 - t_1) \cdot \left(\frac{T_1 + T_2}{2} - T_{ambient}\right)}$$

where A is the convection area, T_1 and T_2 are the surface temperatures in the cooling time t_1 and t_2, $T_{ambient}$ is the ambient temperature, which is approximately the outlet cold-air temperature in cold-air cutting and room temperature in dry cutting.

The results are shown in Figures 8~11.

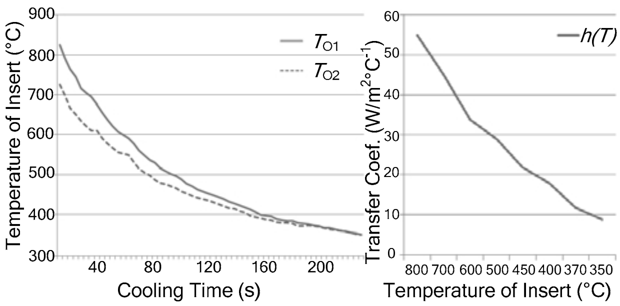


Figure 8: Temperature and heat transfer coefficient in natural cooling.

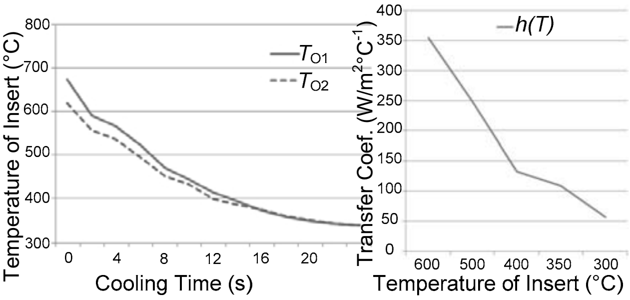


Figure 9: Temperature and heat transfer coefficient in cold-air cooling with pressure 0.05MPa, temperature -20°C.

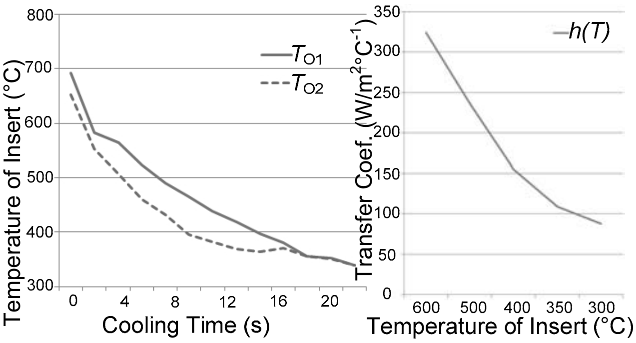


Figure 10: Temperature and heat transfer coefficient in cold-air cooling with pressure 0.1MPa, temperature -20°C.

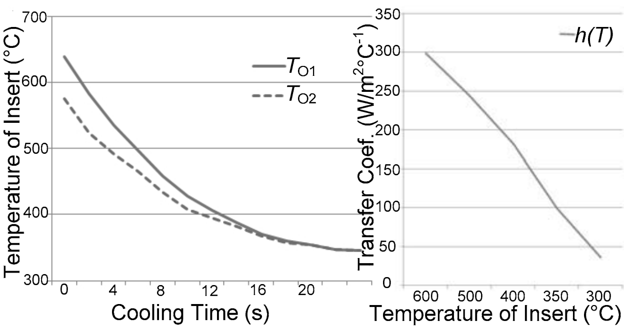


Figure 11: Temperature and heat transfer coefficient in cold-air cooling with pressure 0.15MPa, temperature -20°C.

3.4 Discussion

According to the principle of air compression, if the air pressure goes up, the flux goes up. Subsequently, the air velocity in the outlet increases and heat transfer coefficient increases [15]. However the results in this study show that the coefficients increase in some extent only when the temperature is high, such as 600°C or even higher. The reason may be that the air velocity in the outlet is high, up to about 60 m/s. The velocity change in this range has little influence on the thickness of boundary layer, and little influence on heat transfer coefficient too [15].

Because the blow speed has little influence on the heat transfer coefficient, higher pressure can be selected to increase the outlet pressure and decrease the velocity of air flow during transport, which may decrease the adhesion of oil mist and carry oil mist to the closer position of the cutting point.

4 PROCESS SIMULATION

4.1 2D simplification of milling process

Milling process is simplified to 2D cutting tool machining material with cutting speed $V_{cutting}$ and vertical speed $V_{vertical}$, as shown in Figure 12. The latter is used to simulate the feed motion and the vertical speed is expressed as:

$$V_{vertical} = \frac{doc \cdot V_{cutting}}{loc}$$

where *doc* is the depth of cut which is equal to the feed per tooth; *loc* is the length of cut which is equal to the contact arc of milling.

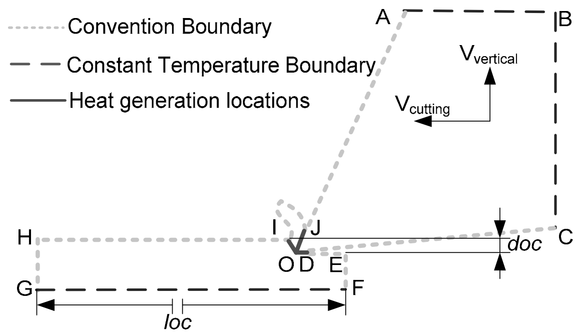


Figure 12: CAE model of simulation.

4.2 Thermal boundary conditions

The thermal boundary condition is important for cutting temperature evolution simulation. As shown in Figure 12, all the surfaces exposed to the environment are defined as the convection boundary with the heat transfer coefficient *h*, which is estimated above. Surfaces marked as AB and BC, respectively, are the cutting tool surfaces which contact with the tool holder. Surface marked as GF is the workpiece surface which contacts with the machine table. These three surfaces are defined as the constant temperature boundary with the room temperature 20°C. OI represents the heat source in the primary shear zone; OJ represents the heat source in the rake face; OD represents the heat source in the relief face.

4.3 Simulation results and its validation

With the finite element analysis software AdvantEdge, the temperature distribution is predicted, as shown in Figure 13.

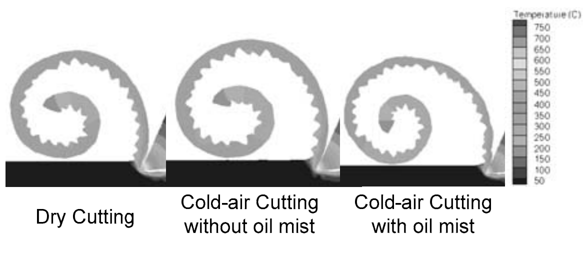


Figure 13: Temperature distribution.

Figure 14 shows the peak temperature in different cooling conditions.

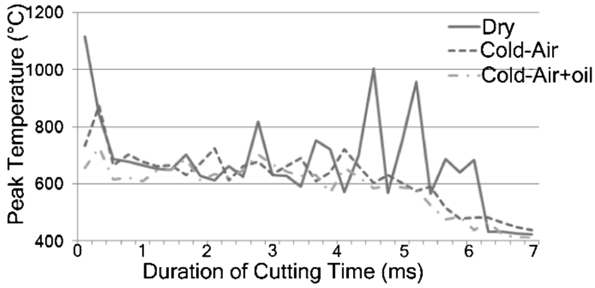
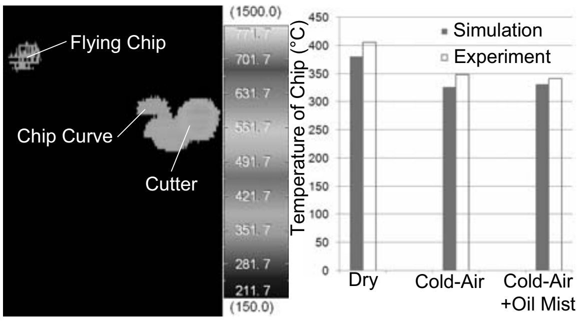


Figure 14: Peak temperature in different cooling conditions.

It can be seen in Figure 14 that the cold-air milling decreases the peak temperature significantly and makes the temperature change smoothly. One reason is that the heat transfer coefficient is many times in cold-air cutting as dry cutting. The higher temperature it is, the higher heat transfer coefficient becomes. Therefore the temperature changes smoothly. In addition, the oil mist decreases the friction between cutting tool and chip/workpiece, which decrease the cutting temperature.

Chip temperature is relatively constant in the cutting so it is measured by using IR pyrometer to validate the simulation results, as shown in Figure 15(a). The temperature comparisons of measurement and simulation are shown in Figure 15(b).



(a) (b)

Figure 15: Comparison of simulated and experimental results.

5 CONCLUSIONS

- Cutting forces decrease significantly with oil mist cooling, but change little with cold-air only. Oil mist can decrease the friction coefficient greatly but cold-air itself has little influence on the friction coefficient in milling Ti6Al4V.
- Heat transfer coefficient in cold-air cutting is many times larger than the one in dry cutting. But the cold-air pressure has little influence on the heat transfer coefficients.
- Cold-air cooling can decrease the peak temperature significantly and make the temperature change smoothly, which is the benefit for the tool life and surface quality.

6 ACKNOWLEDGMENTS

This work is supported by Fundamental Technology Foundation of Chinese Education Department (No. A1420060196), Specialized Research Fund for the Doctoral Program of Higher Education (No. 20070422033) and Scientific Research Foundation for the Outstanding Young Scientist of Shandong Province (No. 2007BS05001).

7 REFERENCES

[1] Su, Y., He, N., Li, L., Li, X.L., 2006, An experimental investigation of effects of cooling/lubrication conditions on tool wear in high-speed end milling of Ti6Al4V, Wear, 261:760-766.

[2] Yokogawa, K., Yokogawa, M., Okumura, S., 1998, Cooling air grinding and cutting technology, Tool Eng., 42/7: 57-62.

[3] Yamazaki, T., Miki, K., Sato, U., 2003, cooling air cutting of Ti-6Al-4V alloy, J. Jpn. Inst. Light Metals, 53/10:416-420.

[4] Ueda, T., Hosokawa, A., Yamada, K., 2006, Effect of oil mist on tool temperature in cutting, ASME J. Manuf. Sci. Eng., 128/2:130-135.

[5] Machado, A.R., Wallbank, J., 1997, The effect of extremely low lubricant volumes in machining, wear, 210:76-82.

[6] Wakabayashi, T., Suda, S., Terasaka, K., Musha, Y., Toda, Y., 2007, Triblogical action and cutting performance of MQL Media in Machining of Aluminium, CIRP Ann., 56/1:97-100.

[7] Espinosa, H. D., Patanella, A. J., Fischer, M., 2000, Dynamic friction measurements at sliding velocities representative of high-speed machining processes, ASME J. Manuf. Sci. Eng., 122/10:834-848.

[8] Hong, S.Y., Ding, Y., 2001, cooling approaches and cutting temperatures in cryogenic machining of Ti-6Al-4V, Int. J. Machine Tools Manufacture, 41/10:1417-1437.

[9] Liu, P.D., 1991, New developments in mechanics of cutting, Press of Dalian Technical University, Dalian, China.

[10] Merchant, M.E., 1945, Mechanics of the Cutting Process, Journal of Applied Physics, 16: 318-324.

[11] Oxley, P.L.B., 1989, Mechanics of machining, John Wiley & Sons, Inc., NY, United State.

[12] Rabinowicz, E., 1995, Friction and wear of materials (second edition), John Wiley & Sons, Inc., NY, United State.

[13] Albert, J.S., Henry, T.Y., 1993, Experiment and finite element predictions of residual stress due to orthogonal metal cutting, Int. J. Numerical Methods in Engineering, 36/9: 1487-1507.

[14] Lin, J., and Liu, C.Y., 2001, Measurement of cutting tool temperature by an infrared pyrometer, Meas. Sci. Technol. 12: 1243–1249.

[15] Holman, J.P., 2002, Heat transfer (ninth edition), McGraw-Hill Companies, Inc., MI, United State.

Environmentally Friendly Machining of Aluminum Using Minimal Quantity Lubrication System

Toshiaki Wakabayashi[1] and Satoshi Suda[2]

[1] Faculty of Engineering, Kagawa University, Japan

[2] Lubricants and Specialties Business Division, Nippon Oil Corporation, Japan

Abstract

Near-dry machining with the minimal quantity lubrication (MQL) system has been recognized as a representative environmentally friendly manufacturing operation. In machining of aluminum and its alloys, since they have highly adhesive characteristics compared with steels, effective lubrication is often necessary. Then, in order to understand the effectiveness of MQL system in machining of aluminum, this paper mainly deals with the experimental work where various MQL machining operations have been carried out using a MQL lubricant of synthetic ester. Further, this study discusses the connection between the MQL cutting performance of the ester and its tribological behaviour with different carrier gases.

Keywords:

Cutting; Minimal Quantity Lubrication; Environment

1 INTRODUCTION

Near-dry machining with the minimal quantity lubrication (MQL) system has been recognized as one of the representative environmentally friendly manufacturing operations [1, 2]. The MQL operation is particularly successful in machining of steels and the lubricating action of a very small amount of cutting fluids plays an important role [3-7]. In those applications, the tribological action of MQL lubricants is considerably important because they should be supplied to the cutting zone in a very small amount with the aid of atmospheric carrier gases. Under the circumstances, synthetic polyol esters, rather than vegetable based oils, have provided the optimal characteristics as MQL media [8-11].

In recent years, the application of aluminum and its alloys to automotive part materials is growing. In machining of them, since they have highly adhesive characteristics compared with steels, more effective lubrication is often necessary. Then, in order to understand the effectiveness of MQL system in machining of aluminum, this paper mainly deals with the experimental work where various cutting operations have been carried out using the MQL system with the synthetic esters. Further, this study discusses the connection between the resultant MQL cutting performance of those esters and their tribological behavior.

2 END MILLING OF ALUMINUM ALLOY

2.1 Experimental methods

Under the Cutting conditions shown in Table 1, this study first carried out end milling of aluminium workpiece and evaluated the cutting performance by using surface finish roughness, Ra, with various lubrication methods. Table 2 presents the specifications of the synthetic polyol ester used in MQL machining. This ester was developed as an optimal MQL lubricant and its successful MQL performance has already been well appreciated [8]. A water soluble type of 10% emulsion containing some sulfur-type extreme pressure (EP) additives was also chosen in the case of flood coolant supply, hereafter designated as WET.

Table 1: Cutting conditions of end milling.

Workpiece	JIS* AC8A aluminum alloy casting
Tool	JIS* SKH55 High speed steel Diameter = 10mm Number of flutes = 2 Helix angle = 30^{o}
Cutting speed	63, 110, 157 m/min
Feed	0.03, 0.06, 0.09 mm/tooth
Depth of cut	Axial: 10 mm, Radial: 3 mm
Mode	Down cutting
Lubrication	Dry, Air blow, MQL, Flood coolant (WET)
Fluid supply	MQL: 24-25 ml/h Air pressure: 0.4 MPa
	WET: 14 l/min
Evaluation	Surface finish roughness (Ra)

*Japanese Industrial Standard

Table 2: Specifications of synthetic polyol ester used in MQL machining.

Density (15 ^{o}C)	0.95 g/cm^3
Kinematic viscosity (40 ^{o}C)	19.1 mm^2/s
Biodegradability (CEC L-33-A-93)	100 %

2.2 Cutting performance evaluation

Figure 1 illustrates the results of the surface finish roughness at different cutting speeds with various lubrication methods. The ranking of the surface finish quality given by lubrication methods is not so largely changed if the cutting speed becomes higher: the best is for WET and the second best is MQL. These experimental facts demonstrate that the lubricating effects of the MQL ester and the WET coolant can improve the surface finish quality compared with air blow and dry operations. The possible reason why WET was better than MQL could be that the cooling action was further effective in improving the surface roughness. However, since WET contained some sulfur-type EP additives, one should also consider another possibility that the boundary lubricating effect of those EP additives worked effectively.

Figure 2 shows the results of the surface finish roughness at the different values of feed with various lubrication methods. The ranking of the surface finish quality given by lubrication methods is almost the same as that obtained in the case of different cutting speeds. There exists the tendency where the surface finish roughness increases as the feed increases except for the case of MQL, that is, MQL can maintain its cutting performance to provide the stable surface finish quality even when the feed becomes higher.

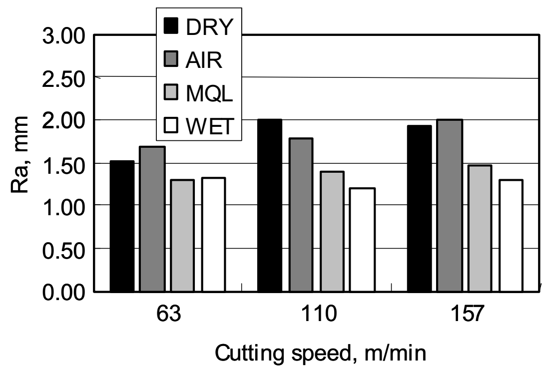


Figure 1: Surface finish roughness (Ra) at different cutting speeds with various lubrication methods.

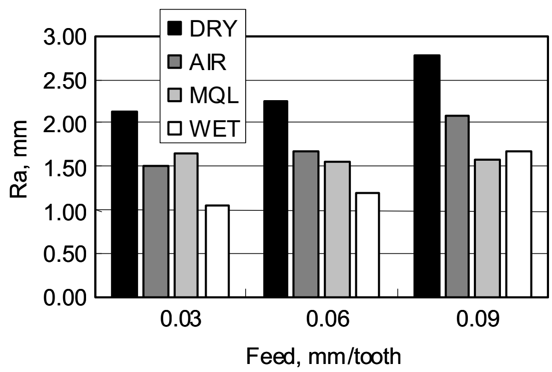


Figure 2: Surface finish roughness (Ra) at different feed values with various lubrication methods.

3 INFLUENCE OF CARRIER GASES ON THE ACTION OF MQL MEDIA

3.1 Experimental methods

Figure 3 illustrates one version of the controlled atmosphere cutting apparatus [10, 11]. This apparatus was prepared to examine the cutting performance in actual MQL machining with various carrier gases. The workpiece is attached to the spindle and rotates counter clockwise. A cutting tool is installed in a rod which is connected with the cutting force dynamometer. Since this dynamometer is mounted on a servomotor drive, the rod can give a radial feed motion to the tool. The mode of cutting is orthogonal and the conditions of this orthogonal cutting are presented in Table 3.

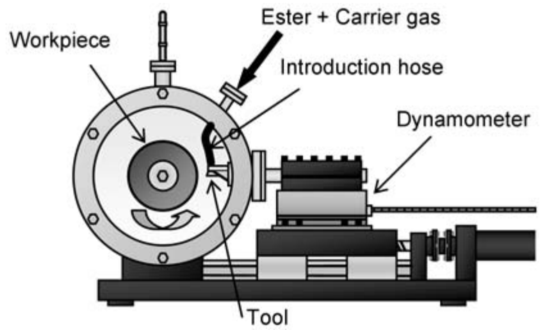


Figure 3: Controlled atmosphere cutting apparatus.

Table 3: Conditions of orthogonal cutting.

Workpiece	JIS A6061 aluminum alloy
Tool	Uncoated cemented carbide
Cutting speed	20 m/min
Depth of cut	0.1 mm

3.2 Importance of atmospheric carrier gas

Figure 4 provides the values of resultant cutting force obtained in the above orthogonal cutting. If one compares the results obtained by the same gas, MQL machining always presents the lower cutting force than gas blow alone and this means the ester provides preferable lubricating effects on the machining phenomena, probably because of its adsorption and the resultant lubricating film formation. Further, this figure obviously demonstrates the better cutting performance in the order of the lower oxygen concentration in the gas. Interesting thing is that the completely reversed results are obtained in machining of steels, that is, the better cutting performance in the order of the higher oxygen concentration in the gas [9].

Figure 5 is also the evidence that the combination of the lubricant ester and atmospheric carrier gases should be important to understand the MQL cutting performance. Metal transfer on the tool flank face is considerably large when the MQL ester is absent, whereas the adhesive characteristics of aluminium can greatly improved in the case of MQL. Again, the cutting situations are better with nitrogen than with oxygen.

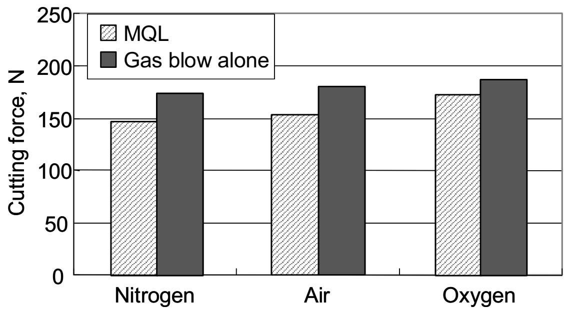


Figure 4: Cutting force obtained in orthogonal machining.

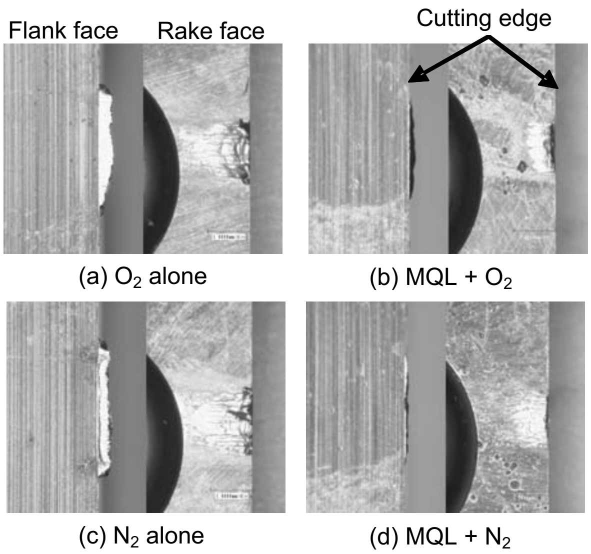


Figure 5: Observation of tool flank and rake faces with different atmospheres.

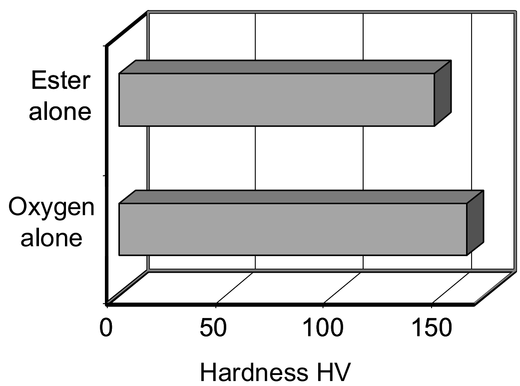


Figure 6: Surface hardness measured after machining.

In successful MQL machining of aluminum, therefore, the main point is how to exclude the influence of oxygen from the cutting atmosphere. The reason why the presence of oxygen is unfavorable may presumably be due to the formation of aluminum oxide (alumina) in the cutting zone. Since alumina is a highly hard material, it can make machining difficult. Figure 6 shows the results of surface hardness measured after machining and may support the influence of alumina formation.

Another important thing is that the introduction of synthetic polyol ester by MQL mist can always decrease the cutting force compared with the case of gas blow alone. This is true even when oxygen is supplied together, meaning the polyol ester can also work as an effective lubricant with a very small amount in the case of MQL machining of aluminum.

4 EFFECT OF LUBRICANT AND INFLUENCE OF CARRIER GASES IN DRILLING

4.1 Experimental methods

Practical MQL drilling was carried out using drill of 6 mm in diameter. The length of the hole (depth) was 50 mm. MQL mists were delivered through the center of the machine tool spindle and supplied to the cutting zone from the oil holes of the drill tip by a pressurized gas of 0.2-0.25 MPa. 100 holes were drilled in one evaluation test and the cutting performance was determined by means of thrust force during drilling. Table 4 presents the cutting conditions of this drilling test.

Table 4: Cutting conditions of drilling.

Workpiece	JIS AC8A aluminum alloy casting
Tool	Cemented carbide with DLC coating Diameter = 6 mm
Rotational speed	7000 rpm
Feed	0.1 mm/rev.
Hole depth	50 mm
Carrier gases	Air, O_2, N_2, Ar

4.2 Results of thrust force measurement

In drilling, compared with other open machining operations, since MQL mists are usually supplied to the cutting zone from the oil holes of the drill tip, the atmosphere inside the hole being drilled can easily become that of a carrier gas. Hence, the cutting phenomena in MQL drilling may strongly be influenced by the carrier gas.

Figure 7 shows the records of thrust force measurement during drilling of the last 20 holes, that is, between 81st. and 100th. drilling. As expected, the values of thrust force indicate the considerably successful results when oxygen is excluded from the atmosphere, and the influences of nitrogen and argon are the same. Further, when oxygen exists, regardless of its concentration in the atmosphere, significant negative thrust can be observed at the end of each drilling, probably because of poor lubrication, and such behavior is remarkably improved if no oxygen exists in the carrier gases.

5 CONCLUSIONS

Using various operations of machining aluminum, this study investigated experimentally the cutting performance of the environmentally friendly MQL system with a synthetic ester lubricant. In end milling, both MQL ester and WET coolant improved the surface finish quality compared with air blow and dry operations. MQL was also able to maintain its cutting performance to provide the stable surface finish quality even when the feed becomes higher. This study further demonstrated the importance of the action of atmospheric carrier gas in combination of the ester. In successful MQL machining of aluminium, in particular, the main point was how to exclude the influence of oxygen from cutting atmosphere.

6 ACKNOWLEDGMENTS

This work is supported in part by Grant-in-aid for Scientific Research (C) from Japan Society for the Promotion of Science.

7 REFERENCES

[1] Klocke, F., Eisenblatter, G., 1997, Dry Cutting, Annals of the CIRP, 46/2:519-526.

[2] Weinert, K., Inasaki, I., Sutherland, J. W., Wakabayashi, T., 2004, Dry Machining and Minimum Quantity Lubrication, Annals of the CIRP, 53/2:511-537.

[3] Heisel, U., Lutz, M., Spath, D., Wassmer, R., Walter, U., 1994, Application of Minimum Quantity Cooling Lubrication Technology in Cutting Processes, Production Eng., Vol.II/1:49-54.

[4] Wakabayashi, T., Sato, H., Inasaki, I., 1998, Turning Using Extremely Small Amounts of Cutting Fluids, JSME International J. (Series C), 41/1:143-148.

[5] Sutherland, J. W., Kulur, V. N., King, N.C., 2000, An Experimental Investigation of Air Quality in Wet and Dry Turning, Annals of the CIRP, 49/1:61-64.

[6] McCabe, J., Ostaraff, M. A., 2001, Performance Experience with Near-Dry Machining of Aluminum, Lubrication Eng., 57/12:22-27.

[7] Rahman, M., Kumar, A. S., Salam, M. U., 2002, Experimental Evaluation on the Effect of Minimal Quantities of Lubricant in Milling, International J. Machine Tool and Manufacture, 42:539-547.

[8] Suda, S., Yokota H., Inasaki I., Wakabayashi, T., 2002, A Synthetic Ester as an Optimal Cutting Fluid for Minimal Quantity Lubrication Machining, Annals of the CIRP, 51/1:95-98.

[9] Wakabayashi, T., Inasaki I., Suda, S., Yokota H., 2003, Tribological Characteristics and Cutting Performance of Lubricant Esters for Semi-dry Machining, Annals of the CIRP, 52/1:61-64.

[10] Min, S., Inasaki, I., Fujimura, S., Wakabayashi, T., Suda, S, 2005, Investigation of Adsorption Behaviour of Lubricants in Near-dry Machining, Proc. of IMechE, Part B: J. Eng. Manufacture, 219:665-671.

[11] Min, S., Inasaki, I., Fujimura, S., Wada, T., Suda, S., Wakabayashi, T., 2005, A Study on Tribology in Minimal Quantity Lubrication Cutting, Annals of the CIRP, 54/1:105-108.

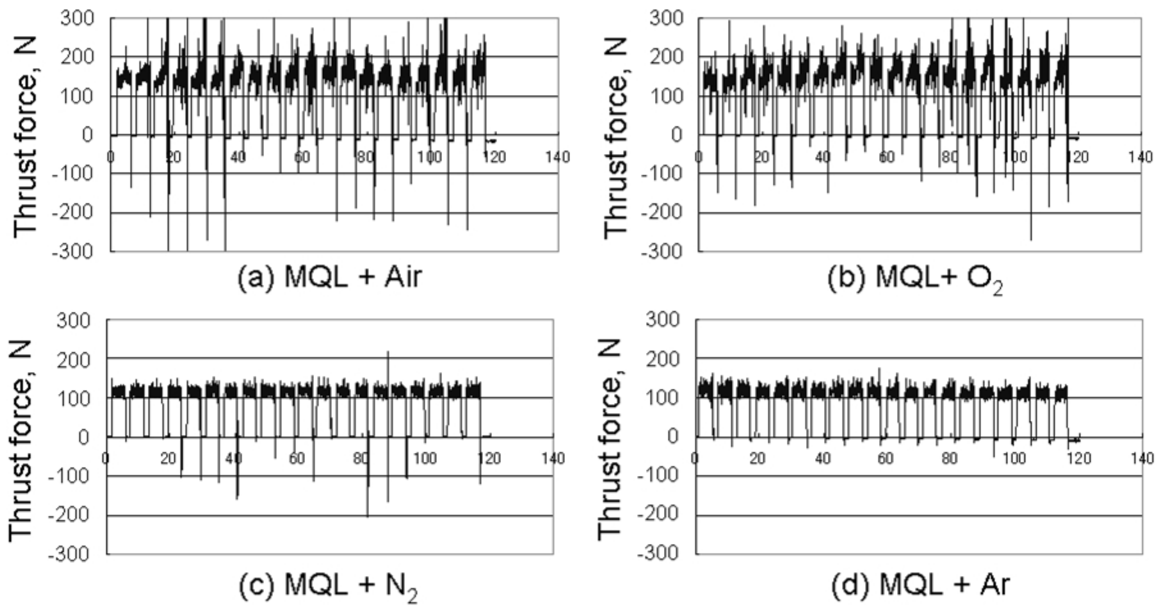


Figure 7: Results of thrust force measurement in drilling.

An Evolutionary Strategy for Optimizing Multi-pass Turning Operation under Dry Machining Conditions

Yi-Chi Wang and Yi-Che Chiu
Department of Industrial Engineering and Systems Management, School of Engineering, Feng Chia University, Taichung, Taiwan

Abstract
Determination of optimal cutting parameters for satisfying certain technological and economical conditions has been one of the most important elements in process planning for metal cutting operations. However, none of the earlier studies for solving such machining problems took into account the influence of the use of cutting fluids. In recent years, many researchers and experts have urged to eliminate the use of cutting fluids in machining operations to avoid casuing environmental problems. Machining without cutting fluids has then become a new trend on technological innovation. This paper presents an evolutionary strategy based optimization technique to optimize the cutting conditions in multi-pass turning processes under dry machining application.

Keywords:
Evolutionary Strategy ; Multi-pass Turning Operation ; Dry Machining

1 INTRODUCTION

One problem existing in metal cutting for nearly a century was to determine proper cutting conditions for a given machining operation. In practice, machining parameters for the cutting condition are usually determined based on the machine operator's experience or by following the cutting handbooks offered by the machine tool or equipment manufacturers. However, taking the constraints such as limited power, force, temperature, required surface finish, etc. into consideration the best cutting condition is difficult to determine. The decision concerning selection of the proper values for machining parameters such as cutting speed, feed rate, and depth of cut have significants influence on product quality, production rate, and the manufacturing costs. With wide application of CNC machines in many global manufacturers during the past few decades, optimizing the cutting conditions has become increasingly important since the CNC machines are much more expensive than the conventional ones and can not operate effectively without optimal cutting parameters.

In turning operations, a cutting process can possibly be completed with a single pass or multiple passes. Multi-pass turning is more preferable than single-pass turning for economic reasons in industry [1]. A multi-pass cutting operation involves in several roughing cuts and a single finishing cut. That makes the problem of determining the optimal cutting conditions more difficult and complicated. To obtain the optimal parameter cutting conditions for a particular multi-pass turning environment, developing a mathematical model for multi-pass turning operations has then become a useful tool for the determination of the optimal cutting conditions.

In metal cutting industry, cutting fluid is one of the most widely used agents. The two main purposes of using cutting fluid are to reduce the cutting forces by providing lubrication and to decrease temperature in cutting zone in order to increase tool life. However, the use of cutting fluid also brings many detrimental effects. Some of the fluids contain environmentally harmful or potentially damaging chemical constitutes. Some of them may be difficult to dispose and expensive to recycle and can even cause skin and lung disease to the operators [2]. A survey conducted in the Germany automotive industry illustrate that around 7%-17% of the unit manufacturing costs is attributable the deployment of cutting fluid, while the tool costs account for only 2%-4% of unit manufacturing costs [3]. Therefore, it seems to be the best option to eliminate cutting fluids usage.

In this paper, a new cutting model for dry turning processes is presented. An optimization algorithm based on the evolutionary strategy is proposed for solving such complicated non-linear machining problems. The proposed approach is first verified by comparing its solutions with the results of the other studies in the literatures and further demonstrated its searching capability in an illustrated example.

2 MULTI-PASS TURNING OPERATION

Shin and Joo [4] developed a mathematical model for multi-pass turning operations. They provided an example and proposed an approach combining the Fobonacci search and dynamic programming to solve it. Gupta et al. [5] solved the same illustrative problem with two stages. The first step is to calculate the minimum production cost for roughing and finishing operations for various depths of cut. The second step is to determine the optimal combination of depth of cut for roughing and finishing operations. Their approach determines a production cost which is substantially less than that using Shin and Joo's method. Alberti and Perrone [6] suggested a multi-objective possibilistic programming model in which the optimal solution is obtained by using a genetic algorithm. They compared their results with the results obtained by Gupta et al. [5] for the same examples. Their results were never worse than Gupta's, but in two of the six examples their solutions show a very tiny violation of the cutting force constraint.

Chen and Tsai [1] extended the multi-pass turning operation model of Shin and Joo by adding seven more constraints. They also proposed an approach that combined the simulated annealing algorithm and a pattern search technique for solving the extended model. Onwubulu and Kumalo [7, 8] proposed a technique based on a genetic algorithm to determine the optimal machining parameters for the extended model of Chen and Tsai. They arrived at a production cost, which was lower than that of Chen and Tsai [1]. However, it must be noted here that the optimal value obtained by Onwulolu and Kumalo [7, 8] has been argued impractical by Chen and Chen [9]. They pointed out that Onwubolu and Kumalo incorrectly handled the machining model since the number of rough cuts in their method is not limited to an integer value. For the same problem, Vijayakumar et al. [10] proposed an ant colony optimization method and claimed that their ant colony based approach found an even better solution with less runs than either the approach of Chen and Tsai or the GA-based approach of Onwubulu and Kumalo. However, their solution later has been proven invalid [11].

3 CUTTING PROCESS MODEL

In the model of Chen and Tsai [1], the minimum unit production cost criterion is adopted as the objective. The production cost for machining one unit piece is represented by the sum of the following four terms:

Minimize $UC = C_M + C_I + C_R + C_T$ (1)

$$C_M = k_0\left[\frac{\pi DL}{1000V_r f_r}\left(\frac{d_t - d_s}{d_r}\right) + \frac{\pi DL}{1000V_s f_s}\right] \quad (2)$$

$$C_I = k_0\left[t_c + (h_1 L + h_2)\left(\frac{d_t - d_s}{d_r} + 1\right)\right] \quad (3)$$

$$C_R = k_0 \frac{t_e}{T_P}\left[\frac{\pi DL}{1000V_r f_r}\left(\frac{d_t - d_s}{d_r}\right) + \frac{\pi DL}{1000V_s f_s}\right] \quad (4)$$

$$C_T = \frac{k_t}{T_P}\left[\frac{\pi DL}{1000V_r f_r}\left(\frac{d_t - d_s}{d_r}\right) + \frac{\pi DL}{1000V_s f_s}\right] \quad (5)$$

In this study, we modified the model of Chen and Tsai [1] by taking serious consideration about the dry cutting effects. As mentioned before, one of the main functions of cutting fluids is to reduce the cutting force and the required power. Instead of using the stead cutting force model depending only on feed rate and depth of cut, we adopt a more complex and precise cutting force model developed by Thomas and Beauchamp [12]. This cutting force model Eq.(11 & 17) takes into account the steady, random, and harmonic forces acting on the cutting tool for the dry turning. It explains 99.4% of the variability in the steady cutting force. In addition, another important function of the cuttings fluids is to extend the tool life. El-Hossainy [13] conducted a study of investigating the effects of using cutting fluid on the Taylor's tool life equation constants (C, n). According to his findings, the use of cutting fluids can increase the constant C by 18% but reduce the constant n by 7%, which overall leads to a percentage in crease in tool life of about 214%. We adopted the conclusion and modified the values of the constants (C_0, *p, q, r*) in the following equations:

$$T_r = \frac{C_0}{V_r^{\,p} f_r^{\,q} d_r^{\,r}}, \quad T_s = \frac{C_0}{V_s^{\,p} f_s^{\,q} d_s^{\,r}} \quad (6)$$

Table 1. Notations of the Proposed Optimization Approach.

Q_r Q_s	chip-tool interface temperatures during rough and finish machining.
Q_U	maximum allowable chip-tool interface temperature.
k_2 τ ψ δ	constants related to equation of chip-tool interface temperature.
k_3 k_4 k_5	constants for roughing and finishing parameter relations, k_3 , k_4 , $k_5 \geqq 1$
SC	limit of stable cutting region constraint.
λ ν	constants related to expression of stable cutting region.
UC	unit production cost not including material cost ($/piece).
$K_1 \mu \nu \kappa$ $\varepsilon \omega$	constants of cutting force equation.
P_U η	maximum allowable cutting power (kw) & power efficiency.
P_r P_s	cutting powers during roughing and finishing (kw).
F_v F_r F_s	maximum allowable cutting force, cutting forces during roughing and finishing.
L_T R	length of the cutting tool, nose radius of the cutting tool.
SR	maximum allowable surface roughness.
p q r C_0	constants of tool-life equation.
T_U T_L T_p	upper and lower bounds of tool life (min), tool life determined by the operator.
T_r T_s	expected tool life for roughing, and expected tool life for finishing.
h_1 h_2	constants pertaining to tool travel and approach/depart time.
t_e t_c	time required to exchange a tool, total machine idle time.
k_t k_0	cutting edge cost ($/edge), direct labor cost & overhead ($/min).
D L d_t	diameter and length of workpiece (mm), depth of metal to be removed.
V_r V_s	cutting speeds in rough and finish machining.
V_L V_U	lower and upper bounds of cutting speed in machining.
f_r f_s	feed rates in rough and finish machining.
f_L f_U	lower and upper bounds of feed rate in machining.
d_r d_s	depths of cut for each pass of rough and finish machining.
d_L d_U	lower and upper bounds of depth of cut.

For each decision variable, there are upper bound and lower bound that limit the cutting condition within a feasible range. Besides, the tool life is expected to be within a certain range. The cutting force should not exceed the maximum force to avid chatter or any deflection on work part or tool. The cutting power required can not be higher than the maximum power of the machine tool. Totally, six constraints are considered for the rough cutting operation.

$$V_{rL} \le V_r \le V_{rU} \quad (7)$$

$$f_{rL} \le f_r \le f_{rU} \quad (8)$$

$$d_{rL} \le d_r \le d_{rU} \quad (9)$$

$$T_L \le T_r \le T_U \quad (10)$$

$$F_r = \frac{k_1 \times f_s{}^{\mu} \times d_s{}^{\nu} \times L_T{}^{\varpi}}{r^{\kappa} \times V_s{}^{\varepsilon}} \le F_U \quad (11)$$

$$\frac{F_r V_r}{6120\eta} \le P_U \quad (12)$$

The other six constraints plus another constraint for surface roughness are considered for the finish cutting operation

$$V_{sL} \le V_s \le V_{sU} \quad (13)$$

$$f_{sL} \le f_s \le f_{sU} \quad (14)$$

$$d_{sL} \le d_s \le d_{sU} \quad (15)$$

$$T_L \le T_s \le T_U \quad (16)$$

$$F_s = \frac{k_1 \times f_s{}^{\mu} \times d_s{}^{\nu} \times L_T{}^{\varpi}}{r^{\kappa} \times V_s{}^{\varepsilon}} \le F_U \quad (17)$$

$$\frac{F_s V_s}{6120\eta} \le P_U \quad (18)$$

$$\frac{f_s{}^2}{8R} \le SR_U \quad (19)$$

Stable cutting region constraints:

$$(V_r)^{\lambda} f_r (d_r)^{\nu} \ge SC \,,\; (V_s)^{\lambda} f_s (d_s)^{\nu} \ge SC \quad (20)$$

Chip-tool interface temperature constraints:

$$k_2 (V_r)^{\tau} (f_r)^{\phi} (d_r)^{\delta} \le Q_u \,,\; k_2 (V_s)^{\tau} (f_s)^{\phi} (d_s)^{\delta} \le Q_u \quad (21)$$

Constraints for roughing and finishing parameter relations:

$$V_s \ge k_3 V_r \;,\; f_r \ge k_4 f_s \,,\; d_r \ge k_5 d_s \quad (22)$$

4 EVOLUTIONARY STRATEGY

The evolutionary strategy was initially proposed by Rechenberg and Schwefel in the 1960s. The classic ES consists of several key operators: initialization, selection, recombination, and mutation. The starting population is initialized by an algorithm-dependent method, and evolves towards successively better regions of the search space by means of (more or less) randomized processes of selection, recombination, and mutation [14]. The ES was categorized as one of the three evolutionary algorithms: genetic algorithm (GA), evolutionary programming (EP), and evolutionary strategy. The GA may be the most popular algorithm among these three. However, one of the advantages of the ES is the use of real values for the decision variables, rather than the binary coding favored in GA implementations. In addition, recombination (crossover) in GA is emphasized more than mutation, while recombination is not applied at all in GP. In ES both operators are key elements for successful implementation.

The first step of the proposed ES is to define the initial parent population. Solving this problem involves searching the best values the six decision variables (V_r, f_r, d_r, V_s, f_s, d_s) which result in the lowest unit production cost. Each individual parent can be randomly generated as a vector with 12 elements representing the values of the 6 decision variables and their corresponding standard deviation, σ. For example,

$$V_r = V_{rL} + U(0,1)(V_{rU} - V_{rL}) \quad (23)$$

$$\sigma_r = \left| V_r - (V_{rL} + \frac{V_{rU} - V_{rL}}{2}) \right| \frac{1}{\sqrt{n}} \quad (24)$$

Where $U(0,1)$ denotes a random variable of uniform distribution in the interval [0,1]. The initialization of the standard deviation is based on the approach of Franco et al. [15].

Selection in the proposed ES is completely deterministic, selecting the μ best individuals out of the set of λ offspring individuals ((μ, λ)-selection) or out of the union of parents and offspring ((μ+ λ)-selection). Based on the suggestions of Bäck and Schewefel [16], the proposed ES employs the ((μ, λ)-selection) in this case.

There are two ways commonly used for recombination in ES: “discrete recombination” sometimes referred to as “dominant recombination” and the “intermediate recombination.” In discrete recombination, the features of individual offspring may come intact or mutated from one parent or the other. While in intermediate recombination, the features of individual offspring are computed as a kind of average of the two parents’ feature. The intermediate recombination in their global form as proposed by Back and Schwefel [16] is adopted in the proposed ES.

After recombining two selected parents to form a new individual. This individual has to be mutated to yield an offspring. The entire mutation procedure consists of two stages: mutating the standard deviations first and then mutating the decision variables.

$$\sigma_r{}' = \sigma_r \cdot \exp(\tau' \cdot N(0,1) + \tau \cdot N_r(0,1)) \cdot (1/g_a) \quad (25)$$

$$V_r{}' = V_r + N(0, \sigma_r{}') \quad (26)$$

According to Back and Schwefel [16], the values for the parameters τ and τ' are set as follows:

$$\tau = \left(\sqrt{2\sqrt{n}}\right)^{-1}, \quad \tau' = \left(\sqrt{2n}\right)^{-1} \quad (27)$$

The value of g_a is set to be 1 for the first 100 generations. After that, g_a is increased by 1 every 100 generations. With such idea, the proposed ES will move its search with the larger “step size” in the early stages to avoid being trapped in the local optimal solutions. The searching step size later is narrowed to prevent jumping out the good solution area.

5 NUMERICAL RESULTS

The proposed approach is used to solve the problems of the other studies [4]. A lower unit production cost can be reached than the ones obtained in Shin and Joo’s study [4]. The best solution found by the proposed method for Chen and Tsai’s model [1] is only $0.015/piece higher than the best one found by them. That demonstrates the potential of the proposed approach. In the proposed cutting model, the optimum cutting conditions for each pass in a turning operation will be determined to demonstrate how the proposed ES performs. A cemented carbide cutting tool with throw-away inserts is used in this case. The data for this case is presented in Table 2.

After running through a thousand generations of the proposed ES(μ, λ)-(10, 70) for various depth of cuts, the solutions obtained are shown in Table 3.

Table 2. Data for the Multi-pass Dry Turning Operation Case.

$d_{rL} = 1$ mm	R =1.59mm	$h_1 = 7 \times 10^{-4}$
$d_{rU} = 3$ mm	$T_L = 25$ min	SC = 140
$d_{sL} = 1$ mm	$T_U = 45$ min	$\delta = 0.105$
$d_{sU} = 3$ mm	$Q_U = 900^\circ$ C	$\phi = 0.2$
$k_t = 2.5\ \text{\$/edge}$	$k_0 = 0.5\ \text{\$/min}$	r = 0.6975
L = 300 mm	$P_U = 200$ kw	q = 1.6275
$d_t = 8$ mm	$F_U = 5$ kgf	p = 4.65
$k_3 = 1$	$\upsilon = -1$	$k_4 = 2.5$
$\lambda = 2$	$\tau = 0.4$	$\eta = 0.85$
$V_{rL} = 50$ m/min	$f_{rL} = 0.1$ mm/rev	$k_2 = 132$
$V_{rU} = 500$ m/min	$f_{rU} = 0.9$ mm/rev	$k_5 = 1$
$V_{sL} = 50$ m/min	$f_{sL} = 0.1$ mm/rev	$k_1 = 1677$
$V_{sU} = 500$ m/min	$f_{sU} = 0.9$ mm/rev	$h_2 = 0.3$
$\varepsilon = 0.08$	$\varpi = 0.05$	$S_r = 10$ μm
$t_c = 0.75$ min/piece	$t_e = 1.5$ min/edge	D =50mm
$\mu = 0.8$	$\nu = 0.96$	$\kappa = 0.07$
L_T = 48.74mm	$C_0 = 41632196900$	

6 CONCLUSIONS

This paper presents a model for multi-pass turning operations under the circumstances of dry machining. Unlike most models in the literature, the mathematical model contains a comprehensive set of practical machining constraints especially focusing on dry cutting conditions in which several key elements such as the cutting force, cutting power, tool-chip interface temperature, and more importantly the tool life are significantly influenced. An evolutionary strategy optimization approach is proposed for solving this problem. The proposed approach demonstrates its potential by comparing with the results of the other cases in literature. Various depths of cut for the dry turning operations are then used to verify the ease of its applications.

7 REFERENCES

[1] Chen, M.-C., Tsai, D.-M., 1996, A Simulated Annealing Approach for Optimization of Multi-pass Turning Operations, International Journal of Production Research, 34/10:2803-2825.

[2] Sreejith, P.S., 2008, Machining of 6061 Aluminium Alloy with MQL, Dry and Flooded Lubricant Conditions, Material Letters, 62/2:276-278.

[3] Klocke, F., Eisenblätter, G., 1997, Dry Cutting, Annals of the CIRP, 46/2:519-526.

[4] Shin, Y.C., Joo, Y.S., 1992, Optimization of Machining Conditions with Practical Constraints, International Journal of Production Research, 30/12:2907-2919.

[5] Gupta, R., Batra, J.L., Lal, G.K., 1995, Determination of Optimal Subdivision of Depth of Cut in Multipass Turning with Constraints, International Journal of Production Research, 33/9:2555-2565.

[6] Alberti, N., Perrone, G., 1999, Multipass Machining Optimization by Using Fuzzy Possibilistic Programming and Genetic Algorithms, Proceedings of the Institution of Mechanical Engineers Part B, 213:261-273.

[7] Onwubolu, G.C., Kumalo, T., 2001, Multi-pass Turning Operations Optimization Based on Genetic Algorithms, Proceedings of the Institution of Mechanical Engineers Part B, 215:117-124.

[8] Onwubolu, G.C., Kumalo, T., 2001, Optimization of Multipass Turning Operations with Genetic Algorithm, International Journal of Production Research, 39/16:3727-3745.

[9] Chen, M.-C., Chen, K.-Y., 2003, Optimization of Multipass Turning Operations with Genetic Algorithms: A Note, International Journal of Production Research, 41/14:3385-3388.

[10] Vijayakumar, K., Prabhaharan, G., Asokan, P., Saravanan, R., 2003, Optimization of Multi-pass Turning Operations using Ant Colony System, International Journal of Machine Tools and Manufacture, 43/15:1633-1639.

[11] Wang, Y.-C., 2007, A Note on 'Optimization of Multi-pass Turning Operations using Ant Colony System', International Journal of Machine Tools and Manufacture, 47/12-13:2057-2059.

[12] Thomas, M., Beauchamp, Y., 2003, Statistical Investigation of Model Parameters of Cutting Tools in Dry Turning, International Journal of Machine Tools and Manufacture, 43/11:1093-1106.

[13] El-Hossainy, T.M., 2001, Tool Wear Monitoring under Dry and Wet Machining, Materials and Manufacturing Processes, 16/2:165-176.

[14] Bäck, T., 1996, Evolutionary Algorithms in Theory and Practice, Oxford University Press.

[15] Franco, G., Betti, R., Lus, H., 2004, Identification of Structural Systems using an Evolutionary Strategy, Journal of Engineering Mechanics, 130/10:1125-1139.

[16] Bäck, T., Schwefel, H.-P., 1993, An Overview of Evolutionary Algorithms for Parameter Optimization, Evolutionary Computing, 1/1:1-23.

Table 3: Results for the Dry Cutting Model with the proposed ES(10, 70)

Depth of cut	Nr	V_r (m/min)	f_r (mm/rev)	d_r (mm)	V_s (m/min)	f_s (mm/rev)	d_s (mm)	\$/piece
6 mm	2	99.0083	0.684092	2	136.443	0.27363	2	2.8118
7 mm	2	103.071	0.570853	2.33333	142.041	0.22834	2.333	3.0644
8 mm	2	106.725	0.488031	2.66667	147.076	0.19521	2.667	3.3140
9 mm	3	102.098	0.595752	2.25	140.7	0.23830	2.25	3.7447
10 mm	3	104.943	0.526442	2.5	144.621	0.21058	2.5	3.9819

Machines and Forming

Proposal of Real-Time Balancing mechanism Using Magnetic Fluid for Machine Tool Spindle

Keiichi Nakamoto[1], Kazuhiko Adachi[2] and Keiichi Shirase[2]

[1] Department of Mechanical Engineering, Graduate School of Engineering, Osaka University, Japan

[2] Department of Mechanical Engineering, Graduate School of Engineering, Kobe University, Japan

Abstract

Spindle vibration of machine tools causes premature tool wear and deterioration of machining accuracy. However, the vibration changes depending on the used tooling system and the rotational speed during various machining operations. In this study, a smart spindle with a real-time balancing mechanism using a magnetic fluid is proposed. The magnetic fluid is enclosed inside the rotational part, and the distribution of the circumferential fluid mass is actively controlled. A test spindle with the real-time balancing mechanism was designed and developed. The experimental results proof the feasibility of the proposed real-time balancing mechanism for achieving high accuracy and efficiency machining.

Keywords:

Machine Tool Spindle, Real-Time Balancing, Magnetic Fluid, Smart Spindle, Vibration Control

1 INTRODUCTION

Machining of high precision parts has been increasingly demanded and the life cycle of products has been rapidly shortened. Therefore, it is always required for machine tools to achieve high accuracy and efficiency machining. Especially, machine tool spindle is the key component affecting the performance of machine tools. In order to realize high speed machining, some development approaches of machine tool spindles have been proposed [1, 2].

In high speed machining, even a small unbalance weight of the spindle makes large whirl. Spindle vibration of machine tools causes premature tool wear and deterioration of machining accuracy. On the other hand, rotor dynamics of industrial machines has been studied [3] and already standardized [4, 5]. However, the vibration of machine tool spindles changes depending on the used tooling system and the rotational speed during various machining operations. Due to this complexity, a real-time balancing mechanism of a machine tool spindle has never realized successfully.

In order to reduce unbalance response of the spindle for achieving high accuracy and efficiency machining, a smart spindle with a real-time balancing mechanism using magnetic fluid is proposed in this study. Magnetic fluid is a type of smart fluid, which is a stable colloidal suspension of magnetic nano particles in a liquid carrier. Its viscosity and behavior can be controlled accurately by varying the magnetic field intensity. The magnetic fluid is enclosed inside the rotational part, and the distribution of the circumferential mass is actively controlled to cancel the unbalance weight of the machine tool spindle.

In this paper, a real-time balancing mechanism using magnetic fluid is proposed. Then, a developed test spindle for the evaluation of the real-time balancing by controlling the magnetic fluid is described. The feasibility of the proposed real-time balancing of machine tool spindles is discussed based on the obtained experimental results.

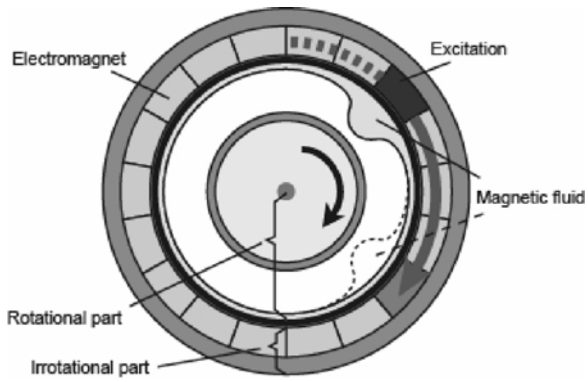


Figure 1: Proposed real-time balancing using magnetic fluid.

2 PROPOSED REAL-TIME BALANCING

It is expected that machine tool spindles are rotated with high accuracy. Therefore, unbalance weight of machine tool spindles is carefully compensated in the field site in order to assure the rotational accuracy. However, in practice, the influence of the used tooling system and the rotational speed can not be ignored during various machining operations. Spindle vibration changes significantly depending on the situation of the rotational part.

A smart spindle with a real-time balancing mechanism using magnetic fluid is proposed in this study. Figure 1 illustrates the overview of the proposed real-time balancing mechanism. The magnetic fluid is enclosed inside the rotational part of the spindle, and electromagnets are located in the stationary (irrotational) part. The magnetic fluid is drawn into an arbitrary position depending on the magnetic field varied by the surrounding electromagnets. Then, the distribution of the circumferential mass of the magnetic fluid is actively

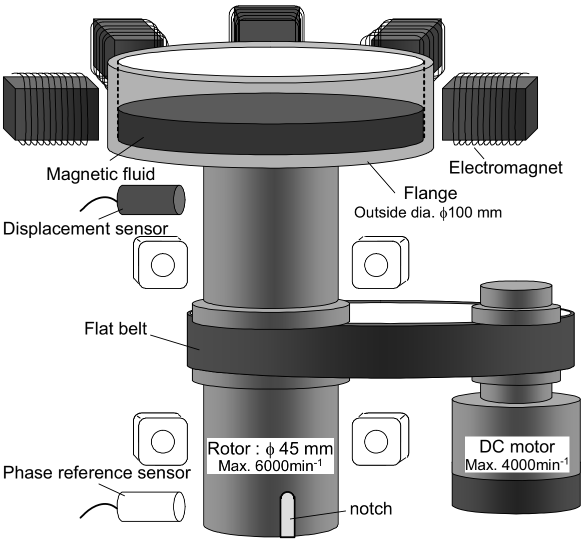


Figure 2: Outline drawing of developed test spindle.

controlled to be the counter weight and to cancel the unbalance weight in synchronization with the spindle rotation.

In previous researches, in order to improve the dynamic characteristic of machine tool elements, smart fluids such as electrorheological fluid have been employed [6 - 8]. The aspect of smart fluids varying the viscosity is applied to induce high damping forces in these approaches. In this study, the changeable mass distribution of the fluid is effectively used to compensate the unbalance weight of the spindle. Moreover, an oscillation force caused by changing the mass distribution may act like a cutting force in irregular tooth pitch end milling and avoid chatter vibration by disturbing the chatter frequency.

3 DEVELOPED TEST SPINDLE

A test spindle with the proposed real-time balancing mechanism was designed and developed in order to experimentally evaluate the real-time balancing controlling the magnetic fluid. Various experiments by using the developed test spindle are conducted to verify the feasibility of a machine tool spindle with the real-time balancing mechanism.

Figure 2 shows the outline drawing of the developed test spindle in this study. The spindle height is 340 mm, and its diameter is 45 mm. The spindle rotates up to 6000 min^{-1} on a DC motor through a flat belt. The upper flange, in which magnetic fluid is enclosed, is made of nonmagnetic stainless steel and has 100 mm diameter (Inside dia. 97 mm). The flange is surrounded by 16 electromagnets for controlling the distribution of the circumferential mass of the magnetic fluid. The unbalance response of the spindle is detected with an eddy current displacement sensor located below the flange. The terminal part of the spindle is notched to detect the rotational phase reference for detecting the position of unbalance weight. The phase angle of unbalance weight is defined in a counterclockwise direction viewed from the flange based on the reference signal (0 deg.). Also, tapped holes are arranged at 22.5 deg. intervals in the flange in order to forcibly add unbalance weight.

The measured unbalance response is analyzed using "Vector Monitor VM-13V1(SHINKAWA Sensor Technology, Inc.)",

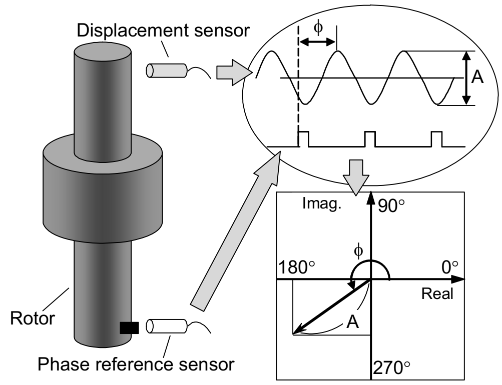


Figure 3: Unbalance response vector.

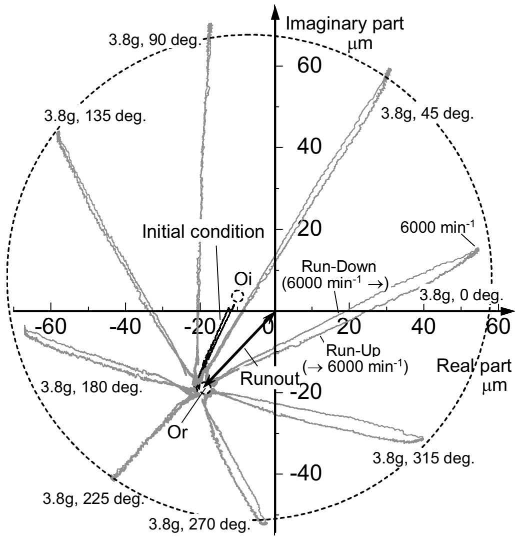


Figure 4: Measured unbalance response vector.

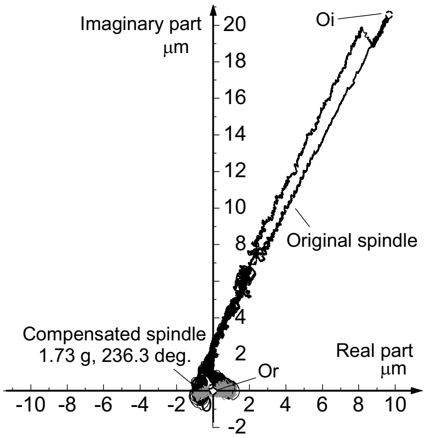


Figure 5: Compensation of initial unbalance weight.

which diagnoses the dynamic characteristics of rotating machinery. As shown in Figure 3, it receives the phase reference of the rotating spindle (one pulse/revolution) and detected vibration signals, outputs unbalance response vector (amplitude and phase angle) values in real-time.

Initial unbalance weight of the spindle is identified prior to verification experiments of the proposed real-time balancing. At first, the rotational speed is continuously changed with diagnosing the unbalance response vector from 0 min^{-1} to

Table 1: Specifications of evaluated magnetic fluids.

	Saturation magnetization at 25°C (mT)	Viscosity at 27°C (mPa·s)	Density at 25°C (g/cm³)
A	11.3	98.9	0.936
B	10.6	56.5	1.047
C	62.0	5.3	1.369

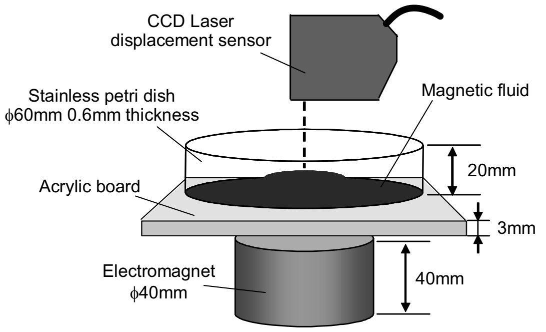


Figure 6: Characteristic test of magnetic fluids.

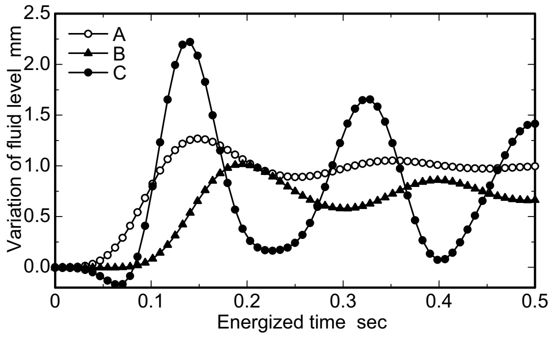


Figure 7: Comparison of response of magnetic fluids.

6000 min^{-1}, and then the rotational speed is returned to 0 min^{-1} in the same way. A known trial mass (3.8 g) is sequentially replaced to the original test spindle at 45 deg. intervals in the flange. Figure 4 shows the measured unbalance response vector. In this figure, the results obtained on low revolution conditions are eliminated because runout strongly affects the values. Runout is caused by machined error or residual magnetism of the spindle, the compensation is not considered in this study.

In Figure 4, the distance between the origin 0 and O_r, which is the start and end point of the unbalance response vector under the whole conditions, indicates the runout component. The edge points (6000 min^{-1}) of the whole conditions make the circle. Its center becomes O_i, which is the edge point of the original spindle condition. The radius corresponds to the impact causd by the trial mass (3.8 g). Based on these results, the initial unbalance mass is identified as 1.73 g at 56.3 deg.

In order to verify the identified initial unbalance mass, the calculated counter balance mass 1.73 g at 236.3 deg. is added to the flange. The measured unbalance response vector of the compensated spindle is shown in Figure 5. In this figure, the origin of the unbalance response vector is offset so that the runout component is eliminated. It is obviously recognized that the unbalance response of the compensated spindle is considerably reduced compared to that of the original spindle in the whole speed range.

The circumferential mass of magnetic fluid should be controlled rapidly to realize the proposed real-time balancing. Therefore, the distinguishing characteristics of 3 kinds of magnetic fluids are investigated in this study. Table 1 lists the specifications of prepared magnetic fluids. As illustrated in Figure 6, the variation of the fluid level of the magnetic fluids excited by an electromagnet (KANETEC KE-4B) is measured with a laser displacement sensor (KEYENCE LK-G30) in the characteristic test.

Figure 7 shows the comparison of the measured fluid levels. Each volume of magnetic fluid is 20 ml and the input voltage is 12 V. It is recognized that magnetic fluid C, which has the highest saturation magnetization, responds quickly to the changed magnetic field intensity. Consequently, magnetic fluid C is adapted for achieving the real-time balancing mechanism controlling the counter weight.

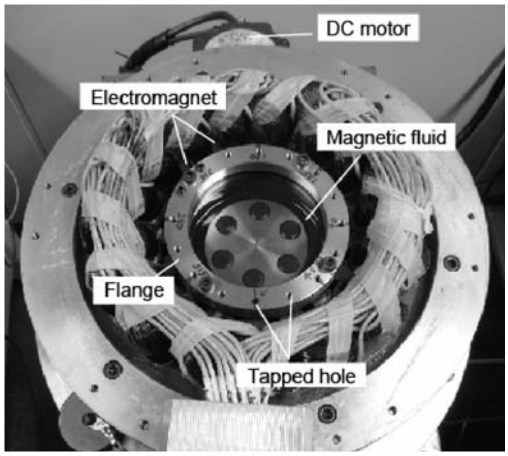
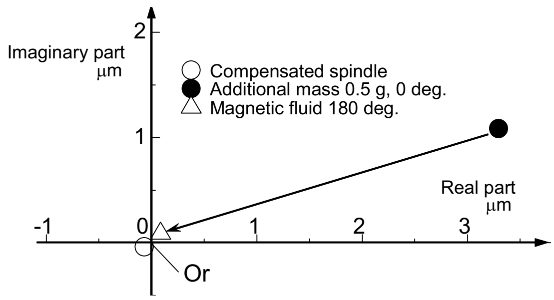
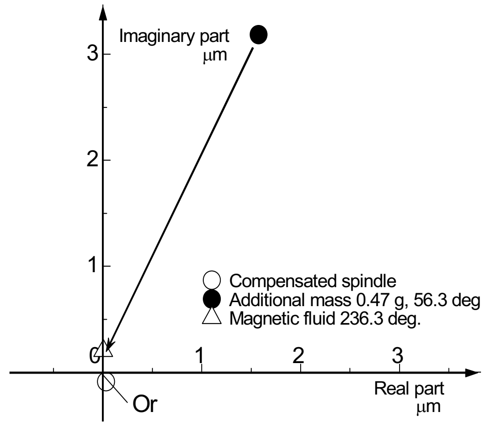


Figure 8: Developed test spindle for verification experiments.



(a) Added mass 0.5 g, 0 deg.



(b) Added mass 0.47 g, 56.3 deg.

Figure 9: Unbalance response in experiment I (4200 min^{-1}).

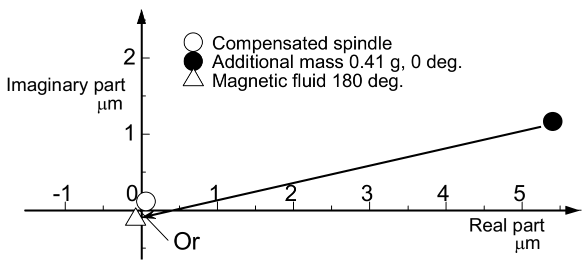


Figure 10: Unbalance response in experiment II (6000 min^{-1}).

4 VERIFICATION EXPERIMENTS

By using the developed test spindle, verification experiments of the proposed real-time balancing mechanism are conducted. Figure 8 shows the appearance of the developed test spindle for the verification experiments. Spindle vibration is evaluated based on the measured unbalance response vector in these experiments. The runout component is eliminated and the average values of unbalance response vector during 10 rotations are compared. Electromagnets are excited by input voltage, which is square wave (5 V), during 1/6 period of each spindle rotation.

As experiment I(a), unbalance mass 0.5 g (Unbalance: 26.3 g-mm) is added to the location 0 deg. of the compensated test spindle rotating at 4200 min^{-1}. In order to reduce the unbalance response, it is attempted that the counter weight in the location 180 deg. is generated by controlling magnetic fluid. Figure 9(a) shows the measured unbalance response vector. In this figure, it is confirmed that unbalance response increases when the unbalance mass is added to the compensated spindle. Then, the unbalance response vector becomes quite small and close to the origin O_r because of the effect of operating the proposed real-time balancing mechanism.

In addition, the location of unbalance mass 0.47 g (Unbalance: 24.7 g-mm) is changed to 56.3 deg. as experiment I(b). As shown in Figure 9(b), the unbalance response vector obtained in the added mass situation is rotated approximately 56.3 deg. in a counterclockwise direction compared to experiment I(a). By actively controlling the magnetic fluid inside the test spindle, the unbalance response is reduced again in this case. These results show the ability of the proposed real-time balancing mechanism, which can arbitrarily control the distribution of the circumferential mass of magnetic fluid.

The rotational speed of the spindle is changed to the maximum speed 6000 min^{-1}, and unbalance mass 0.41 g (Unbalance: 21.5 g-mm) is added to the location 0 deg. of the compensated test spindle as experiment II. Figure 10 shows the measured unbalance response vector in this experiment. It is recognized that the unbalance response with unbalance weight is larger than those obtained in experiment I. Though centrifugal force becomes larger than experiment I, the unbalance response is reduced greatly in case of generating the counter weight by the proposed real-time balancing mechanism.

According to these verification experiments, it is confirmed that unbalance response of the spindle can be improved significantly by the proposed real-time balancing mechanism using magnetic fluid. Moreover, the availability of the proposed real-time balancing mechanism is proved by the results obtained under the different rotational speeds and the various unbalance weights.

5 SUMMARY

In order to reduce unbalance response of the spindle for achieving high accuracy and efficiency machining, a smart spindle with a real-time balancing mechanism using magnetic fluid was proposed in this study. A test spindle with the proposed real-time balancing mechanism was designed and developed for experimental verification to evaluate the real-time balancing by controlling the magnetic fluid. Initial unbalance weight of the spindle was identified prior to verification experiments of the proposed real-time balancing. The unbalance response of the compensated spindle was considerably reduced compared to that of the initial spindle in the whole speed range. The characteristics of 3 kinds of magnetic fluids were investigated. It was recognized that magnetic fluid, which had the highest saturation magnetization, responds quickly to the changed magnetic field intensity. By using the developed test spindle with the proposed real-time balancing mechanism, experimental verifications were conducted. It was confirmed that unbalance response of the spindle can be improved significantly. The experimental results presented in this paper proof the feasibility of the proposed real-time balancing mechanism for machine tool spindles.

6 ACKNOWLEDGMENTS

This study was supported by Industrial Technology Research Grant Program in 2005 from New Energy and Industrial Technology Development Organization (NEDO) of Japan. The authors would like to express their sincere thanks to Dr. S.Kohda and Mr. M.Kumagai (OKK Co.) who contributed to develop the test spindle.

7 REFERENCES

[1] Altintas, Y., Cao, Y., 2005, Virtual Design and Optimization of Machine Tool Spindles, Annals of the CIRP, 54/1: 379-382.

[2] Brecher, C., Spachtholz, G., Paepenmüller, F., 2007, Developments for High Performance Machine Tool Spindles, Annals of the CIRP, 56/1: 395-399.

[3] Ehrich, F.-F., 2004, Handbook of Rotordynamics, Krieger Publishing Co.: 3.1-3.117.

[4] ISO 1940-1, 2003, Mechanical Vibration - Balance Quality Requirements for Rotors in a Constant (rigid) State - Part 1: Specification and Verification of Balance Tolerances.

[5] ISO 11342, 1998, Mechanical Vibration - Methods and Criteria for the Mechanical Balancing of Flexible Rotors.

[6] Slocum, H.-A., Marsh, R.-E., Smith, H.-D., 1994, A New Damper Design for Machine Tool Structures: the Replicated Internal Viscous Damper, Precision Engineering, 16/3: 174-183.

[7] Aoyama, T., Inasaki, I., 1997, Application of Electrorheological Fluid Dampers to Machine Tool Elements, Annals of the CIRP, 46/1: 309-312.

[8] Milecki, A., 2001, Investigation and control of magneto–rheological fluid dampers, International Journal of Machine Tools and Manufacture, 41/3: 379-391.

Prototyping of Autonomous CNC Machine Tool Based on Digital Copy Milling Concept

Keiichi Shirase[1], Takashi Shimada[1] and Keiichi Nakamoto[2]

[1] Department of Mechanical Engineering, Graduate School of Engineering, Kobe University, Japan

[2] Department of Mechanical Engineering, Graduate School of Engineering, Osaka University, Japan

Abstract

A prototype of CNC machine tool which realizes autonomous milling operation was developed. The machine tool developed is controlled by the digital copy milling concept. In the digital copy milling concept, a tracing probe and a master model in traditional copy milling are represented by 3D virtual models, and cutter locations are generated dynamically according to the motion of virtual tracing probe in real time. Therefore, not only feed speed, but also radial and axial depths of cut can be adapted for milling process control. Also, new tool paths can be added to avoid and recover from cutting troubles without any NC program. In the experimental tool crash condition, the cutting tool was successfully retracted to avoid tool breakage.

Keywords:

Milling, NC Control, Autonomous CNC Machine Tool, Digital Copy Milling

1 INTRODUCTION

Computer Numerical Control (CNC) machine tools contribute to achieve machining automation, besides to achieve high accuracy and high productivity of machining operations. However, conventional CNC machine tools require instructions written in an NC program to perform whole machining operations, and have no function to change cutting parameters for machining process control and cutting trouble avoidance.

Additionally, the machine tool and the cutting tools to be used have to be predetermined to generate instructions or NC commands properly. Naturally, the NC program prepared is special for the predetermined machine tool and cutting tools. It is the cause to the reduction of machining operations' flexibility.

In order to solve these problems, a new architecture to develop an autonomous CNC machine tool which does not require any NC program is proposed. Figure 1 shows the autonomous CNC machine tool proposed in comparison with conventional machine tool. A conventional CNC machine tool has no function to change instructions or NC commands dynamically during machining operation. Autonomous functions such as machining process control, cutting trouble avoidance and flexible machining operation are not required to operate a conventional CNC machine tool.

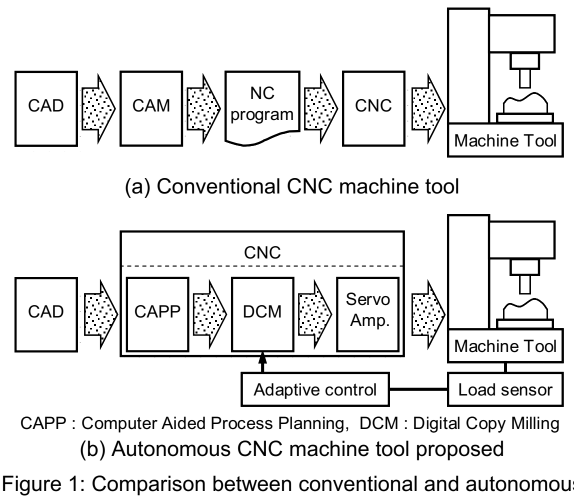


Figure 1: Comparison between conventional and autonomous machine tools.

The autonomous CNC machine tool proposed utilizes a system called Digital Copy Milling (DCM) to generate tool paths in real time. In DCM, a tracing probe and a master model in traditional copy milling are represented by 3D virtual models in a computer. A virtual tracing probe is simulated to follow a virtual master model, and cutter locations or tool paths are generated dynamically according to the motion of virtual tracing probe in real time. Therefore, NC program is not required. Additionally, not only feed speed, but also radial and axial depths of cut can be adapted for machining process control. Also, new tool paths can be added using the DCM to avoid and recover cutting troubles without any NC program.

In our previous study, our concept was verified using a commercial CNC machining centre [1]. However, the tool breakage avoidance function could not be verified because of the lack of response speed. In this study, a prototype of autonomous machine tool which has a PC based CNC controller was developed to verify our concept. Both the cutting load adaptation and the tool breakage avoidance were performed successfully in the feasibility milling test.

2 PREVIOUS RESEARCHS ON CNC CONTROLLOR

Altintas et. al. developed a hierarchical Open-Architecture CNC system to achieve adaptive cutting force control and tool breakage detection [2]. This system was improved as the open and reconfigurable modular tool kit named ORTS for the design of CNC systems [3]. Tool path generation using spline interpolation function and adaptive milling force control are demonstrated as sample applications. Mori et. al. developed an open servo control system to support custom control functions [4]. The software model reference adaptive control is demonstrated to show the effectiveness of the open servo control system. In these researches, CNC systems were developed to achieve easy implementation of custom control functions. Open architecture is important to accomplish flexible CNC controllers which satisfy the requirements of the end users. However, these researches focus on functionality and flexibility of CNC controllers. Therefore, the concept of DCM proposed in this study to achieve an autonomous CNC machine tool is unique as compared with the previous researches mentioned.

3 AUTONOMOUS MILLING PROCESS CONTROL

3.1 Digital Copy Milling

The Digital Copy Milling system (DCM) generates cutter locations or tool paths in real time. Therefore, it is possible to change cutting parameters during milling operation.

In traditional copy milling, the feed velocity vector of tool motion is calculated from the maximum displacement ε which is detected by a physical tracing probe. On the other hand, the feed velocity vector of tool motion is calculated from the maximum collision ε between the virtual tracing probe and the virtual master model in the DCM, as shown in Figure.2.

The displacement ε is kept to be the reference displacement ε_0, and the tracing probe is controlled to follow the master model. The relation among the vectors of the tangential feed velocity V_T, the normal feed velocity V_N and the resultant feed velocity V are shown in Figure 3. The tangential feed velocity V_T and the normal feed velocity V_N to the surface of the master model are calculated by following Equations (1), (2) and (3).

$$\left|\vec{\varepsilon}_d\right| = \left|\vec{\varepsilon}\right| - \varepsilon_0 \quad (1)$$

$$\left|\vec{V}_N\right| = G_N \cdot \left|\vec{\varepsilon}_d\right| \quad (2)$$

$$\left|\vec{V}_T\right| = V_C - G_T \cdot \left|\vec{\varepsilon}_d\right| \quad (3)$$

In these equations, ε_d is the difference between the displacement ε and the reference displacement ε_0, G_N is called V_N gain which defines the relation between ε_d and V_N, G_T is called V_T gain which defines the relation between ε_d and V_T. V_C is the commanded feed velocity.

In the DCM, the maximum collision ε between geometric models of the virtual tracing probe and the virtual master

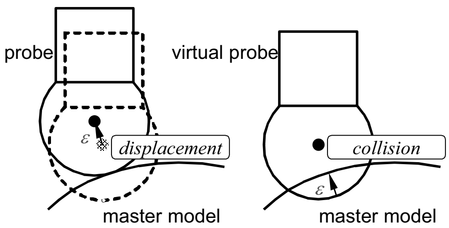


Figure 2: Displacement for traditional copy milling and collision for digital copy milling.

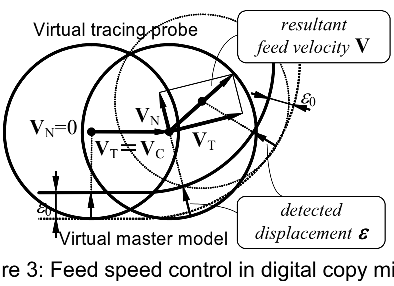
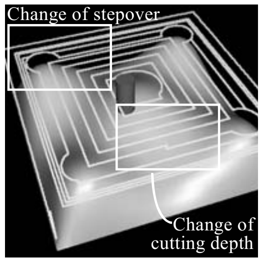


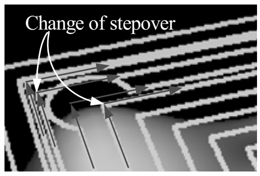
Figure 3: Feed speed control in digital copy milling.

model is used to calculate the resultant feed velocity V to control the tool motion. The cutter location is calculated from the resultant feed velocity V at each control interval t.

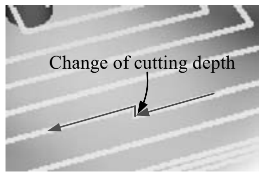
In the DCM, unilateral, bilateral (zigzag) and contour scanning tool path patterns are available. Figure 4 shows an example of contour scanning tool path pattern. Even radial and axial depths of cut can be changed during milling operation. It is a distinctive feature of the DCM.



(a) Example of tool path generation



(b) Change of stepover or radial depth of cut



(c) Change of cutting depth or axial depth of cut

Figure 4: Tool path generation in digital copy milling.

3.2 Machining strategy of milling process control

A machining strategy to adapt feed speed, radial and axial depths of cut is integrated with the DCM for the autonomous milling process control. The flow chart of the proposed machining strategy is shown in Figure.5.

The machining strategy consists of "Cutting trouble avoidance" and "Cutting load adaptation". Cutting load is monitored by a force sensor. "Critical Load" is settled for the cutting trouble avoidance and "Desired Load" is settled for the cutting load adaptation. Once the cutting load is over the "Critical Load", the feed speed will be reduced and the cutting tool will be retracted to avoid the cutting trouble.

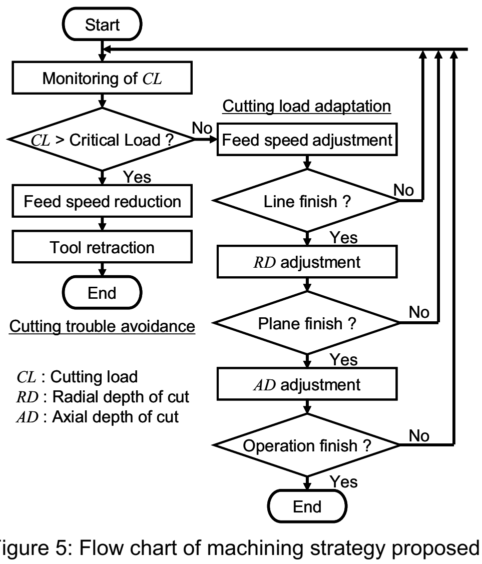


Figure 5: Flow chart of machining strategy proposed.

When cutting load is under the “Critical Load”, the cutting parameters are adjusted to keep the cutting load between the minimum and the maximum “Desired Load”. Milling operation is performed by successive planes which have a displacement, equivalent to axial depth of cut, from the previous plane machined. One plane is machined by successive lines which have a displacement, equivalent to radial depth of cut, from the previous line machined. The feed speed is adjusted during the milling of one line, the radial depth of cut RD is adjusted for the milling of the next line, and the axial depth of cut AD is adjusted for the milling of the next plane.

For the adjustment of cutting parameters, reference values such as minimum, maximum and reference values are defined for feed speed, radial and axial depths of cut, respectively. These values depend on machining know-how. Reference values are set as default values of the cutting parameters to start the milling operation. The cutting parameters are adjusted between the minimum and maximum values of them.

To adjust the radial depth of cut RD for the next line milling, the maximum and the average cutting loads are memorized as L_m and La during the current line milling. The radial depth of cut for the next line milling RD_n is determined by the ratio of the “Desired Load” L_d and the maximum spindle load L_m in the current line milling, when feed speed is maximum. Also, the RD_n is determined by the ratio of the “Desired Load” L_d and the average spindle load L_a in the current line milling, when feed speed is minimum.

Similarly, to adjust the axial depth of cut AD for the next plane milling, the maximum and the average cutting loads are memorized as L_M and L_A during the current plane milling. The axial depth of cut for the next plane milling AD_n is determined by the ratio of the “Desired Load” L_d and the maximum cutting load L_M in the current plane milling, when radial depth of cut is maximum. Also, the AD_n is determined by the ratio of the “Desired Load” L_d and the average cutting load L_A in the current plane milling, when radial depth of cut is minimum.

Figure 6: Prototype of autonomous machine tool developed.

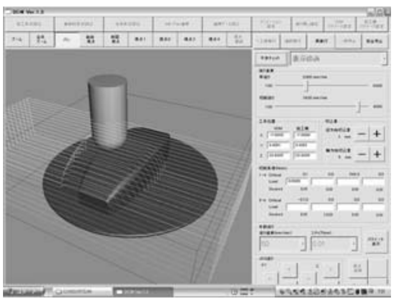


Figure 7: Graphic user interface of DCM.

4 PROTOTYPE OF AUTONOMOUS MACHINE TOOL

In this study, a prototype of autonomous machine tool which has a PC based CNC controller was developed to verify our concept. Figure 6 shows the prototype of autonomous machine tool which is a desktop vertical 3-axis machining centre, It has two CPUs, Intel Pentium IV 2.4GHz CPU is working for the DCM with graphical user interface, Renesas SH4 240MHz CPU is working as a conventional CNC controller. Additionally, a force sensor to detect cutting load in X and Y axis is equipped in this machine tool.

Figure 7 shows the graphical user interface of the DCM. In this version of the DCM, a Computer Aided Process Planning (CAPP) system have not been implemented yet. Therefore, milling sequence, cutting tool and cutting conditions have to be determined manually. Tool paths are generated autonomously by the tool control principle of traditional copy milling.

5 VERIFICATION OF AUTOMONOUS MILLING PROCESS CONTROL

5.1 Cutting load adaptation

In order to verify the effectiveness of the cutting load adaptation, the experimental milling was conducted. Cutting conditions, the range of “Desired Load”, feed speed and radial depth of cut for adaptive control are summarized in Table 1. In the Cases A and B, the reference radial depths of cut RD is varied to change the cutting load at the immersion of milling.

Table 1: Experimental conditions of cutting load adaptation.

<table>
<tr><td colspan="2"></td><td>Case A</td><td>Case B</td></tr>
<tr><td rowspan="2">Tool</td><td>Shape</td><td colspan="2">Square end mill</td></tr>
<tr><td>Diameter</td><td colspan="2">6 mm</td></tr>
<tr><td colspan="2">Workpiece</td><td colspan="2">Cast iron (JIS FC250)</td></tr>
<tr><td colspan="2">Tool path pattern</td><td colspan="2">Bilateral (zigzag) scanning</td></tr>
<tr><td rowspan="3">Feed speed</td><td>Maximum</td><td colspan="2">1300 mm/min</td></tr>
<tr><td>Reference</td><td colspan="2">1000 mm/min</td></tr>
<tr><td>Minimum</td><td colspan="2">700 mm/min</td></tr>
<tr><td rowspan="3">Radial depth of cut RD</td><td>Maximum</td><td colspan="2">5 mm</td></tr>
<tr><td>Reference</td><td>3 mm</td><td>0.8 mm</td></tr>
<tr><td>Minimum</td><td colspan="2">0.5 mm</td></tr>
<tr><td colspan="2">Axial depth of Cut AD</td><td colspan="2">3 mm</td></tr>
<tr><td colspan="2">Spindle speed</td><td colspan="2">6000 min^{-1}</td></tr>
<tr><td rowspan="3">Desired load</td><td>Upper</td><td colspan="2">118 N (98N + 20N)</td></tr>
<tr><td>Reference</td><td colspan="2">98 N</td></tr>
<tr><td>Lower</td><td colspan="2">78 N (98N - 20N)</td></tr>
</table>

Figure 8 shows the experimental results of the Case A. Cutting load shown in Figure 8 (a) is decreased to approach the “Desired Load” successfully, according to the adaptive control of feed speed and radial depth of cut shown in Figure 8 (b). At the first line milling, feed speed is decreased, then radial depth of cut is decreased before the second line milling to adjust cutting load. Figure 9 shows the experimental results of the Case B. Cutting load shown in Figure 9 (a) is increased to approach the “Desired Load” successfully, too. At the first line milling, feed speed is increased, then radial depth of cut is increased before the second line milling to adjust cutting load as shown in Figure 9 (b).

It is confirmed that not only feed speed, but also radial depth of cut can be adapted for milling process control. Also, new tool paths are generated to change radial depth of cut.

5.2 Tool breakage avoidance

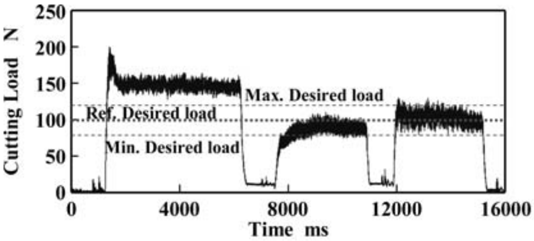
In order to verify the effectiveness of the tool breakage avoidance, the experimental milling was conducted. Cutting conditions are summarized in Table 2.

The tool breakage avoidance was performed successfully as shown in Figure 10. Feed speed is reduced quickly when cutting load is over the “Critical Load”, and the cutting tool was retracted before the tool break. But a time delay about 40msec still remained.

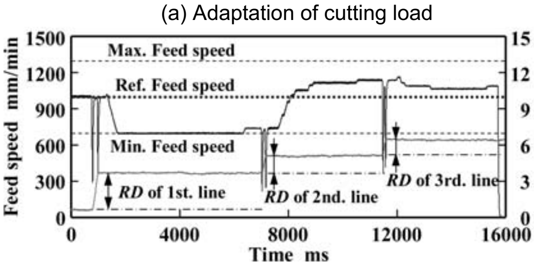
6 CONCLUSIONS

A prototype of CNC machine tool which realizes autonomous milling operation was developed. The machine tool developed is controlled by the digital copy milling concept. The results are summarized as follows.

1. The digital copy milling system based on the principle of traditional copy milling has been developed. The real-time tool path generation and the cutting parameter adaptation during milling operation are available.

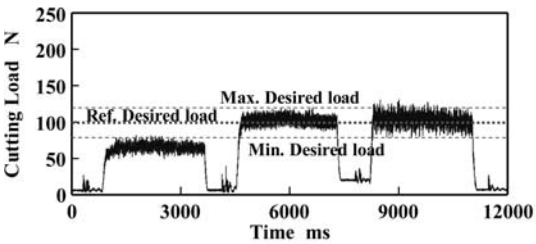


(a) Adaptation of cutting load

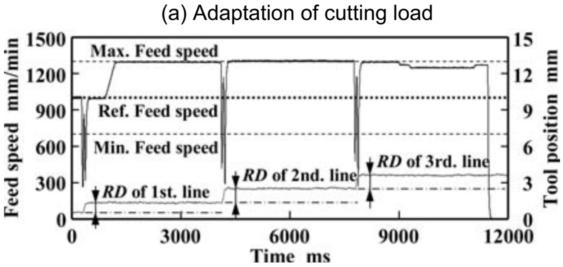


(b) Adaptive control of feed speed and radial depth of cut

Figure 8: Adaptation to decrease cutting load.



(a) Adaptation of cutting load



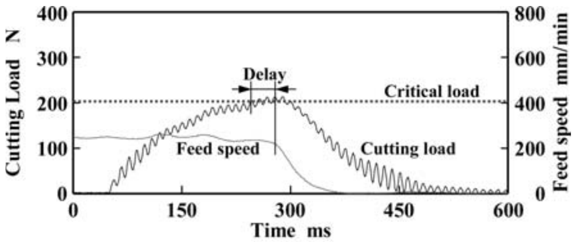
(b) Adaptive control of feed speed and radial depth of cut

Figure 9: Adaptation to increase cutting load.

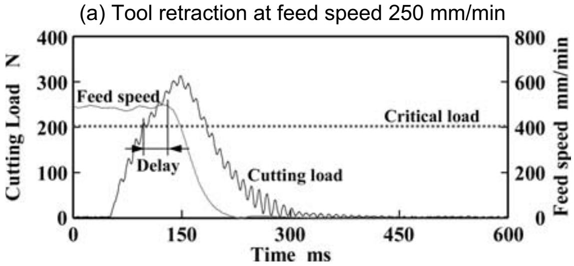
2. The machining strategy of milling process control is integrated with the digital copy milling. The cutting load adaptation was performed successfully in the experimental cutting.
3. The tool breakage avoidance was performed successfully, in the experimental tool crash condition. The cutting tool was retracted before the tool break.

Table 2: Experimental conditions of tool breakage avoidance.

		Case C	Case D
Tool	Shape	Square end mill	
	Diameter	4 mm	
Workpiece		Cast iron (JIS FC250)	
Feed speed		250 mm/min	500 mm/min
Radial depth of Cut *RD*		Full immersion	
Axial depth of Cut *AD*		10 mm	
Spindle speed		6000 min^{-1}	
Critical load		200 N	



(a) Tool retraction at feed speed 250 mm/min



(b) Tool retraction at feed speed 500 mm/min

Figure 10: Experimental results of tool breakage avoidance.

7 ACKNOWLEDGMENTS

This study was supported by a grant of Ministy of Economy, Trade and Industry (METI) of Japan. The authors would like to express their thanks to Mr. Mishima and Mr. Furumura who contributed to develop the prototype machine.

8 REFERENCES

[1] Shirase, K., Nakamoto, K., Arai, E., Moriwaki, T., 2005, Digital copy milling - autonomous milling process control without an NC program, Robotics and Computer-Integrated Manufacturing, 21/4-5, 312-317.

[2] Altintas, Y., Munasinghe, W.K., 1994, A Hierarchical Open-Architecture CNC System for Machine Tools, Annals of the CIRP, 43/1, 349-354.

[3] Altintas, Y., Erol, N.A., 1998, Open Architecture Modular Tool Kit for Motion and Machining Process Control, Annals of the CIRP, 47/1, 295-300.

[4] Mori, M., Yamazaki, K., Fujishima, M., Liu, J., Furukawa, N., 2001, A Study on Development of an Open Servo System for Intelligent Control of a CNC Machine Tool, Annals of the CIRP, 50/1, 247-250.

Intricate Shaped Manufacturing by 6-Axis Non-rotational Cutting with the Application of Ultrasonic Vibrations

Feliciano H. Japitana[1], Koichi Morishige[2] and Yoshimi Takeuchi[3]

[1] Metals Industry Research and Development Center – Department of Science and Technology, Manila, Philippines

[2] Morishige Laboratory, Department of Mechanical and Intelligent System, The University of Electro-Communications, Tokyo, Japan

[3] Department of Computer-Controlled Mechanical Systems, Osaka University, Osaka, Japan

Abstract

Complex mold or intricate shaped product can be manufactured by rotational cutting tool with an increased in cutting speed, a process now known as high speed cutting; however, shaving remnants are formed according to the tool shape at the adjoining surfaces or end corner of the product. This limitation is inevitable, especially when employing rotational tool in the process. If this limitation can be removed by cutting, the flexibility of the product form, which can carried out in manufacturing can be extended greatly.

This paper presents a new manufacturing method to produce an intricate shaped product utilizing the 6-axis control machining using a non-rotational cutting tool with the application of ultrasonic vibrations. The result of experiment shows that intricate shape product has been manufacture using the new manufacturing method.

Keywords:

Non-rotational Tool, 6-Axis Control Machine Tool, Ultrasonic Vibrations Cutting, Intricate Shaped

1 INTRODUCTION

With the increasing demand for new product in the market nowadays, many sophisticated designs or intricate shaped parts and components are adopted to meet the latest requirement particularly in the die and mold industry. Thus, molds must be manufactured to cope with production requirements of the market and must consider the following: low cost; fast delivery and flexible in manufacturing process due to intricacy of shape. Under the circumstances, the concept of simplified molds was formed and one of these is by using aluminum materials. Aluminum alloy can be utilized in producing molds with low molding pressure, such as by blow molding, vacuum forming, rubbing forming, etc. This material is also suitable to utilize in prototyping a product. Intricate shaped product or complex mold can be manufactured by rotational cutting tool to increase the cutting speed, a process now known as high speed cutting; however, shaving remnants are formed according to the tool shape at the adjoining surfaces or end corner of the product as shown in Fig. 1.1(a). This limitation is inevitable, especially when employing rotational tool in the process. The efficiency and accuracy of machining a product is a critical factor in many manufacturing process. The most efficient machining would mean the cheapest machining of parts with required dimensions, tolerances and quality in the shortest possible time. The completion of machining in just one setting would lead to the production of high accuracy, quality and efficiency. Since the result of processing with rotational tool is symmetrical with the rotation, processing of a intricate part shape or asymmetrical form is difficult to achieve. For example, if the intricate part or form has a character line, sharp corners or groove on side surfaces, processing of this part or form using a rotational tool will produce an arc-like radius remains. This is the limitation acquired from this processing with rotational tool. If this limitation can be removed by cutting, the flexibility of the product form, which can carried out in the creation can be extended greatly. Some of the important elements of inticate shapes are, character line, sharp corner or groove on adjoining or side surfaces and these are usually produced by electric discharge machining (EDM) as shown in Fig. 1.1(b). However, such process requires two or more setup which will affect the efficiency and accuracy of the product due to the lenghtening of machining time and giving limitations on the type of materials to be machined since only electrically conductive materials are suitable for such machining process. The 6-axis control cutting using a non- rotational tool has been applied to solve the problem, as shown in Fig. 1.1(c), but still, another problem arises - i.e. low surface quality of the product due to the low cutting speed in nonrotational cutting as shown in Fig. 1.1(d).

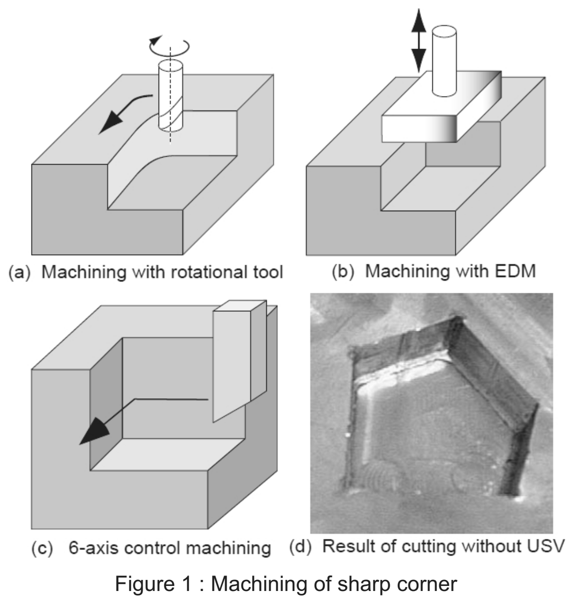


Figure 1 : Machining of sharp corner

High cutting speed is required in order to produce a good surface roughness on the product especially in machining of soft or malleable materials like aluminum.

The needs of new machining technology to deal with the problem in machining of intricate shapes are necessary. The multi-axis machining, which is the 6-axis control machining together with the ultrasonic vibration will be used to solve the problem. The 6-axis control machine tool using the non-rotational cutting tool will be used to deal with the problem caused by using the rotational cutting tool. The ultrasonic vibration will be applied to solve the problem in non-rotational cutting tool which is the low cutting speed. The low cutting speed in the non-rotational cutting will be offset by applying the ultrasonic vibration duringcutting operation. Low cutting speed greatly affects the surface roughness of the product especially if the material is malleable like aluminum alloy. High cutting speed is needed to produce a good surface roughness of the product. The developed CAM software will also be used in machining a complex or complicate shape since it is indispensable in the process. Multi-axis has been used to lower cost of processing of a given part or assembly to a required specification in metal manufacturing. It helps to achieve one's goal, to increase productivity. It also minimize the number of machining needed to complete the processing of parts, as well as the needs for a more uniform machining accuracy and improved product quality.

2 EXPERIMENTAL MECHANISM

2.1 Ultrasonic vibration cutting

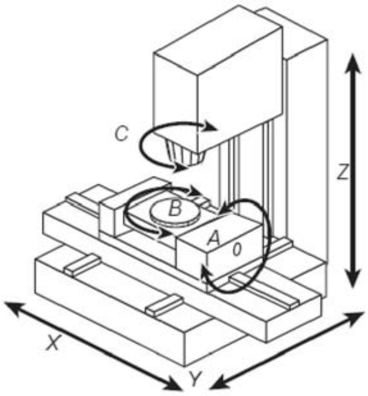
The development of the ultrasonic vibration cutting process was instigated largely by the more extensive use in recent years of hard, brittle materials and the necessity of finding means to machine them effectively and efficiently. Among other difficult machining problems it has solved, ultrasonic vibration cutting is being used successfully to machine carbides, stainless steels, and glass [19]. The process is best suited in machining hard and brittle materials which may be conductors or insulators.

2.2 6-axis control machine tool

Flexible multi-functional but cost effective production equipment is becoming more and more important to cover the market turbulence. In order to realize a fast and flexible single-hit production process, multi-axis machine tools have already achieved a considerable degree of acceptance. Most of them are based on multi-axis structures, where individual rotary and linear axes are connected in a serial way. Their advantage is their dexterity, thus allowing for a large number of applications [2-4].

One of the typical example of multi-axis machine tool is the 6-axis control machine tool. Shown in Fig. 2(a) is the actual photo of the 6-axis control machine tool used in the study and also shown Fig. 2(b) is the schematic illustration of the said machine tool. The 6-axis control machine tool provides multiple-axis machining capabilities beyond the standard five-axis tool path movement. In order to realize the processing of the required intricate shape with ease, 6-axis control machine tool is required, since its flexibility is needed for expressing all the postures in the space for the tool. The 6-axis control machine tool is constructed by adding the rotational function *Cs* on the main spindle of the 5-axis control machine tool which has 2 rotational axes, namely; A, which is the rotary tilting of the table and B, which is the rotary index table.

(a) Actual 6-axis control machine tool (Makino GN-107A)



(b) Schematic view of 6-axis control machine tool

Figure 2 : 6-axis control machine tool

2.3 Non-rotational tool

Cutting tools maybe broadly classified as single point tools, which have one active cutting edge, and multipoint tools, which have multiple active cutting edges. Single point cutting tool can be defined as the tools which have a cutting edge at one end and it has always been widely used for metal removal. Single point tools are commonly used for turning and boring operations, while multipoint tools are used for drilling and milling operations. In the study, non-rotational tool is used to solve the limitation in using a rotational tool. One of its typical applications is on non- rotational cutting like planing, in which this tool is also applicable in 6-axis control ultrasonic vibration cutting. This type of tool is usually used in turning operation but we tried to use it in 6-axis control cutting. This tool is quite unique as compared with other tools because of its size and shape.

2.4 CAD/CAM for 6-axis control cutting

The CAM software for 6-axis control ultrasonic vibration cutting is indispensable [39] in order to fully machine the required shape of the product with ease. One of its key applications is the extraction of the data for use in the analysis and in manufacturing. And since there is no commercially available 6-axis CAD/CAM software yet in the market, the software for cutting of complex shape product must be developed by ourselves. The main method for developing such software are as follows; first the 3-dimensional model is drawn using a Computer Aided Design (CAD) program namely, DESIGNBASE. This model is used as a basis for generating a collision-free (CL) data on the basis of tool information, tool orientation and the target shape of the product using the main processor.

In either of 3 or 5-axis milling, the cutter location problem is essentially one on offsetting the cutter contact locus on the surface such that the cutter center follows the desired path. As an example, if the cutting tool used is a rotational tool like ball end mill, the center of the spherical end is offset by the radius of ball end mil in a direction equal to the surface normal at the contact point as illustrated in Fig. 3(a). In the case of 6-axis control cutting using the non-rotational tool, the cutter location is based on the cutting tip as shown in Fig. 3(b).

The CL data generated is converted into 6-axis numerical control (NC) data using the post processor on the bases of machine information, setting information and cutting condition. The NC data generated will be employed in controlling machining operation.

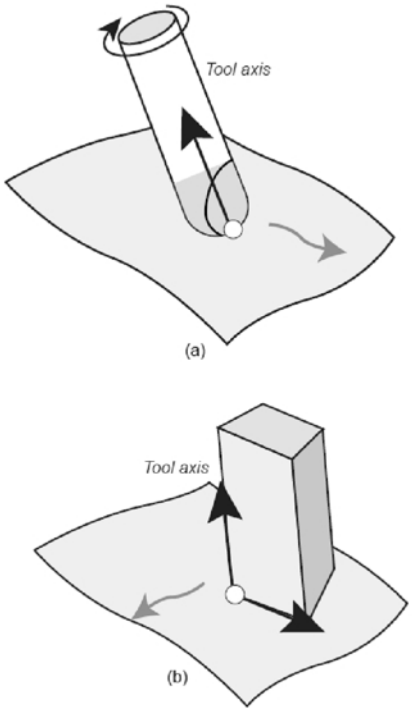


Figure 3 : Difference of cutter location: (a) CL of 3 to 5-axis control rotational tool, (b) CL of 6-axis control non-rotational tool.

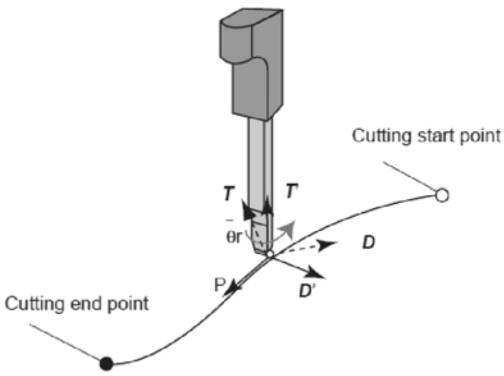


Figure 4 : Define tool path for non-rotaional cutting tool.

Figure 4 shows a tool path transversing a defined edge wherein the tool tip contact points for a given position and orientation. The tool position and orientation is based on the CL data generated by the main processor.of the cutting points and the angle of expression of the tool posture. The NC data converted from the CL data by the post processor, in turn, are used in the machining operation of intricate shape product. The NC data controls the movement of the tool and all of the axes produces such NC data.

3 GENERATIONS OF CUTTER LOCATION (CL) DATA

3.1 Main processor for generating CL data

The main processor was developed utilizing C language with the basic library files provided by DESIGNBASE (Ricoh Co., Ltd) software. Based on the geometric information of the OHG obtained from the 3D-CAD data of the workpiece, the main processor generates the position of the cutting point (P), the normal vector (***N***) and the feed vector (***F***). The PNF data are generated on the basis of the cutting tool shape, the cutting direction and the depth of cut. The cutting direction is decided based on the structure of the tool and the location of the intricate shape like overhanging groove (OHG). If the opening of overhanging groove is located on the right side of the surface and close to the end, as shown in Fig. 5(a), the cutting direction to be used for generating the PNF data will depend on the left- handed non-rotational tool, and if the opening of the OHG is on the opposite side as shown in Fig. 5(b), the same procedure will be used; however the bases of generating the PNF data for cutting direction will be right-handed non-rotational tools.

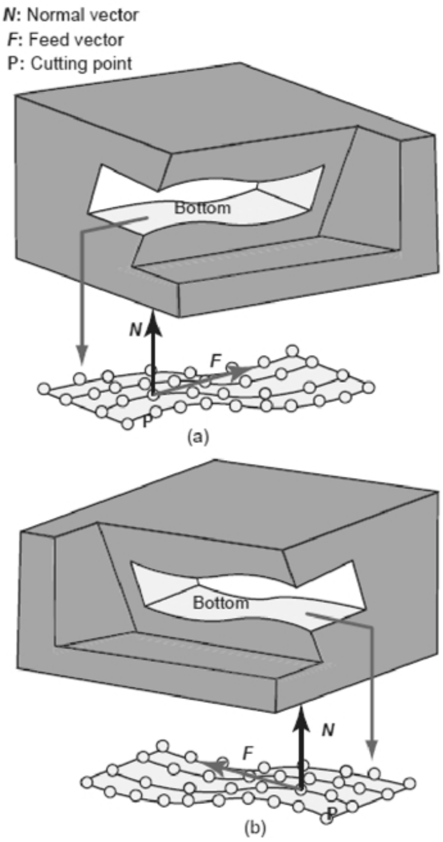


Figure 5 : Generation of PNF data: (a) PNF data for left side surface (b) PNF data for right side surface.

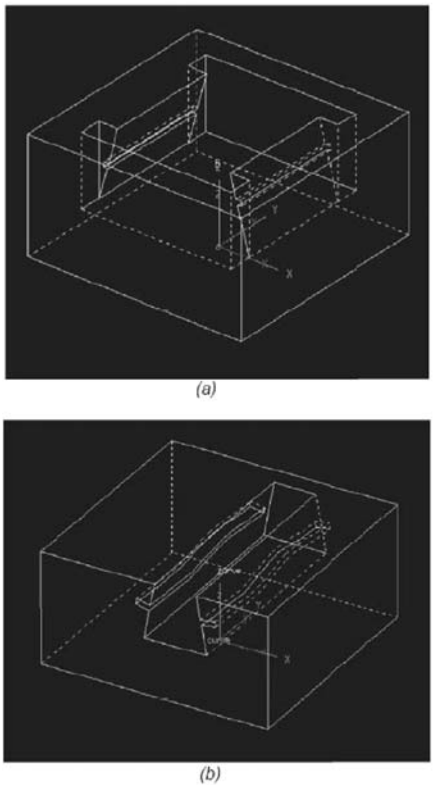


Figure 6 : Target shape: (a) groove with overhanging surface (OHG), plain type, (b) groove with overhanging surface (OHG), curve type.

The number of divisions in the direction of the groove determines the depth of cut. As the number of divisions is being input, the system automatically computes the depth of cut. In a curved OHG, the number of divisions on the lenght of surface greatly affects the curvature of the OHG. When the number of divisions increases to create the curved OHG accurately, the cutting points generated by the system also increases. On the other hand, in linear OHG, the number of divisions is not so important since it deals only with a straight line.

After the PNF data for OHG is generated, they will be converted into cutter location (CL) data using the same main processor. The main processor generates the collision-free CL data, based on the structure of the OHG, the cutting tool information and the vibration conditions. The CL data is obtained by selecting the bottom surface of the groove. The depth of cut and the number of divisions in the PNF data are then inputted again. Collision check is carried out in the process, but collision avoidance is not performed in this operation. The entry and exit points of the cutting tool are very important, especially in dealing with an OHG that is close to the end surface. The entry and exit points are added to the CL data before converting to NC data.

3.2 Postprocessor for OHG

The CL data generated by the main processor, the P, *T* and *D* vectors, are changed into specific NC data by the postprocessor. The NC data are identical to the expression with regards to tool posture. They consist of the coordinate valuesof the cutting points and the angle of expression of the tool posture. The NC data converted from the CL data by the post processor, in turn, are used in the machining operation of OHG. The NC data controls the movement of the tool and all of the axes produces such NC data.

4 MACHINING AND RESULTS OF EXPERIMENT

The cutting experiment has been done in order to verify the validity of the developed software as well as the new machining method for overhanging grooves. The overall size of the workpiece is 80 x 80 x 40 mm and the material is an aluminum alloy (A5052). Two type of OHG have been cut in the study, namely the straight type groove within the pocket and the curved type groove with a closed-end surface, as shown in Fig. 6(a) and (b), respectively. The total depth of cut and the minimum width of cut were limited to only 2.0 and 2.0 mm, respectively, due to the size and construction of the cutting tool. Rough cutting has been done by 3-axis control cutting on the plain surfaces using a rotational tool ball end mill with a radius of 3mm while the overhanging portion of the surfaces was cut using the 5- axis control cutting. After rough cutting was done, the OHG was machined using 6-axis control machining with the application of USV. The generated 6-axis control NC data are used to cut the OHG. During cutting, USV is also employed in the cutting tool to ensure the smoothness of the machining operation. In cutting a straight line type OHG, the A-axis tilts to a certain degree so as to follow the required inclination angle of the product. In the case of a curved type OHG, all of the axes (X, Y, Z, A, B and C) are driven synchronously to obtain the required shape. In this situation, the function Cs plays a vital role, since it makes the vibration direction parallel to the cutting direction so as to continue with an effective and efficient cutting.

5 CONCLUSIONS

Machining of complex shape grooves, which is impossible to machine in conventional manner was successfully done even if it is situated adjacent to the obstruction without the collision of the tool and the workpiece in one setup. Potential cost saving is observed since the use of expensive fixture is eliminated, and only one machine tool with 6-axis control function is also used all throughout the process. It also gives a new idea that the multi-axis machining broadens its manufacturing capabilities.

6 ACKOWLEDGEMENTS

The authors would like to express their sincere appreciation to the assistance given by the Ministry of Education, Culture, Sports, Science and Technology (Monbukagakusho) in providing funds on some part of this study.

7 REFERENCES

[1] T., Moriwaki, E., Shamoto and K., Inoue: Ultraprecision ductile cutting of glass by applying ultrasonic vibration, Annals of the CIRP, 1992, Vol.41, pp.141-144.

[2] Y. Koren: Control of machine tools, Journal of Manufacturing Science and Engineering, 1997, Vol.119, pp.749-755.

[3] K. Srinivasan and T.C. Tsao: Machine tool feed drives and their control-A survey of the state of the art, Journal of Manufacturing Science and Engineering, 1997, Vol.119, pp.743- 748.

[4] Y. Takeuchi, H. Shimizu, T. Idemura and T. Watanabe, 5-axis control machining based on solid model, Journal of Advanced Automation Technology, Vol.2, No.6, 1990, pp.23-28.

Expanding the possibilities of position error compensation in CAM for PKM milling machines

Jan Los[1], Tetsuo Tomiyama[2], Tetsuro Shibukawa[3] and Yoshimi Takeuchi[4].

[1] Dep. of Precision Mechanical Engineering/PT, Delft University of Technology, The Netherlands

[2] Department of Bio-Mechanical Engineering, Delft University of Technology, The Netherlands

[3] Technical R&D Center, JTEKT Corporation, Japan

[4] Department of Mechanical Engineering, School of Engineering, University of Osaka, Japan

Abstract

PKM milling machines have promising advantages compared to SKM milling machines, such as a higher stiffness-to-mass ratio, enabling a better dynamic behaviour. Errors in the PKM driving chains cause position errors of the end effecter. Its parallel construction and non-uniform position error patterns make finding the error sources a challenging task. This paper presents the application and expansion of a position error reduction method. Measured and predicted error patterns, the latter generated by a kinematic model, are compared and the most likely error source is found. With this information, the NC code is systematically modified and the PKM machine can be adjusted.

Keywords:

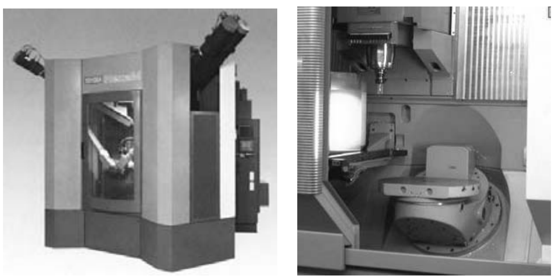
Parallel Kinematic Mechanisms, Multi-axis Milling Machines, Position Accuracy, Error compensation, CAM

1 INTRODUCTION

1.1 Background

Within the field of milling there is a continuous search for higher accuracy and productivity. In practice, a compromise has to be made between these two goals.

In general, milling machines need to be very rigid to archive a high accuracy. Rigid machine components tend to have a high mass. This is especially true for the Serial Kinematic Mechanism (SKM) Milling Machines. The construction of a SKM is sometimes referred to as 'stacked': each consecutive axis is stacked on the next. Because the axes bear eachother, they must be very ridged and strong.



(a) (b)

Figure 1: Hexam, prototype 2 from Toyoda Machine Works, Japan; A SKM milling machine

On the other hand, high speeds and accelerations of the milling tool is desired for a competitive productivity. Components with a high inertia limit the dynamic behaviour of the milling machine. There will always be a trade-off between accuracy and productivity.

Parallel Kinematic Mechanism (PKM) Machines have the potential to make this trade-off much less limiting. Due to a fundamentally different construction, PKM milling machines can sport a high stiffness to mass ratio, enabling a highly dynamic behaviour. In the HexaM milling machine (Figure 1a) for example, all actuators are connected to the end-effecter with a separate drive-chain (Figure 2). In other words, all drive-chains are contributing to the overall stiffness.

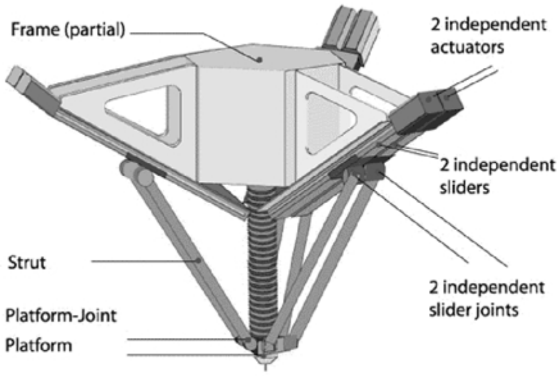


Figure 2: Lay out of the driving chains of the HexaM

1.2 Problem description

This paper describes a continuation of the PKM research preformed by S. Pastoor [1]. In his research report '*Improvement of HexaM, a Parallel kinematic machine tool, using error compensation in CAM*' [1], Pastoor describes how to use a kinematic model of the milling machine to find the most likely causes for the position error of the end effecter, and achieve higher tool position accuracies. The main goal is to gain a better understanding of which error sources are present in a PKM milling machine and what effects these error sources have on tool position accuracy. Applying this knowledge will lead to better milling accuracy and productivity. In short, this research tries is to find a universal method for improving position accuracy and gaining insight in the error sources of a PKM milling machine.

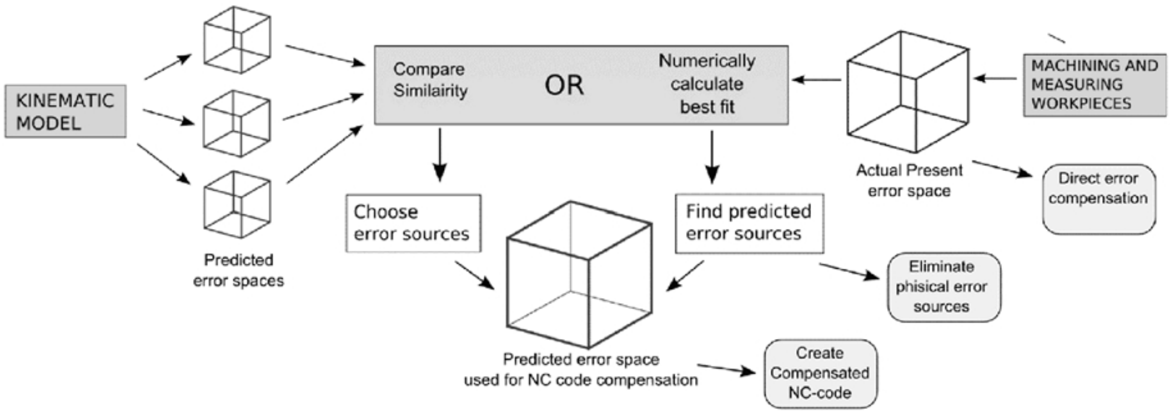


Figure 3: Scheme of the research approach.

2 RESEARCH APPROACH

In Figure 3 an overview of the research approach is given. In general, the research approach is similar to Pastoor's, but the focus is different. This research focuses on application, validation and expansion of the proposed research methods by Pastoor.

The first phase of this research consists of two separate parts. A kinematic model is made of the milling machine, predicting which 3D error pattern originates from a fixed set of error sources, and 3D measurements are carried out on the milling machine to investigate the actual present tool position error pattern. Here, the kinematic model calculates a characteristic error field belonging to error sources. In general, the position error is not caused by one error source only.

The single-source error patterns now are compared with the measured error pattern. When the two error fields show a similar pattern, the corresponding error source is likely to be a pysical error source. The combination of error sources which causes an error pattern most similar to the measured error pattern is determined. This is done by manualy compairing the characteristic patterns, or numerically as described in the recommendations.

The newly derived multi-source predicted error field acts as a blueprint for the compensation of the NC code through CAM. All nominal coordinates of the work piece NC-code are changed with the interpolated value of this error field (indirect compensation). This is the main research approach and the proposed method in this paper. This approach is universal and can be applied to any PKM milling machine There are more possibilities for position error compensation, some of them used in this research due to practical limitations.

2.1 Direct Compensation

The measured error space can be used to directly compensate the NC code (direct compensation). This is not a universal method, since the errors found are only relevant to the one machine under consideration, and no insight is gained in the source of the errors.

2.2 Eliminate error sources in the Machine

The set of error sources found by the numerical comparison method can be eliminated in the milling machine itself. This may be effective, bet very costly.

3 KINEMATIC MODEL OF THE MILLING MACHINE

This kinematic model describes the position of the 6 sliders and the tool tip, thus predicting the tool position error. The kinematic model used in this research is described by Pastor. The model is build according the geometrical information provided by the milling machine manufacturer J-Tekt [2]. The model uses forward and inverse kinematics. Inverse kinematics are used to calculate the position of the actuators when the tool position and orientation are given. Forward kinematics does exactly the opposite: calculating the position of the tool when the position of the actuators is given.

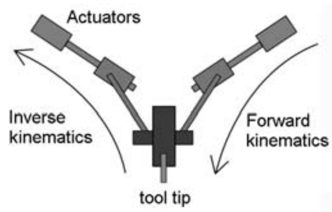


Figure 4: Direction of calculation in the Kinematic model.

3.1 Input of the Kinematic Model: Error sources

The Kinematic model has several error sources as an input:

- Geometrical errors
  - Calibration errors
  - Manufacturing and Assembly errors
- Errors due to gravity
  - Spindle bending
  - Ball-screw compliance
  - Strut elongation

3.2 Non-linearity and iteration

The forward kinematic calculation is more complex than the inverse calculation: The equations are not linear and there is no unique solution. Therefore the forward calculation uses a Newton-Rapson iteration described by

$$X_k = X_{k-1} - J^{-1}\left(U_k - U_{k-1}\right) \quad (1)$$

Where X_k is the calculated tool position, X_{k-1} the previous tool position, U_k the position of the sliders and U_{k-1} the previous slider positions. The Jacobian matrix J describes the relation between slider and tool velocities.

A example of a cross-section of one of the predicted error fields is shown in Figure 5.

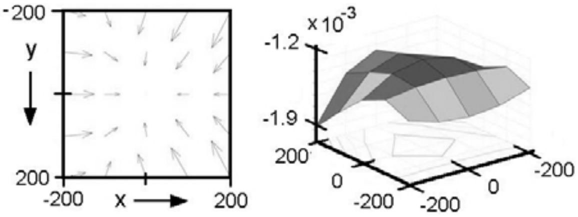


Figure 5: Predicted error pattern with axial strut deformation due to gravity.

4 EXPERIMENTS

A convenient method for investigating the position error of a milling machine is a direct position measurement with a Double Ball Bar, as described in [3] and [4]. Such a measurement device was not available, so this method was not used. Instead an indirect measurement method is used: work pieces are milled, and the machined geometry is measured with a Coordinate Measuring Machine (CMM). This method inherently reduces the achievable accuracy of the measurement.

4.1 Indirect Measuring of the Position Errors

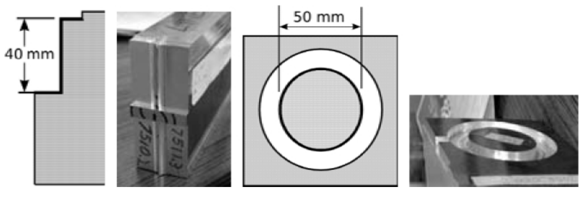
Two standard work pieces, for the vertical (Z axis) and horizontal (X and Y axis) direction, are designed to investigate the present position accuracy. Cutting conditions are chosen with the focus on keeping the cutting forces and tool vibrations to a minimum, preventing an additional reduction in tool position accuracy (Table 1).

Table 1: Used cutting conditions

Tool	10 mm flat end mill
Material	Aluminium
Spindle Speed	2500 [rev/min]
Feed	200 [mm/min]
Pick feed	0.5 mm

4.2 The Two Work Pieces

The vertical work piece has a single step geometry of 40 mm, including a reference surface. (Figure 5a and 5b). The horizontal work piece has a circular slot geometry, with a inner diameter of 50 mm (Figure 6c and d).



(a) (b) (c) (d)

Figure 6: The vertical and horizontal work pieces.

The horizontal work piece is measured with the CMM, sensing the inner vertical surface of the circular slot, as shown in Figure 7. The vertical work piece is measured by sensing the straight step geometry.

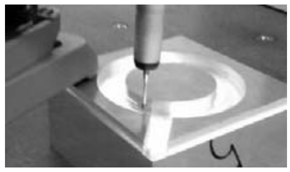


Figure 7: Measuring the milled geometry with the CMM.

4.3 Measurement Locations

The workspace of the HexaM PKM milling machine is shown in Figure 8. The investigated workspace is a rectangular volume, measuring 400 x 400 x 300 mm. The workspace is partitioned into a 3x3x3 mesh, resulting in a layout with 27 locations. Both work pieces are machined on each measurement location, indicated by (x,y,z) coordinates, as in Figure 9.

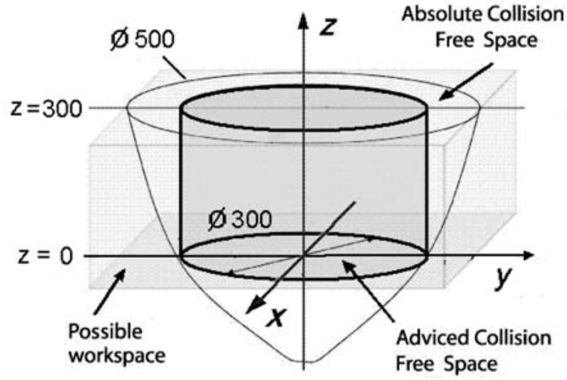


Figure 8: The workspaces of the milling machine

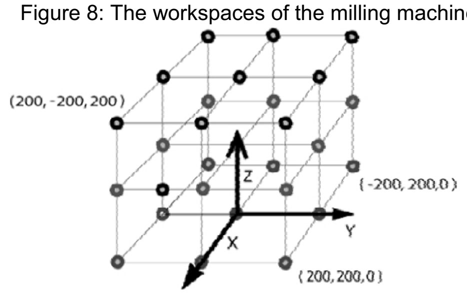


Figure 9: The 27 measurement positions in the workspace

5 DATA PROCESSING

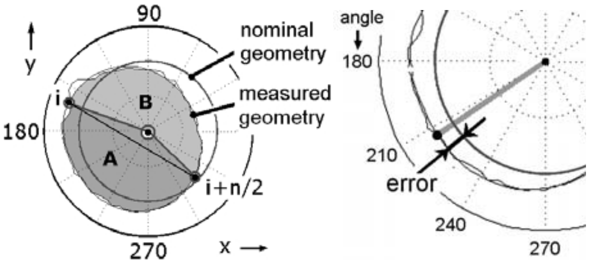
In order to effectively process and interpret all measurement data, some automatic data processing is applied. Because the two work pieces have a different geometry, they are dealt with separately.

5.1 Data processing The Horizontal Work Piece

The direction of the horizontal position error is found using the circular CMM measurement, consisting of n data points. As shown in Figure 10a, areas A and B are spanned by the origin and the datapoints ***i*** and ***i+n/2***. The value of ***i*** is calculating with the formula

$$\min_{i\in\{1 \ldots n\}} \left(A - B \right) \qquad (2)$$

The error direction and value are determined by the data point in de middle of i and i+n/2 (Figure 10b). A higher order polynominal is fit to the circular measurement data to rule out significant influence of the surface roughness on the error value.



(a) (b)

Figure 10: Horizontal error calculation method

5.2 Data Processing the Vertical Work Piece

Processing the data of the vertical work piece is less complex, because the desired result is an error value in one direction only (the z-direction). The average value of the line coordinates is the used vertical error value.

5.3 The Total Error Field and Interpolation Method

The original mesh of measurement locations (3x3x3) is not very dense: the vertical and horizontal increment are 200 and 150 mm respectively. A denser distribution of the error throughout the workspace desired, especially when machining larger work pieces or milling at locations between the original error data points. Therefore the original 3x3x3 mesh of data points is interpolated using cubic spline interpolation, resulting in a 9x9x9 mesh. This is a reliable method for creating more data points in the error field because the error fields are assumed to be continues in the first and second order and do not have rapid changing gradients. A vertical cross-section of the resulting interpolated error field is shown in Figure 11.

6 MEASUREMENT RESULTS

6.1 First Observations

The measured error field has a distinguishable pattern, with position errors directed away from the origin, and increasing with increasing distance from the origin. However, the pattern in the Z-direction is not very consistent. The absolute value of the horizontal and vertical error ranges from 0 to 30 μm, and 0 to 50 μm respectively.

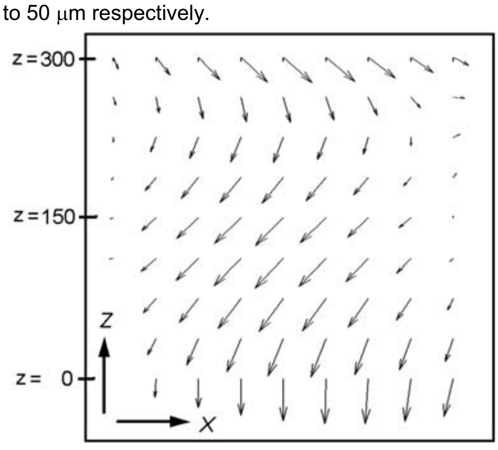


Figure 11: Vertical XZ measured error field at Y = 200.

6.2 Reproducibility

The reproducibility of the measurement is tested by machining the same work piece several times at the same location (Figure 12). Unfortunately, the reproducibility of the vertical work pieces seems to be poor. The variance in position error is high compared to the absolute position error. Since the CCM measurement itself has a good reproducibility, the large variance must have other sources.

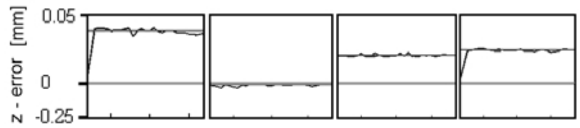


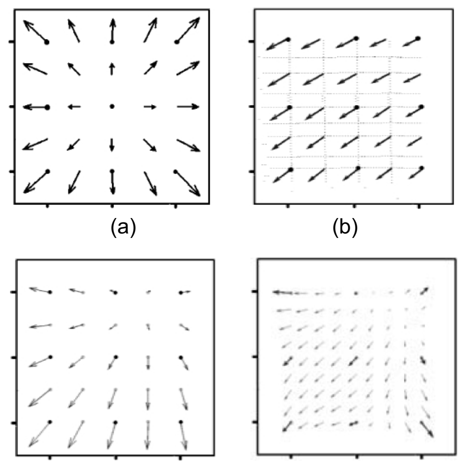
Figure 12: An example of poor reproducibility

7 POSITION ERROR COMPENSATION

The predicted and the actual present error fields are known and can be compared with each other. When the pattern of one of the predicted error fields is similar to the measured error field, the corresponding error source is likely to be a real error source in the Milling machine.

7.1 Comparing the Measured and Predicted Error Field

The predicted and measured error space in the horizontal direction will be investigated. Taking in consideration that a) the vertical error field does not show a characteristic pattern, and b) the reproducibility of this error is poor, there is no attempt made to use the indirect compensation method for the z-error. As for the horizontal error fields, the following observations are made. The horizontal cross-section of the measured error field (z=200 mm) shown in Figure 13d does not originate from one error source. The origin of the error in the milling machine may not be a single one, there may be several error sources simultaneously causing the measured position error. The kinematic model is given more than one error source as input and calculates the combined effect of these errors: all struts elongated with 1 mm (Figure 13a); struts 1, 2 and 6 elongated with 8 mm (Figure 13b); the errors of a and b added up in Figure 13c. The measured error field is shown in Figure 13d. When proceeding with the indirect error compensation, the error field in 13c is used to alter the NC code throughout the workspace.



(a) (b)

(c) (d)

Figure 13: predicted and measured error fields compared

Comparing the multi-source predicted error field with the measured error field (compare Figure 13c with Figure 13d) shows what seems to be a reasonable fit, but a certain misfit remains. A numerical method to find a better fit is briefly described in the recommendations. This fitting method can be convenient, but are not used in this research. The remaining misfit implies that there are error sources affecting the position accuracy, which are not yet taken into account by the kinematic model.

It is for this reason no attempt have been made to compensate the NC-data with the predictions of the kinematic model. The fit is simply not accurate enough: in many locations throughout the workspace, this indirect compensation will result in a bigger error than when machining the original work piece without compensation.

7.2 Basic Compensation of the XY-error

The indirect compensation method is not yet accurate enough to generally use when creating compensated NC-code. The other method, direct compensation, can still be used. First of all, the direct compensation of the X-Y error will be executed.

The point with coordinates (x,y,z) = (200,200,150) is chosen for testing this method, because here the measured error is relatively large. Since this position is exactly between two measurement positions (z = 200 mm, z = 200 mm) insight in the accuracy of the interpolating method can be achieved. The X and Y errors are measured in this position, together with the interpolated error value The basic compensation is executed according to

$$\boldsymbol{v}' = \boldsymbol{v} + \boldsymbol{e}_{com} = \boldsymbol{v} - \boldsymbol{e}_{err} \qquad (3)$$

v' is the compensated position coordinate, ***v*** the nominal coordinate. $\boldsymbol{e}_{comp}$ and $\boldsymbol{e}_{err}$ are the compensation and error vector. In Figure 14, The difference between the error pattern of the uncompensated (a) and compensated (b) work piece are shown. Figure 14c shows the expected compensated error pattern: the original error pattern translated over a distance equal to the compensation. In these plots, the inner circle indicates the nominal value; the outer circle indicates an error of +15 µm in radial direction.

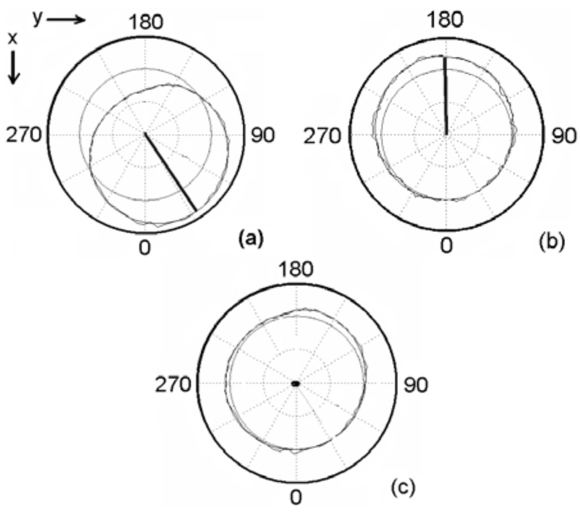


Figure 14: Direct compensation in the X-Y direction

7.3 Basic Compensation of the Z-error

For the z-direction the same compensating strategy is used as for the x and y direction. In this case, we have only one error to compensate. Because of the reproducibility problems, finding the z-error after compensation is done by machining and measuring not one but four work pieces; and calculating the average of the found errors.

All position errors, before and after compensation, are listed in Table 2. The relative errors are calculated by dividing the absolute error with the measurement distance (circle: 50 mm; step 40mm).

Table 2: position error before and after direct compensation

	Average Measured error				
Error Direction	**Before Compensation**		**After Compensation**		**Error reduction**
	[mm]	%	[mm]	%	%
X	0.0479	0.096	- 0.0312	0.062	56
Y	0.0359	0.072	0.0004	0.001	99
Z	0.0172	0.043	- 0.0173	0.043	- 1

The z-error after compensation shows the compensation method is less effective in this direction: the error is now in the opposite direction. Apparently, the machine's behaviour has changed and now has a small position error: it has shifted over a distance of 17 µm. The intended compensation has introduced all the remaining error; in other words, approximately all the error found after compensating is the compensation itself.

Considering the reproducibility problems encountered, this behaviour is not a big surprise. To this point, the poor reproducibility holds back further progress in indirect and direct compensation.

8 CONCLUSIONS

Using the kinematic model to gain insight to the position error sources of the milling is a promising method. It can be a very effective way to find error sources and achieve a accurate milling process. It can also be used as a calibration tool. When this method is fully mature, the accuracy of the machine may be improved greatly.

The compensation method seems to be effective; the error has been reduced with more than 50%. However, the compensation of the x and y-error seems to be significantly more effective than the z-error. The remaining Y-error is not significant; it can be caused by the sample interval of the measurement or by the data processing. It is not certain if the remaining x-error is a reproducibility error or not. The expected pattern does predict a slight remaining error in the positive x-direction, but not as big as measured after compensation.

This model is only suitable for low speeds and accelerations, because it is a static kinematic model. Since a PKM milling machine has highly dynamic properties, this model has its limitations.

The fairly good reproducibility of the horizontal work piece, together with the misfit made by the predictions of the kinematic model, suggest that there are error sources present that are not yet taken in account by the kinematic model.

The used indirect measurement method has a poor total reproducibility. The reproducibility of the horizontal work piece is reasonable, but the reproducibility of the vertical work piece is poor.

As for the measured error field, two observations can be made. In the horizontal direction some characteristic patterns can be distinguished. The vertical errors, on the other hand, do not show any clear trends.

The CMM measurement has a good reproducibility and does not need further investigation in future research.

Improving the error investigation method does not improve the reproducibility of the milling process. The last practical step of this research is to mill a error compensated work piece. If the measurement method were flawless, the poor machining reproducibility will still remain. If this reproducibility is poor, extra attention to improve it is needed.

9 RECOMMANDATIONS

This research did not reach an endpoint and further research is needed. In order to accomplish a better understanding of all errors affecting the accuracy of a PKM milling machine, several recommendations are made, listed in different categories below.

9.1 Experimental Method

Using a Double Ball Bar as a measurement device will probably increase the reproducibility of the position error investigation. This is only part of the solution, because in some cases the milling process itself has a poor repeatability. The performance of the compensation method using the predictions of the kinematic model still depends on the final product: a compensated work piece. If this final product has a poor repeatability, the whole method is a much less convenient.

If the error compensation is built in the machine controller, the NC data do not have to be changed manually or by a CAM program.

9.2 Improving Machine Behaviour

Better control of environmental influences on the milling machine, including temperature control, will probably give a much better overall reproducibility. A better understanding of internal and external thermal effects will facilitate in improving the reproducibility of the milling experiments. Climate control of the machines environment may improve its reliability. Another option is to equip some of the components of the machine with thermal sensors or strain gauges.

9.3 Improving the Kinematic Model

The kinematic model is very promising but has its limitations. There might be more relevant error sources present that are not yet implemented into the machine. It is recommended to search for these extra error sources.

The PKM machine's advantage is its high dynamic behaviour. The model only calculates static situations, and it is recommended to expand it with dynamics. When inertias and velocities of the machine components are included, the model will reach it full potential.

Furthermore, the Kinematic model could greatly benefit from a simple users interface. Considering the fact that 1) the model has a limited number of input parameters (the magnitude of each error source) and 2) these errors have to be adjusted frequently to produce the predicted error fields, such a user interface will make the use of the kinematic model much more convenient.

9.4 Numerical calculation of the error sources

When all single-source predicted error fields–in this example a total of 3–are assumed to be linear independent, a linear combination of these error fields exist which results in a best fit to the measured error field. As described in [5], the formula used is equivalent to an Eigen value problem:

$$a\boldsymbol{v}_1 + b\boldsymbol{v}_2 + c\boldsymbol{v}_3 = \boldsymbol{J}_{calc} \tag{4}$$

With $\{a,b,c\} \in \mathbb{R}$ and $\boldsymbol{v}_i$ vectors containing all data points of the predicted single-source error fields, and $\boldsymbol{J}_{calc}$ the calculated multi-source error field.

With $[\mathrm{A}] = [\boldsymbol{v}_1\ \boldsymbol{v}_2\ \boldsymbol{v}_3]$ and $\boldsymbol{x} = [a\ b\ c]^T$ (5)(6)

we get the more compact notation:

$$[\mathrm{A}]\boldsymbol{x} = \boldsymbol{J}_{calc}$$

Using least-square optimization, the desired $\boldsymbol{x}$ can be found with

$$\min_{\boldsymbol{x}\in\mathbb{R}^3} [\boldsymbol{J}_{calc}(\boldsymbol{x}) - \boldsymbol{J}_{meas}]^2 \tag{7}$$

This proposed numerical method delivers the best fit of $\boldsymbol{J}_{calc}$ to $\boldsymbol{J}_{meas}$, the measured error field. Future research may benefit from this numerical method, because is saves time and effort finding a good fit.

10 ACKNOWLEDGEMENTS

The authors would like to thank Mr. Y. Ueda and Mr. Teramoto. of the Department of Mechanical Engineering, Osaka University, for their advice and practical assistance with the machining experiments.

11 REFERENCES

[1] Pastoor, S., 2005, Improvement of Machining Accuracy for Parallel Kinematic Machine Tools using Error Prediction and Compensation for CAM Application, Proceedings of the 9th International Conference on Mechatronics Technology (ICMT2005), Kuala Lumpur, Malaysia, Dec 5-8, 2005.

[2] See http://www.jtekt.co.jp/e/products/machine.html

[3] ISO 230-4, Test code for machine tools, Part 4: Circular tests for numerically controlled machine tools.

[4] Kakino et al., 2000, Kinematic calibration of a hexapod machine tool by using circular test, Proceedings of the 2000 Japan USA flexible Automation Conference, July 23-36, Ann Arbor, Michigan. (2000JUSFA-13100)

[5] Lay, D.C., 1997, Linear Algebra and its Applications, 2nd Edition, Pearson.

Optimization of Toolpath Generation in Medical CAM for a Machine Tool for Orthopedic Surgery

Naohiko Sugita[1], Takayuki Osa[1], Yoshikazu Nakajima[1] and Mamoru Mitsuishi[1]

[1] Department of Engineering Synthesis, School of Engineering, The University of Tokyo, Japan

Abstract

Toolpath generation and optimization is considered as a challenging problem in the minimally invasive orthopedic surgery with a milling robot. The objective of this paper is to minimize the collision of the cutting tool with the soft tissues. A novel approach of toolpath generation and optimization is proposed. A redundant axis is implemented to avoid the collision in the robot. Some important components are modeled based on the physical requirements. A geometric optimization approach based on the model is proposed to improve the toolpath. Case studies show the validity of this approach. Software is developed for this application and the effectiveness is evaluated with a cadaveric bone.

Keywords:

CIRP International Conference; Manufacturing System; Paper Instructions

1 INTRODUCTION

In the surgical procedure of joint replacement, the setting position of artificial joint affects the inferior limb direction after the operation. Therefore, postoperative pain, reduction in the useful lifespan of the artificial joint, and organization necrosis due to abrasive wear particles shed by the bone-joint system will occur if the artificial joint is not properly fixed. The accuracy of the cut surface conventionally depends on the surgeon's skill, because the bone is shaped by hand. Therefore, the use of a bone-cutting robot is expected to increase the quality of the surgery, and many robots have been developed worldwide.

Meanwhile, the number of surgical procedures with minimally invasive technique has been increased also in orthopedic field. The minimally invasive surgical approach utilizes small incisions and offers several advantages over traditional open surgery as reduced pain and trauma to the body, faster recovery and shorter hospital stay. New ways to open the knee may be more important than the length of the incision. However, the difficulty of the procedure increases with the smaller incisions, and the result of the operation depends on the surgeon's skill. Therefore, it is hoped that a mechanical or robotic assisted surgery system is introduced in the procedure.

Invasiveness, accuracy of cut surface, high efficiency and safety of machining are required in the minimally invasive orthopedic surgery mainly. The minimally invasive knee replacement technique attempts to accomplish the procedure through a smaller incision, and the knee joint is accessed through the quadricep tendon. This is said to be less invasive to soft tissues and or bone. The shape accuracy of the setting plane is important to fit the artificial joint, and the setting absolute position of the artificial joint should be precise. Operation time for bone cutting is limited to about 15 minutes in terms of necrosis. The safety of machine and tools is also required.

Many of developed robots so far including our multi-axis bone cutting robot uses an end mill as a cutting tool, and the following problems should be solved to apply them to the minimally invasive orthopedic surgery.

The minimally invasive surgery makes the incisions smaller, reduces pain and trauma to the body and helps faster recovery. The smaller incision means the small and narrow opening area. This results that the robot attitude for the bone resection is restricted, and it can result in the collision of tool with the surrounding tissue, the existence of untouched area and the degradation of joint position accuracy. Toolpath generation technique specialized for the bone cutting is expected to resolve these issues.

Also the collision of the cutting edge with the soft tissue should be taken into account as a problem of invasiveness. The end mill is a rotational tool, and all the angles around the shaft function as a cutter. Therefore, it is likely to damage the surrounding soft tissues, vessels, nerves. The protection mechanism to cover the unworking part of the cutting edge will be required and avoid the damage. The necrosis of bone cell caused by the cutting heat or the tool friction heat should be prevented by cooling the cutting edge.

In this paper, a toolpath generation method is proposed to minimize the damage of the surrounding tissues in the robotic assisted minimally invasive orthopedic surgery. With the method, cutting tool can approach the resection area through a narrow opening area, proceed the machining of the bone without any damage and accomplish the procedure.

ROBODOC has been developed as a robotic orthopedic surgery system [1] and is most famous in the orthopedic field. It has been used in many clinical operations. The recent orthopedic robots have unique features. A robot works passively supporting the surgeon, and another is downsized and mounted on the bone directly. For example, "ACROBOT" developed by Davies,B. et al. is former, and it is used clinically [2]. Dombre,E. et al. developed "BRIGHT", and a guide jig for the bone saw is implemented on the tip of robot arm [3]. "ARTHROBOT" by Kwon,D.S. et al. aimed at the minimally invasive joint replacement [4], and a robot by Plaskos,Ch. can be set on the bone directly [5]. The recent tendency is focused on the minimal invasiveness of surgical procedure in addition to the high accuracy.

This study focuses on the issues of the tool attitude control and the toolpath generation due to the narrow opening area, and these problems are resolved in a robotic assisted minimally invasive surgery. Toolpath optimization problems themselves have been discussed in manufacturing systems and robot manipulators for a long term. The toolpath generation in the multi-axis machining tool and the path planning to avoid the interference in the robot manipulator have been studied.

For example, Morishige,K. et al. applied the concept of the configuration space to the collision avoidance for 5-axis control machining. The concept was studied in the robot field originally [6]. Pritschow,G., et al. presented the design and test of a fail-safe numerical control (NC) for robotic surgery, which has assisted in a wide range of surgical treatments [7].

The minimal invasiveness in the robotic assisted orthopedic surgery comes to the toolpath generation study. A cutting tool approaches the target through the small hole and resects the large area inside the joint. This paper proposes a toolpath generation method to cut the bone without damaging the surrounding tissues.

2 MACHINE TOOL DESIGNED FOR ORTHOPEDIC SURGERY

2.1 Concept

Considering the use in the operation room, the size of the machine is designed to 900mm, 1,800mm for width and height, respectively. The weight is approximately 300kg, and the developed machine tool has 7 degrees of freedom as shown in Fig.1. Mechanical and structural features are as follows. (1) High rigidity is realized by adopting a linear guide and a circular guide. The mechanical elements which are used for the robot realize high system rigidity compared with a conventional robot having rotational degrees of freedom. (2) The axes of all rotational degrees of freedom intersect at the same point. When the attitude of a cutting tool is changed, the other axis does not have to move for safety reasons. The bone cutting robot is located beside the operating table. The rigidity is 271 N/mm, 72 N/mm, 65 N/mm for U-axis, V-axis and W-axis at the home position, respectively.

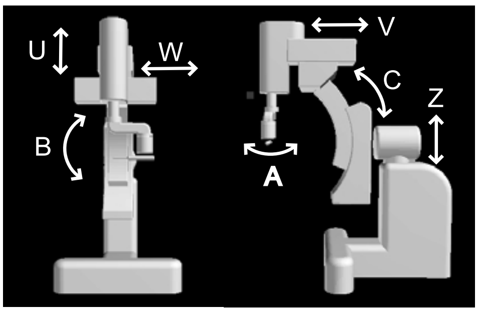


Figure1 : Overview of milling machine tool.

2.2 Redundant Axis for Minimally Invasion

The problems in the minimally invasive surgical procedure are to approach and resect the target bone through the narrow visible area. To solve these problems, the machine tool is equipped with a redundant axis (C-axis in Fig.1), and the cutting tool can avoid the interference like the soft tissues under the minimum change of the robot attitude.

Figure 2 shows the redundant axis and spindle with the cutting tool. The tool tip does not move during the rotation of this redundant axis, and the cutting tool approaches inside the joint and resect the target bone by controlling the tool attitude suitably.

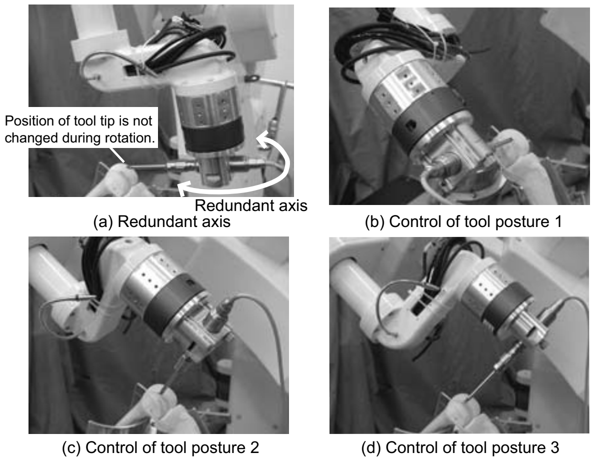


Figure 2 : Redundant axis for minimally invasion.

3 STRATEGIES FOR TOOLPATH OPTIMIZATION

3.1 Toolpath in Orthopedic Surgery

In the minimal invasiveness of the procedure, the control of the tool attitude and the toolpath generation are problems caused by the restricted opening area. As shown in Fig.3, the cutting tool approaches through the opening area (the range where the tool can enter) and resect the large part. In the total knee arthroplasty (TKA), there are 5 planes (anterior, anterior slope, distal, posterior slope, posterior) for femur, and the cutting tool needs to work deep inside the joint.

It is important not to collide with or damage the soft tissues during the bone resection, and it is necessary to calculate the optimum attitude of tool for that. Although the attitude of the cutting tool is restricted due to the small opening area, the degradation of the cutting accuracy is not allowed.

The preoperative CT images don't make clear the contour of the cartilage generally, and the posture of patient changes at pre- and intra-operation. Therefore, the precise shape and position of workpiece is determined at intra-operation, and the toolpath planning should be completed then. These environments are different from the industrial case.

These problems are unique in the surgical procedure, and this paper focuses on the toolpath generation in multi-axis bone machine tool. Specifically, as shown in Fig.4, the opening area, the interference plane, the resection part and the cutting tool are modeled geometrically. The task is to finish the target plane without any collision or damage, and the theme is how to solve this issue.

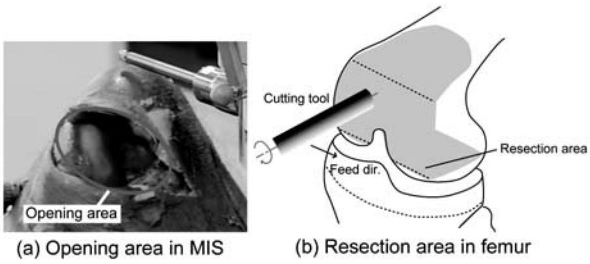


Figure 3 : Toolpath in Minimally Invasive Surgery.

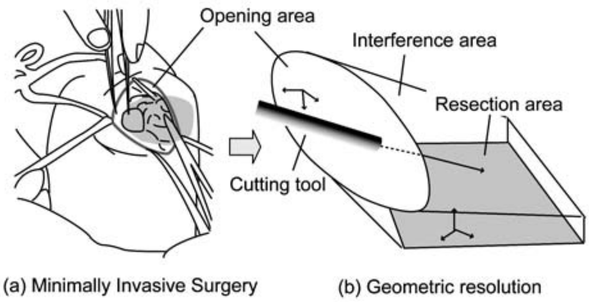


Figure 4 : Geometric model.

3.2 Proposed Toolpath Generation Strategy

Measurement of Opening Area

As described in former sections, the machining process of orthopedic surgery is different from the industrial one. The shape of workpiece is complicated, and the surrounding interferences also have complexity in their shape. This collision avoidance problem comes to a geometric issue. First, the opening area, the interference planes and the resection part are measured and calculated with a 3-dimensional optical position instrument (Northern Digital Inc., Polaris). Based on the data of measured points, a plane is generated statically, and a geometric model is constructed (Fig.5).

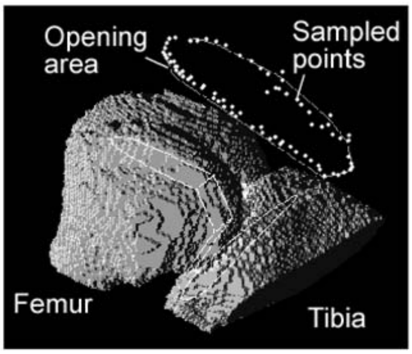


Figure 5 : Measurement of opening area.

Caluculation of Tool Posture

Next, it is determined which direction the cutting tool should approach from, and it is calculated how the attitude should be to cut the bone. Considering the collision with the interference, a tool attitude is obtained to cut all the target area. If there is no attitude to meet the condition, an attitude to maximize the possible area is selected. On the other hand, when some attitudes can attain the purpose, an attitude to minimize the required opening area is selected. The selected one is set to the attitude for the approach and the toolpath generation (Fig.6).

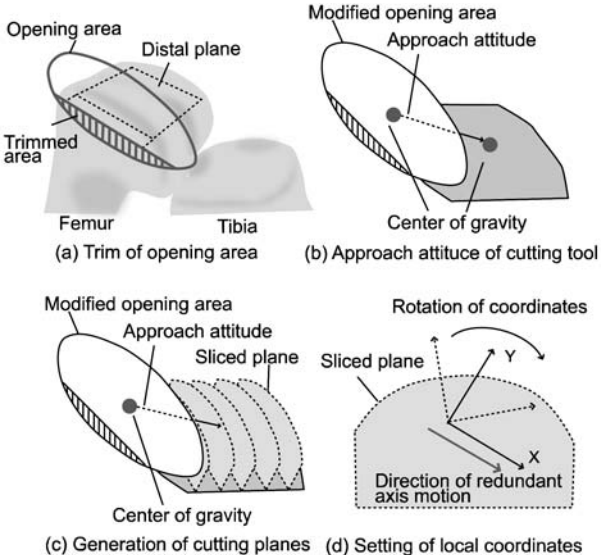


Figure 6 : Calculation of tool attitude.

Generation of Cutter Location

With the attitude calculated in previous step, cutter location (CL) is computed. Zigzag path is used for it, and other parameters like cutting direction (down cutting, up cutting) can be set by the software. By those steps so far, initial tool attitude and cutting location are determined (Fig.7).

Generation of NC Code

Tool attitude is computed for each CL point. Rotational B-axis and C-axis in Fig.1 are used for the initial tool attitude, and the angle of more one rotational and redundant axis fixes the tool posture during the bone cutting. There are two ways to move the redundant axis. One is what the attitude of the tool is modified continuously. The redundant axis is not changed as long as the tool collision with the soft tissue is not occurred. It should be discussed which is better way. In either way, the tool tip is controlled so that it moves along the cutting location. Finally, collision with the interferences is checked, and if the collision is detected, tool attitude is modified to avoid it with other rotational axis.

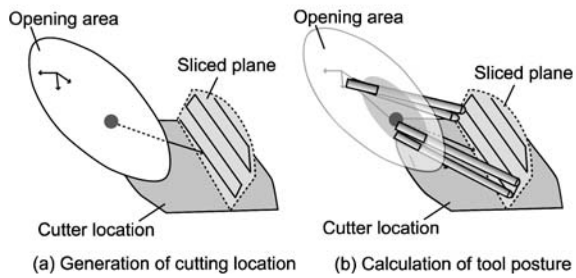


Figure 7 : Generation of cutter location.

4 EXPERIMENTAL RESULTS

When the proposed toolpath generation method above is applied to the robotic assisted surgical procedure, the effectiveness is evaluated by the motion of the cutting tool during the process. The shape of workpiece is unique in the minimally invasive joint replace, and there are many interferences and a small opening area. The toolpath generation algorithm is implemented in the processor for the multi-axis bone cutting robot, and the cutting location data without any collision is generated automatically.

In the experiment, unicondylar knee arthroplasty (UKA) and total knee arthroplasty (TKA) are targeted, and a cadaveric bone was cut. In UKA case, 3 surfaces for femur (distal, posterior slope and posterior) and 1 surface for tibia are cut using the developed system. As shown in Fig.3, most of the resected part is surrounded by the soft tissues, and the collision avoidance is required by controlling the tool attitude optimally.

Figure 8 shows the cutting location (CL) data generated for femur automatically. After resection plane, interference planes and opening plane are measured and computed, the data of the cutting location and the tool attitude is generated along the algorithm. In the CL data generation, some conditions are constrained. The tool attitude is normal to the distal and posterior slope planes, and is parallel to the anterior, anterior slope and distal planes. The calculation of initial approach attitude is simplified.

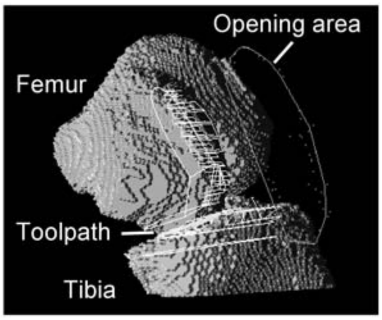


Figure 8: Generated toolpath.

The cutting conditions are set to Tool rotational speed 12000 rpm, Feed rate: 400 mm/min from the experiment so far, and a conventional end milling is conducted.

Figure 9 shows the overview of experiment and the tool motion behaviour for UKA. The length of incision is about 100 mm, though many of the bone cutting robots required the length from 150 mm to 200 mm. This experiment proved the possibility to cut the bone without any damage of the surrounding tissues and that the motion in the collision avoidance is smooth.

5 SUMMARY

In this paper, a redundant axis to avoid the interferences with the minimal attitude change was implemented in the multi-axis bone cutting robot for the minimally invasive joint replacement surgical procedure. A strategy of toolpath generation was proposed to accomplish the procedure, and it aimed at the approach through the narrow opening area and the machining without the damage of soft tissues. To realize this method, some techniques were described. This method proposed in this paper is applicable to the universal milling robots. Finally, a cadaveric experiment was conducted with the incision length 100 mm, and the toolpath and the tool attitude were evaluated in the minimally invasive procedure. Conclusions are as follows.

(1) A problem of tool collision came to the geometric one in the minimally invasive joint replacement, and it was defined as a problem to cut the large area of femur and tibia through the small and narrow opening area.

(2) With geometric models, an algorithm to determine the tool attitude was proposed, and it enabled to generate the toolpath without the collision and to finish the surfaces.

(3) The generated CL data was converted to the NC data in the post processor, and a cutting experiment was conducted. In the experiment, the cutting tool did not collide with the soft tissues, and accomplished the surgical procedure, and the effectiveness of the proposed method was confirmed.

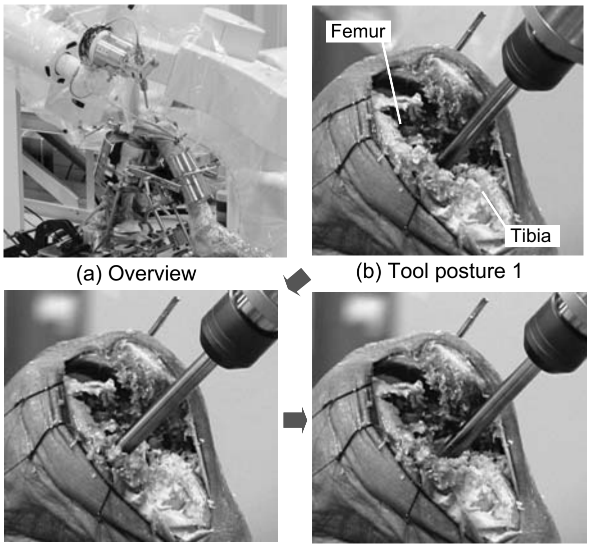


Figure 9 : Overview of experiment.

REFERENCES

[1] L. Joskowicz, R. Taylor, et al., 1995, Computer Integrated Revision Total Hip Replacement Surgery: Preliminary Report, MRCAS'95, 193-202.

[2] F. Rodriguez, S. Harris, M. Jakopec, A. Barrett, P. Gomes, J. Henckel, J. Cobb and B. Davies, 2005, Robotic clinical trials of uni-condylar arthroplasty, Int J Medical Robotics and Computer Assisted Surgery, 1/4:20-28.

[3] P. Maillet, B. Nahum, L. Blondel, P. Poignet, E. Dombre, 2005, BRIGHT, a Robotized Tool Guide for Orthopaedic Surgery, ICRA2005, 212-217.

[4] J.H. Chung, S.Y. Ko, D.S. Kwon, J.J. Lee, Y.S. Yoon and C.H. Won, 2003, Robot-assisted femoral stem implantation using an intramedulla gauge, IEEE Transaction on Robotics and Automation, 19/5:885-892.

[5] Ch. Plaskos, et al., Safety and Accracy Considerations in Developing a Small Sterilizable Robot for Orthopaedic Surgery, 2005, ICRA2005, 942-947.

[6] K. Morishige, K. Kase and Y. Takeuchi, 1996, The Method of Collision Avoidance for 5-Axis Control Machining Using 2-Dimensional Configuration Space, J. JSPE, 62/1:80-84.

[7] T. Bürger, U. Laible and G. Pritschow, 2001, Design and Test of a Safe Numerical Control for Robotic Surgery, Annals of the CIRP, 50/1:295-298.

Development of CAM system to Estimate Tool Wear of ball-end-milling

CheolWon Song[1] and Hideki Aoyama[2]

[1] School of Integrated Design Engineering, Keio University, Japan

[2] Department of System Design Engineering, Keio University, Japan

Abstract

This paper presents a new CAM system to estimate tool wear of ball end milling for cutting of curved surface. In the development of the CAM system, tool wear is predicted at each point on the cutting edge according to tool wear prediction equations derived from experiment data and cutting parameters: real cutting length, real feed-rate, and depth of cut in radial and axial directions, derived by an NC simulator installed in the CAM system. As the result, the effectiveness to estimate tool wear of a ball end mill was confirmed with basic experiments.

Keywords:

CAM system; Tool wear; Ball end mill

1 INTRODUCTION

Development of tool wear prediction systems for cutting process has been well recognized in industry due to the ever-increased demand for high quality production and productivity improvement. More specifically, increasing market demands on productivity and accuracy of the manufacturing process makes it necessary to use new methods and tools when developing NC machine tools [1-3].

Especially, ball end milling is one of the most widely used cutting processes for free form cutting in the automotive, aerospace, die/mold, and machine parts industries, and the tool wear prediction of a ball end mill is an important issue to be realized for high precision machining.

Many researchers have researched about tool wear estimation. Researching methods divided the in-process monitoring method and the analysis method. In-process monitoring method is based on AE signal[4-6], cutting force[7-8] and direct measurement of tool wear. This method is effective to judge the tool life in steady state cutting process. However, the method is not detecting of detailed tool wear. The analysis method is generally estimated tool wear by analyzing cutting force and cutting temperature in flat end milling processes and turning processes. However, there is not so much tool wear prediction in ball end milling.

Moreover, CAM is checked merely geometrical data. From this data, during cutting process inspects geometrical data. For the reason, CAM can not make a plan in real cutting process[9] and tool wear prediction.

Conventionally, cutting tool change timing is mainly determined by the cutting time. This method is useful in the manufacture of the same product by the same tool. But generally the die and mold is produced in single item production. And in curved surface manufacturing by the ball end mill, cutting points depend on the manufacturing shape, and the wear speed on cutting edge does not correspond linearly with cutting time and cutting conditions, which means that the tool change method based on the cutting time is not useful in this case. If tool wear distribution can be predicted, high-accuracy manufacturing can be realized by the generation of cutter paths taking into account processing error, and the most appropriate tool change timing can be determined.

This research therefore aims to develop CAM system that estimates tool wear distribution of ball end milling. In the development of the system, wear forms are investigated in cutting experiments, equations to predict tool wear from the real cutting length and cutting conditions has derived, and a module to calculate the real cutting length under cutting conditions at cutting points with different radius is constructed.

The usefulness of the system was also clarified. The following provides a brief summary of this research.

2 CHARACTERISTIC OF TOOL WEAR

Tool wear is variable according to real cutting length at each cutting radius in ball end milling. In this section, the concept of real cutting length and the relation between characteristic of tool wear and real cutting length will explain.

2.1 Real cutting length

Cutting length is the relative distance moved by the cutting tool. In the turning process, cutting length is the length cut by byte. Cutting length in the end milling process represents the length moved by the cutting tool relatively. Figure 1 shows the cutting length in end milling and ball end milling process. Cutting length is usually marked by feed length. However, the real cutting length is different because the radius of revolution differs according to the cutting edge position similar to the relation between the cutting speed and real cutting speed in ball end milling. Real cutting length is therefore distinguished from the cutting length.

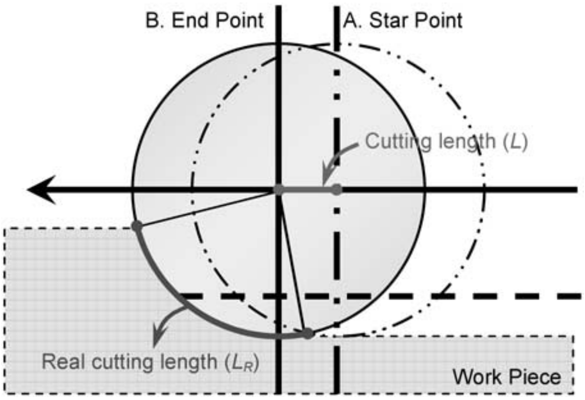


Figure 1: Cutting Length(L) and Real Cutting Length(L_R)

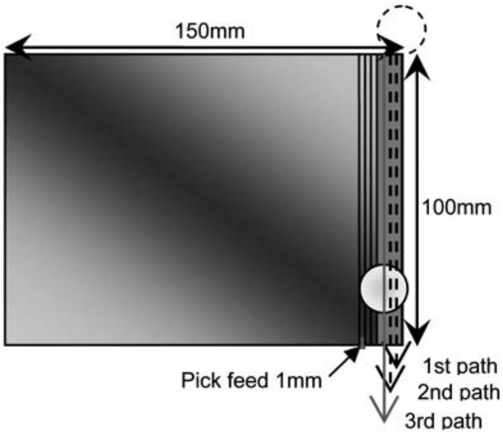


Figure 2: Method of tool wear experiments

Real cutting length indicates the distance actually cut by the cutting edge per radius of revolution. Figure 1 shows real cutting length of the cutting tool per revolution.

2.2 Characteristic of tool wear

For clarifying the relation between tool wear and real cutting length make the tool wear experiments in ball end milling process. Cutting conditions was changed RPM, feeding speed and tool radius by using the MAKINO milling machine. Table 1 shows the cutting condition of tool wear experiments. Figure 2 shows the tool wear experiment method. Tool wear measured in every 5 or 10 paths(1path is 100mm of cut).

Figure3 shows flank wear form at cutting edge in ball end mill. In Figure 3, tool wear speed at center point and edge point of tool is faster than at middle point of tool. The tool wear form was observed all of the cutting condition in experiments. This point is said point P that tool wear speed is the slowest point. P point is changing point of flank wear form in ball end millng.

Caracteristic of tool wear summarized with The following 1~7 from experimental results of Table 1:cutting conditions

1. The width of tool wear is proportional to the real cutting distance and to the second power of proportion or real cutting distance(L_R).
2. Initial tool wear does not depend on RPM(N) and cutter position and progresses more or less consistently.
3. The width of tool wear increases at each cutter position in proportion to the second power of the distance from the P Point.

Table 1: Cutting Conditions

Tool material	HSS(High speed steel)
Workpiece	S55C
Tool radius	5.0, 7.5mm
Rotation per minute	1000, 1500, 2000, 2500, 3000, 3500rpm
Feeding speed	0.05, 0.10, 0.15 mm/tooth
Peak feed	1.0mm/track
Number of cutter	2 cutters
Coolant	Dry cutting
Direction of cutting	Down cut

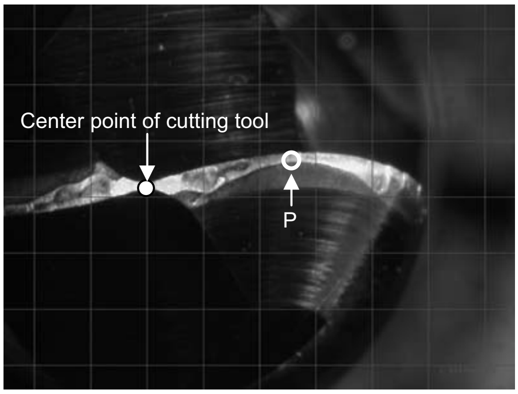


Figure 3: Point P of flank wear form in ball end milling

4. Tool wear changing point(P point) lies on the cutter, tool wear speed decreases with increasing radius of the blade where radiuses are smaller than the P point, and tool wear speed increases with increasing radius of the cutter where radiuses are larger than the P points.
5. P point(cutter radius) depends on RPM(N).
6. The tool wear speed increase with an increasing of cutting feeding(f_z)
7. The tool wear speed in the vicinity of the P point decreases considerably with an increasing in tool diameter(R_0).

3 EQUATION OF TOOL WEAR PREDICTION

The equation of tool wear prediction is derived from the measured data of tool wear distribution by a basic experiment in section 2.2. In this research, the tool wear prediction equation was derived for S55C with the HSS tool by the experiment.

The following prediction equation (1) was derived by using the least square method so that the relations in section2.

$$V_B = aL_R^{\ 2} + bL_R + c = a(L_R - p)^2 + q \qquad (1)$$

Where a, b, c are coefficients derived by tool wear experiments. The tool wear data obtained by tool wear experiment was calculated. The least square method was applied and approximated. The coefficients for each condition (RPM, cutting radius, tool diameter, and depth of cut) were calculated.

The tool wear equation (2) are derived in consequence.

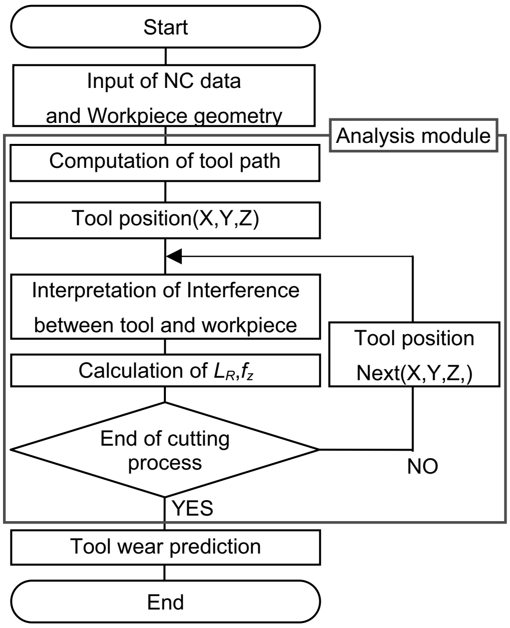


Figure 4: Diagram of the tool wear prediction system

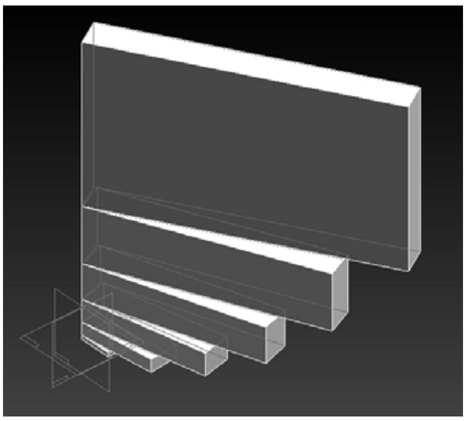


Figure 5: Division of Tool Model in Ball End Milling

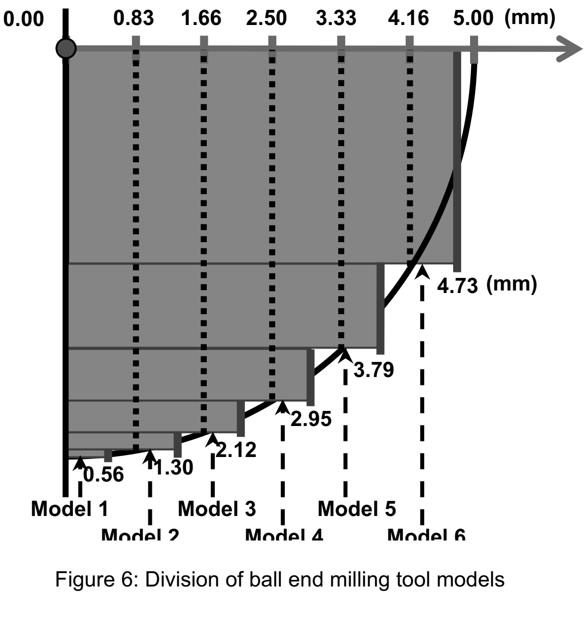


Figure 6: Division of ball end milling tool models

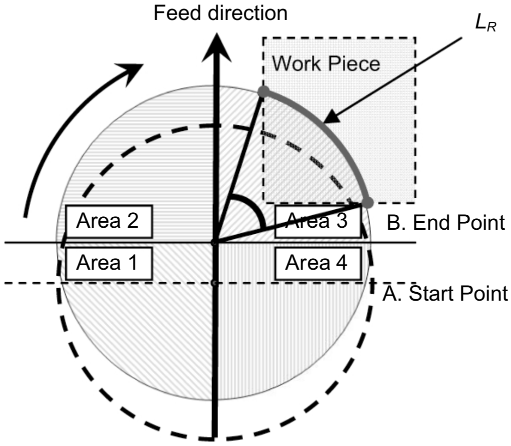


Figure 7: Algorithm for calculation of real cutting length(L_R)

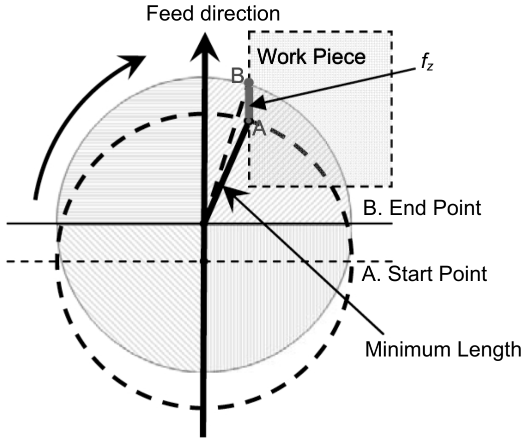


Figure 8: Algorithm for calculation of real feedrate(f_z)

$$V_B = (-0.335 R_0 + 2.67) \times (11.88 f_Z - 0.188) \times (9.103 \times 10^{-10} N + 9.95 \times 10^{-6}) \times \{(r + 7.68 \times 10^{-4} N - 4.65)^2 + 1\} \times (L_R + 2.87)^2 + 0.0249 \quad (2)$$

4 DEVELOPMENT OF CAM SYSTEM

The workpiece and NC data was entered into the analysis module for calculating the real cutting length and real feed-rate. Figure 4 shows the diagram of tool wear system.

The following section explains the analysis module(real cutting length and real feed-rate) and calculation methods.

4.1 Divided tool model

A system was constructed for calculating the real cutting length and real feedrate in ball end milling. The division models used the calculation algorithm of the real cutting length by the Z-map model.

It can predict also the exact tool wear and real cutting length calculation that will increase by the number of divisions. However, processing is complicated with number of divisions become large and the calculation speed will decrease. Therefore, there is no meaning even if do number of division more than critical number of parts. Figure 5 is shown the divided model in this research. The model is divided tool into 6 parts in Figure 6.

4.2 Algorithm for calculation of real cutting length(L_R)

As shown in Figure 7, the tool from area 1 to area 4 is divided by the feed direction line and horizontal lines. The real cutting length is calculated by deciding the tool and workpiece intersection in four areas. In this system, the workpiece model used is the Z-map model. This algorithm searches grids intersecting the tool in each area, and calculates the initial cutting point and final cutting point as shown in Figure 7. The real cutting length is calculated by the angle between those points and each cutting model's average radius of revolution. First, grid point A of the minimum length to the tool center searches from a grid that is intersecting the cutting tool.

4.3 Algorithm for calculation of real feedrate(f_z)

The real feed rate (f_z) is the depth of cut in the feed direction of cutting by each cutter of ball end milling. Depending on the workpiece model and cutting conditions, the real feed rate value per tooth can differ for the same cutting tool. Figure 8 shows the real feed rate(f_z). Then, grid point B searches the point in the tool progress direction from grid point A. Here, the real feed rate is the distance form grid point B further from grid point A and the point along the feed direction of the cutting tool from the tool center.

5 EVALUATION OF CAM SYSTEM

5.1 Tool wear experiment method

To evaluate the tool wear prediction system of free form surface, cutting was performed under the following condition in Figure 9(a); HSS ball end milling of diameter 10mm for the cutting tool, S55C for workpiece material, RPM of 2000, feed rate of 0.1mm/tooth in experiment, and axial depth of cut of 5mm. Figure 9 compared developed system data with experimental data.

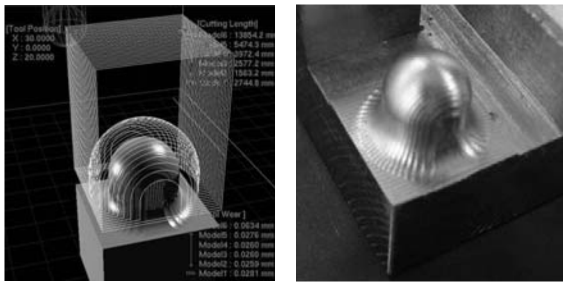
5.2 Evaluation of tool wear prediction system

Figure 10 compared result of tool wear evaluation experiments with result of developed system.

In center area of tool, there is a wide difference results between the system and the experiment. The used cutting condition for derived prediction equation was shoulder cutting. Therefore, those are supposed that evaluation experiment of free form surface cutting and tool wear experiments of sholder cutting is a wide difference cutting conditon. The results of tool wear prediction and the results of tool wear measurement are corresponding well in over 2mm of tool radius.

6 CONCLUSION

In this paper, the tool wear form of ball end milling was clarified. The equation for tool wear prediction was derived from tool wear experiment on ball end milling process. A module was constructed for analysis of cutting condition at each cutter position, and a tool wear prediction system was also developed using this module and tool wear prediction equation, and its usefulness was evaluated and discussed in this paper.



(a) System data (b)Experimental data

Figure 9:Comparison between system and Experiments

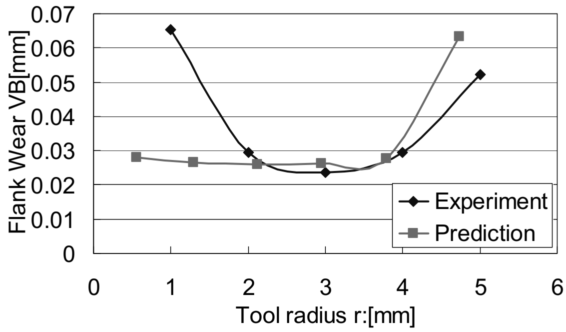


Figure 10: Comparison between system and Experiments

7 REFERENCES

[1] Bianchi.G., Paulucci.F., Van den Braembussche.P., Van Brussel.H., 1996, Towards virtual engineering in machine tool design, Annals of the CIRP 45 (1) pp.381-384.

[2] Miro.J.V., White.A.S., 2002, Modelling an industrial manipulator a case study, Simulation Practise and Theory (9) pp.293-319.

[3] M.J. Hendrik, J. van Brussels, 1996, Mechatronics-a powerful concurrent engineering framework, IEEE/ASME Transactions on Mechatronics 1 (2) pp127-136.

[4] Sampath.A., Vajpayee.S.,1987,Tool health monitoring using acoustic emission, Int.J.Prod.Res,25(5),pp703-719.

[5] Chi.D, Kwon.WT, Chu,CN, 1999, Real-time monitoring of tool fracture in turning using sensor fusion, International Journal of Advanced Manufacturing Technology,15(5),pp305-310.

[6] Roget.J, Souquet.P, Gsib.N, 1988, Application of acoustic emission to the automatic monitoring of tool condition during machining, ASNDT Materials Evaluation,46, pp.225-229.

[7] Altintas.Y, Yellowley.I, 1989, In process detection of tool failure in milling using cutting force models, ASME J.of eng,for Industry, 111, pp.147-157.

[8] Dimlasr D.E., Lister.P.M, 2000, On-line metal cutting tool condition monitoring. I Force and vibration analyses, International Journal of Machine Tools and Manufacture, 40(5),pp.739-768.

[9] Kimura.F., 1993, Product and Process Modelling as a Kernal for Virual Manufacturing Environment, CIRP Annals, Vol 42, pp.147-150,

Model based control of a piezo-actuated axis

Christian Brecher[1], Gunnar Schauerte[1] and Michael Merz[1]
[1] Fraunhofer Institute for Production Technology IPT, Aachen, Germany

Abstract
This paper presents the physical state space model of the static und dynamic behaviour of a piezo-actuated axis. The starting model parameters are fitted to measured data by an optimisation algorithm based on Newton´s method. The state space description of the axis is used to design a controller architecture with a model based state observer. The measured closed loop frequency response of the model based control is compared to simple linear controllers like PID elements. To reduce noise, the feedback signals for the model based control are estimated with a Kalman filter which bases on the linear model of the open-loop system.

Keywords:
control, observer, piezo-electric actuator

1 INTRODUCTION

Drilling bars with great overhang are used for the purpose of deep drilling. Vibrations of the drilling bar and changing cutting forces, caused from position and form errors of the pilot hole, are examples of interfering influences within the process. They result in a deteriorated quality of the drilled hole. Therefore, the development of mechatronic drilling tools with piezo-actuated axes is of increasing interest in order to improve the cutting process. Such piezo-actuated axes compensate for vibrations in the tool structure, minimise form and position errors in the work piece and enable new possibilities in the manufacturing of noncircular drillings.

The design of the presented drilling bar facilitates the use of two different piezo-actuated positioning systems. An integrated sensor-actuator-system in the top of the drilling bar enables a highly dynamic radial deflection of the cutting edge. Thus, the surface of the bore hole can be structured with a maximum stroke of 120 µm and a maximum frequency of several hundred Hertz. Form and position errors of the work piece are compensated for by three radially adjustable guiding pads, which are positioned with a distance of 120° on the circumference of the drilling bar.

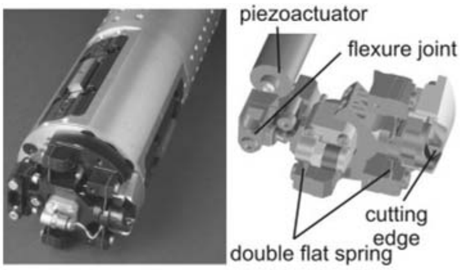


Figure 1: Head of the piezo-actuated drilling tool

2 LINEAR STATE SPACE MODEL

Figure 1 shows the head of the drilling tool (90 mm in diameter) with the axis of the cutting edge. The high voltage piezo-actuator pushes against a solid-state-joint, transferring the axial stroke of the actuator in a radial deflection of the cutting edge carrier which is supported by a double flat spring. The combination of the flexure joint and the double flat spring is defined as cutting-edge-joint. Strain gauges which are mounted on the flat springs detect the movements of the cutting edge.

The dynamical, electromechanical properties of the piezo-actuated axis can be described by an extension of the static sensor and actuator equations which are known from literature [1]. The presented equations are extended by the parallel connected cutting-edge-joint and the additional external process force F_L. The electrical charge q_P of the piezo-actuator is calculated by the sensor equation. It includes the charge q_V fed by the amplifier and the induced charge q_m. The induced charge q_m is calculated by the force F_e acting in the piezo-actuator, the piezo-electric constant d_{33} and the number of wafers n in the piezoelectric stack.

$$q_P = q_V - q_m = C \cdot U_V - d_{33} \cdot n \cdot F_e \quad (1)$$

The actuator equation

$$m_A \ddot{x}_A = k_P d_{33} n \cdot U_V - c_A \dot{x}_A - k_A x_A - F_L \quad (2)$$

characterises the movement x_A of the piezo-actuated axis, depending on the amplifier voltage U_V, the piezo-electric stiffness k_P, the stiffness of the axis k_A and the damping constant c_A. The third differential equation is obtained from the electrical circuit of the amplifier and the piezo-actuator.

$$L\ddot{q}_V = k_V U_e - R\dot{q}_V - \frac{1}{C} q_V \quad (3)$$

It includes the amplifier charge q_V depending on the amplifier inductance L, the electromechanical coupling coefficient k_V, the input voltage U_e of the amplifier, the piezo-electric resistance R and the piezo-electric capacity C. The established equation system can be easily transferred into the following block diagram (Figure 2).

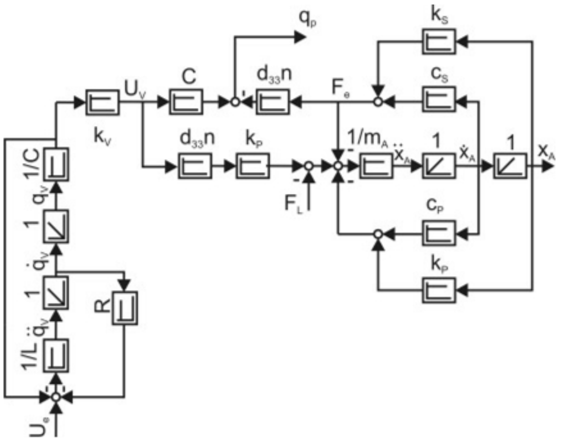


Figure 2: Linear model of the piezo-actuated axis

The matrices of the state space model

$$\boldsymbol{x} = \begin{bmatrix} q_V \\ \dot{q}_V \\ x_A \\ \dot{x}_A \end{bmatrix};\ \mathbf{A} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ \frac{-1}{LC} & \frac{-R}{L} & 0 & 0 \\ 0 & 0 & 0 & 1 \\ \frac{k_P \cdot d_{33}\, n}{C \cdot m_A} & 0 & \frac{-k_A}{m_A} & \frac{-c_A}{m_A} \end{bmatrix} \tag{4}$$

$$\mathbf{B} = \begin{bmatrix} 0 & 0 \\ \frac{k_V}{L} & 0 \\ 0 & \frac{-1}{m_A} \\ 0 & 0 \end{bmatrix};\ \mathbf{C} = \begin{bmatrix} 0 & 0 & 1 & 0 \end{bmatrix};\ \mathbf{D} = 0;\ \boldsymbol{u} = \begin{bmatrix} U_e & F_L \end{bmatrix}$$

with the transfer equation

$$\dot{\boldsymbol{x}}(t) = \mathbf{A}\boldsymbol{x}(t) + \mathbf{B}\boldsymbol{u}(t) \tag{5}$$

and the output equation

$$\boldsymbol{y}(t) = \mathbf{C}\boldsymbol{x}(t) + \mathbf{D}\boldsymbol{u}(t) \tag{6}$$

built from the linear electromechanical equations 2 and 3 of the piezo-actuated axis. The state space representation can be used for the simulation of the linear dynamic behaviour of the piezo-actuated axis and for the synthesis of a model-based control architecture. Figure 3 shows the simulated frequency response of the piezo-actuated axis. The simulation is a first approximation of the measured transfer function from the voltage input of the electrical power amplifier to the strain gauges on the cutting-edge-joint. The measured frequency response shows effects of an additional mass represented by a second natural frequency. This can be modelled as an additional mass damper. The effect on the system can be approximated by the extension of the one mass spring model to a two mass spring model of the piezo-actuated axis. The parameters of the damping oscillator are initialised with estimated values. The simulated frequency response of the two mass spring model shows an improved approximation to the measured data in the range of the natural frequencies.

3 OPTIMISATION OF STATE SPACE PARAMETERS

Using the real system's frequency response, the parameters for the linear two-mass-spring-model of the piezo-actuated axis can be verified and adjusted. The accuracy of the frequency response of the system with the calculated model parameters relating to the measured curve is quantified with the weighted discrete error sum

$$e(\boldsymbol{b}_i) = \sum_{\omega=0}^{\omega_g} \alpha(\omega) \cdot \left|\Delta G(\omega, \boldsymbol{b}_i)\right|^2 \tag{7}$$

with the squared distance of the simulated and the measured complex transfer functions $G(\omega, b_i)$ of the angular frequency $\omega = 2\pi f$ and the parameters of the calculated linear model $\boldsymbol{b}_i$. The weighting factor α allows for measuring the estimation error for the model parameters differently in designated frequency ranges. The endpoint of the summation of the error is the cut off frequency of the measured system ω_g.

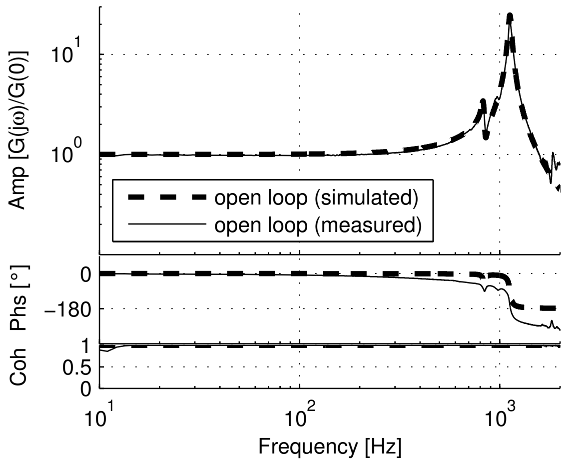


Figure 3: Frequency response of the piezo-actuated axis without position control (open loop)

The model parameters are tuned automatically with an optimisation algorithm based on Newton's method with the error function $e(\boldsymbol{b}_i)$ as objective function. The minimum of the function is found with the iteration step containing the gradient and the Hessian of e

$$\delta_i = -\mathbf{H}^{-1}(\boldsymbol{b}_i) \cdot \nabla e(\boldsymbol{b}_i) \text{ with } \mathbf{H}(\boldsymbol{b}_i) = \nabla^2 e(\boldsymbol{b}_i) \tag{8}$$

which results from a Taylor approximation truncated after the second term. The vector $\boldsymbol{\delta}_i$ points to the minimum in the multidimensional room spanned by the parameter vector $\mathbf{b}_i$ and is used for the update for the next iteration step applying

$$\boldsymbol{b}_{i+1} = \boldsymbol{b}_i + \beta \cdot \delta_i\,. \tag{9}$$

The weighting factor β, which is found with a linear search, stabilises this iteration step [2].

With appropriate starting points, the algorithm converges superlinear in the surrounding area of the minimum value.

Though, the stability of a pure Newton algorithm depends heavily on the preset starting point of the iteration $\boldsymbol{b_0}$. To produce a feasible starting point for the Newton algorithm from suboptimal starting points, a simplex algorithm is used in a preprocess. It converges maximally linear, but is generally more stable with parameter values more distant from the minimum point.

Optimised parameters enable a better approximation of the simulated frequency response to the measured one. A system identification toolbox only calculates a transfer function best fitted to measured data. The advantage of the approach described in this chapter is that the parameters of a physical model are optimised with the measured data. Under small-signal conditions the obtained linear physical system and the real measurement match by 96 %. The model-based approach is the basic requirement for complex control design. Another requirement for this goal is the simulation of nonlinear effects.

4 CONTROL DESIGN

For the piezo-actuated axis a cascaded control structure, usually applied for today's feed-drive systems, is applied. The dynamically slowest control cascade is the outer position control. The subordinate current control is the dynamically fastest control and has P-character. It damps the eigenfrequency of the series resonant circuit of the amplifier inductance *L*, the piezo-actuator resistance *R* and the piezo-actuator capacity *C*. The voltage control is superordinate to the current control and is designed as PI-control in order to achieve static accuracy.

Different control structures are tested for the design of the position control. The approaches are described in detail in [3]. The dynamic bandwidth of a PI-controller can be increased with the help of a Notch filter. With the right parameters it suppresses the resonance of the parallel connection of the piezo-actuator and the cutting edge joint.

A modern control strategy is the Internal Model Control (IMC). It requires the model of the cutting edge axis. The transfer function of the model is switched parallel to the real transfer behaviour. Instead of the controlled variable, the deviation between the real transfer function and the simulated model is fed back. In the ideal case with an accurate model and without disturbances, the feedback is inactive and the system becomes an open-loop control.

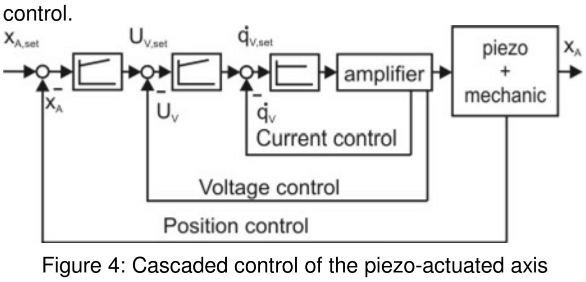


Figure 4: Cascaded control of the piezo-actuated axis

The bandwidth is equivalent to the cut-off frequency, defined as the frequency at an amplitude of -3 dB and a phase of -90°. The bandwidth of the PI-position controlled system, shown in figure 5, is 15 Hz. The PI-control reduces the amplitude of the eigenfrequency at 1122 Hz by a factor of 31 compared to the open-loop system. Using the IMC controller, the bandwidth of the closed loop system is increased to 300 Hz. The Notch filter control rises the bandwidth of the axis further on up to 450 Hz.

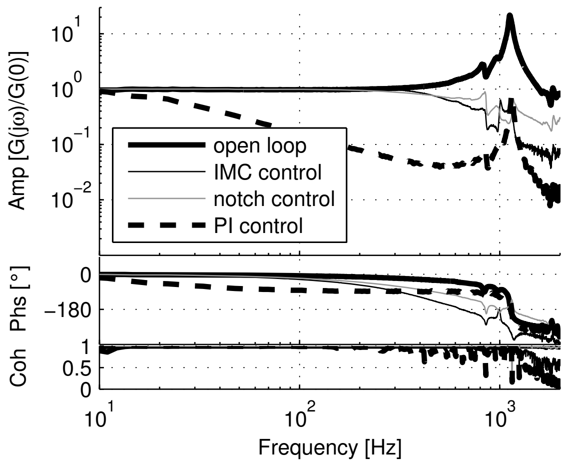


Figure 5: Frequency response of different position controls

The control loop can be improved dynamically by adopting a state regulator which feeds back the linear system's complete state vector [4]. Such a system can be seen as a systematic extension of the cascaded control structure using additional feedbacks. In a real system not every needed feedback represented as state in the model is measurable. A state observer based on the linear state space model (equation 4) can be used to compute the inner states from the inputs and outputs of the according system. Furthermore, the real system is affected by noise and disturbance. To model these influences, equation 5 and 6 are extended to

$$\dot{\boldsymbol{x}}(t) = \mathbf{A}\boldsymbol{x}(t) + \mathbf{B}\boldsymbol{u}(t) + \mathbf{B}_d\boldsymbol{d}(t) \tag{10}$$

$$\boldsymbol{y}(t) = \mathbf{C}\boldsymbol{x}(t) + \boldsymbol{n}(t) \tag{11}$$

with the process noise ***d(t)*** and the measurement noise ***n(t)***. With the assumption of observability and controlability of the state space system [4], the optimised matrix **L** of the observer (Figure 6) is found by minimising the expected value of the quadratic error with $\boldsymbol{e} = \boldsymbol{x} - \hat{\boldsymbol{x}}$ between the real state vector x and the observed state vector $\hat{\boldsymbol{x}}$. This observer is called Linear-Quadratic Gaussian (LQG) observer or Kalman filter [5].

$$J_{LQG} = \lim_{t\to\infty}\left\{\left|\boldsymbol{e}(t)\right|^2\right\} \to \min \tag{12}$$

The design of the optimised observer feedback **L**

$$\mathbf{L} = \mathbf{PCR}_n^{-1} \tag{13}$$

is obtained by solving the static algebraic Riccati equation

$$\mathbf{AP} + \mathbf{PA}^T - \mathbf{PCR}_N^{-1}\mathbf{C}^T\mathbf{P} + \mathbf{B}_d\mathbf{Q}_N\mathbf{B}_d^T = 0\ . \tag{14}$$

P is the positive definite solution of this equation. Since the noise signals are uncorrelated, the covariance matrices for the process noise $\mathbf{R}_N$ and the measurement noise $\mathbf{Q}_N$ become diagonal matrices. The entries of the matrices result from the power density spectra of the different signals. Assuming zero-mean white noise processes, the power density spectra have a constant value. Since the processes are modelled without mean value, the state space control with a Kalman filter alone is not usable for tracking control, where following errors for absolute values cannot be tolerated.

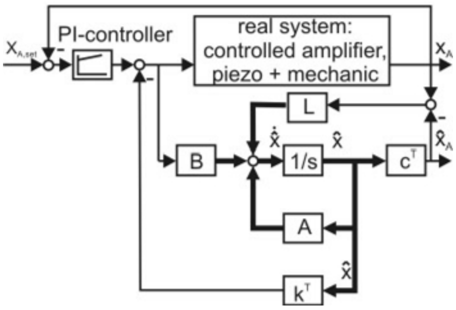


Figure 6: Block diagram of a state space tracking control

In order to use the Kalman filter in a tracking control, a PI-controller is introduced in an outer loop to achieve static accuracy. Figure 7 shows the frequency response of the piezo-actuated axis with a Kalman filter and a PI-tracking-control. The bandwidth is extended to 520 Hz compared to the Notch filter control.

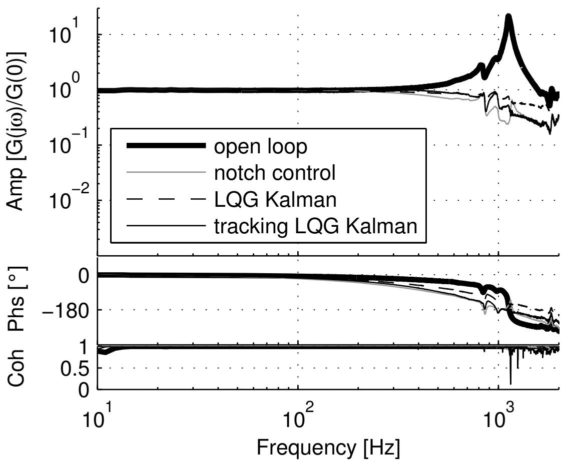


Figure 7: Frequency response of the state space tracking control

5 EXPERIMENTAL INVESTIGATION

In Figure 8 the result of a noncircular manufacturing of a cylinder liner with the piezo-actuated drilling tool is presented (rotation speed: 300 min^{-1}, infeed: 30 mm/min, cutting depth 0.175 mm). The cylinder liner (material: Al, drilling diameter: 93 mm) was predrilled on a conventional lathe with a roundness of 24 µm and a parallelism of 21 µm on the length of 140 mm. The piezo driven cutting axis was actuated with the Notch filter control. The cylinder liner was structured with the rotation speed of 300 min^{-1} and an amplitude of 12.5 µm proceeding setpoints with the frequency of 50 Hz. To synchronise with the basis machine, a setpoint generator reads out the angular position of the manufacturing spindle and the position of the feed-drive axis.

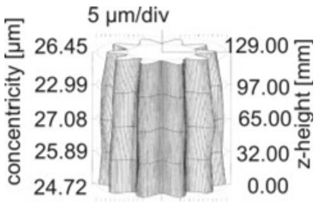


Figure 8: Noncircular manufactured cylinder liner

6 SUMMARY

Exemplified by a drilling bar, a piezo-actuated axis consisting of an amplifier, a piezo-actuator and a solid state joint was introduced. Depending on the physical and mechanical equations, a linear state space model was derived to specify the static and dynamic behaviour of the piezo-actuated axis. The model parameters were fitted to measured data by an optimisation algorithm based on Newton´s method. The linear model of the piezo-actuated system »cutting edge« is a useful support to design simple linear controllers like PID elements. The PID Controller can be easily extended with a Notch filter to damp the natural frequency of the system. The optimised model of the piezo-actuator can be used to design extensions of standard control structures like IMC controllers. The state space model of the piezo-actuated axis can be used to design a state space controller with a Kalman filter and PI-tracking control. This control extends the bandwidth of the axis up to 520 Hz which is 65 Hz better than a Notch filter control. Unfortunately, the design of a Kalman filter is very time consuming and the control structure is very susceptible to parametric changes in the controlled system. Thus, the Notch filter control achieves the best rate between use and effort.

7 ACKNOWLEDGMENTS

The authors would like to thank the DFG for their support which enabled the research program BR 2905/12-2 “Bohrstange mit integrierter hochdynamischer Zustellung sowie Schwingungs- und Formfehlerkompensation” within the program of emphasis “Adaptronik für Werkzeugmachinen”.

8 REFERENCES

[1] Schugt, M.: Aktor-Sensorverhalten von Piezoelementen in Kfz-Diesel-Einspritzsystemen; Shaker Verlag, 2002

[2] Fletcher, R.: Practical Methods of Optimization Vol. 1 (Unconstrained Optimization), Department of Mathematics University of Dundee (Scotland, U.K.), John Wiley & Sons, 1980.

[3] Brecher, C. Schauerte, G: Adaptronic Drilling Tool, Proceedings of the 2. colloquium; Adaptronic for Machine Tools, Braunschweig, 2007

[4] Föllinger, O.: Regelungstechnik 5. Edition; Heidelberg; Hüthig Verlag, 1985

[5] Maybeck, P. S.: Stochastic models, estimation, and control Vol. 1 – Mathematics in Science and Engineering, 1979.

Register Control of Sectional Drive Rotogravure Printing Press

Tsuguaki Yoshida[1], Shinichiro Takagi[1], Yasuhiko Muto[2] and Teilong Shen[2]

[1] Technology Development Center, Dai Nippon Printing Co., Ltd., Ibaraki, Japan

[2] Department of Mechanical Engineering, Sophia University, Tokyo, Japan

Abstract

In the latest type of rotogravure printing presses referred to as the sectional drive system, wherein each gravure cylinder is driven by an individual motor, tension fluctuations by an upstream register control generate register errors in downstream units. Hence, the introduction of new effective control system to prevent the generation of the downstream register errors is required. For this purpose, in this paper, a method of register control for printing presses has been proposed. First, a nonlinear coordinate transformation for the state variable of the printing press dynamics is proposed. Subsequently, based on the new coordinate system, a nonlinear control method is designed using the Lyapunov stability. Finally, the performance of the new control method is demonstrated by simulation and experimental results and comparing with a conventional PD control method.

Keywords:

Rotogravure Printing press; Nonlinear Control; Register

1 INTRODUCTION

In a rotogravure printing press, a picture is printed on a flexible, continuous material referred to as a web. Each color picture is sequentially printed in each unit consisting of a gravure cylinder, an impression roll and a dryer as shown in Figure 1. The majority of conventional rotogravure printing presses comprise more than five units.

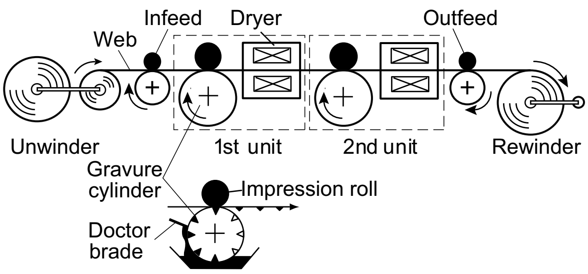
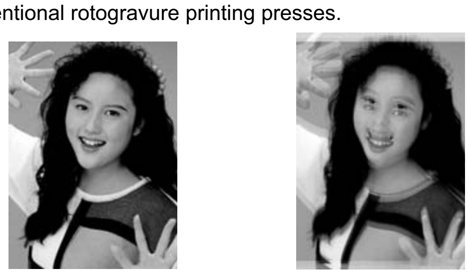


Figure 1: Rotogravure printing press

In a multicolour printed product, when the superposition accuracy of each color, referred to as a register, is not sufficient, the product appears defective, as shown in Figure 2. Therefore, the register is one of the most important qualities in printed products. To maintain a high register accuracy, registers are automatically controlled in the conventional rotogravure printing presses.



Accurate register   Inaccurate register

Figure 2: Register

In the latest type of the rotogravure printing presses referred to as the sectional drive system, each gravure cylinder is driven by an individual motor and all motors are electrically synchronized. The register is regulated by a phase difference between the gravure cylinders. The classical control law between adjacent units is employed for the conventional control systems, as shown in Figure 3.

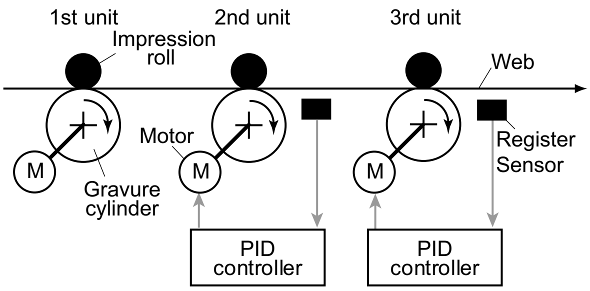


Figure 3: Conventional control method for sectional drive printing press

In this system, the upstream register control causes tension fluctuations that travel towards the downstream units with a time delay. As a result, downstream register errors are generated sequentially with the time delay. Therefore, the introduction of a new control system to prevent the generation of the downstream register errors is required.

For this purpose, in this paper, a method of nonlinear register control for the rotogravure printing press has been proposed. The control method which has a feedforward structure of the web tension and the upstream control input is designed from the dynamics of the sectional rotogravure printing press. First, a nonlinear coordinate transformation for the state variable of printing press dynamics is proposed, and in the new coordinate system, a nonlinear control method is designed using the Lyapunov stability. The performance of the new control method is then demonstrated by simulation and experimental results and compared with the conventional PD control method.

2 DYNAMICS OF ROTOGURAVURE PRINTING PRESS

In this section, we present the dynamics of the sectional drive rotogravure printing press and a coordinate transformation of the variable in order to be able to easily deal with the dynamics. Figure 4 illustrates a model of a sectional drive rotogravure printing press between ith and (i+1)th units.

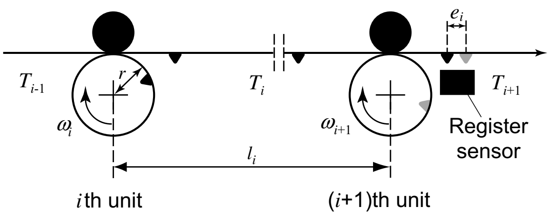


Figure 4: Schematic diagram of sectional drive rotogravure printing press

Black and gray triangular patterns are printed in the ith and (i+1)th units, respectively, for each revolution of the gravure cylinders. When both the color patterns are printed on the same position, the register error is defined as zero.

e_i and T_i(i = 1, 2, …, n) denote the register error and the web tension between the ith and (i+1)th units, respectively, and n denotes the number of units. ω_i denotes the angular velocity of the gravure cylinder composed of the variable $\Delta\omega_i$, which is the control input, and the constant ω^*, which corresponds to the line speed, as shown in equation (1).

$$\omega_i(t) = \omega^* + \Delta\omega_i(t) \tag{1}$$

r is the radius of the gravure cylinder, equal across all units, l_i is the web pass length between the ith and (i+1)th units, and K is the extension ratio of the web.

The dynamics of the register error and the tension between the units are represented as follows [1].

$$\begin{cases} \dfrac{dT_i(t)}{dt} = \dfrac{r}{Kl_i}\Big\{(1+KT_i(t))\big(\omega^* + \Delta\omega_{i+1}(t)\big) \\ \qquad\qquad - \dfrac{(1+KT_i(t))^2}{1+KT_{i-1}(t)}\big(\omega^* + \Delta\omega_i(t)\big)\Big\} \\ \dfrac{de_i(t)}{dt} = \dfrac{r\omega^*}{1+KT_i(t)} - \dfrac{r\omega^*}{1+KT_{i-1}(t-L_i)} \end{cases} \tag{2}$$

where L_i is the delay time defined in equation (3).

$$L_i = \frac{l_i}{r\omega^*} \tag{3}$$

For system (2), we introduce the following coordinate transformation.

$$\begin{bmatrix} T_i \\ e_i \end{bmatrix} \to \begin{bmatrix} \phi_i \\ e_i \end{bmatrix} \tag{4}$$

where ϕ_i is defined as

$$\phi_i(t) = \frac{1}{1+KT_i(t)}$$

Based on the new coordinate system, system (2) can be rewritten as the follows.

$$\begin{cases} \dfrac{d\phi_i(t)}{dt} = -a_i\left\{(\omega^* + \Delta\omega_{i+1})\phi_i(t) - (\omega^* + \Delta\omega_i)\phi_{i-1}(t)\right\} \\ \dfrac{de_i(t)}{dt} = -b\left\{\phi_i(t) - \phi_{i-1}(t-L_i)\right\} \\ a_i = \dfrac{r}{l_i} \quad b = r\omega^* \end{cases} \tag{5}$$

where ϕ_i, i.e., T_i and e_i are assumed to be measured. In comparison with system (2), the system (5) is almost linear. With the help of system (5), a nonlinear controller to make the register error converge to zero can be designed.

3 CONTROL DESIGN

To design the new control law, we introduce a new parameter $\varepsilon_i(t)$, as described in (6).

$$\varepsilon_i(t) = \lambda_i e_i(t) + \frac{de_i(t)}{dt} \tag{6}$$

where λ_i is a positive constant. Equation (6) can be rewritten as equation (7) using the second equation of (5).

$$\varepsilon_i(t) = \lambda_i e_i(t) - b\left\{\phi_i(t) - \phi_{i-1}(t-L_i)\right\} \tag{7}$$

The derivative of equation (7) can be given by equation (8) using equation (5).

$$\begin{aligned} \frac{d\varepsilon_i(t)}{dt} &= \lambda_i\frac{de_i(t)}{dt} - b\left\{\frac{d\phi_i(t)}{dt} - \frac{d\phi_{i-1}(t-L_i)}{dt}\right\} \\ &= a_{i-1}b\big(\omega^* + \Delta\omega_{i-1}(t-L_i)\big)\phi_{i-2}(t-L_i) \\ &\quad - b\left\{a_{i-1}\big(\omega^* + \Delta\omega_i(t-L_i)\big) - \lambda_i\right\}\phi_{i-1}(t-L_i) \\ &\quad - a_ib\big(\omega^* + \Delta\omega_i(t)\big)\phi_{i-1}(t) \\ &\quad - b\big(\lambda_i - a_i\omega^*\big)\phi_i(t) + a_ib\phi_i(t)\Delta\omega_{i+1}(t) \end{aligned} \tag{8}$$

When the control input of the (i+1)th unit $\Delta\omega_{i+1}$ is assumed as follows,

$$\begin{aligned} \Delta\omega_{i+1}(t) = -\frac{1}{a_ib\phi_i}\Big[&\big(1+\lambda_i\mu_i\big)e_i(t) \\ &+ a_{i-1}b\big(\omega^* + \Delta\omega_{i-1}(t-L_i)\big)\phi_{i-2}(t-L_i) \\ &+ b\left\{\lambda_i + \mu_i - a_{i-1}\big(\omega^* + \Delta\omega_i(t-L_i)\big)\right\}\phi_{i-1}(t-L_i) \\ &- a_ib\big(\omega^* + \Delta\omega_i(t)\big)\phi_{i-1}(t) - b\big(\lambda_i + \mu_i - a_i\omega^*\big)\phi_i(t)\Big] \end{aligned} \tag{9}$$

the equation (8) can be expressed as follows.

$$\frac{d\varepsilon_i(t)}{dt} = -\mu_i\varepsilon_i(t) - e_i(t) \tag{10}$$

where μ_i is a positive constant. The closed loop system can then be described as equation (11) using equation (6) and (10).

$$\begin{cases} \dfrac{de_i(t)}{dt} = -\lambda_i e_i(t) + \varepsilon_i(t) \\ \dfrac{d\varepsilon_i(t)}{dt} = -\mu_i \varepsilon_i(t) - e_i(t) \end{cases} \tag{11}$$

With regards to the convergence of system (11), the following conclusion can be derived using the Lyapunov stability theorem.

Proposition:

The closed system (11) is asymptotical stable at the origin $e_i=0,\ \phi_i=0$.

Proof:

The Lyapunov function is written as equation (12).

$$V = \frac{1}{2}\sum_{i=1}^{n} e_i^{\,2} + \frac{1}{2}\sum_{i=1}^{n} \varepsilon_i^{\,2} \tag{12}$$

The derivative of V along the trajectory of the system is given by

$$\begin{aligned} \frac{dV}{dt} &= \sum_{i=1}^{n} e_i \cdot \frac{de_i}{dt} + \sum_{i=1}^{n} \varepsilon_i \cdot \frac{d\varepsilon_i}{dt} \\ &= -\sum_{i=1}^{n} \lambda_i e_i^{\,2} - \sum_{i=1}^{n} \mu_i \varepsilon_i^{\,2} \end{aligned} \tag{13}$$

Clearly, dV/dt is negative definite over $e_i \neq 0,\ \phi_i \neq 0$. Therefore we can conclude that system (11) is asymptotical stable at the origin $e_i=0,\ \phi_i=0$ by the Lyapunov stability theorem.

From the proposition, it can be concluded that all $e_i(t)$ and $\varepsilon_i(t)$ converge to the origin, irrespective of the initial state. Then, based on equation (7), ϕ_i converges to the upstream value $\phi_{i-1}(t-L)$. Eventually, ϕ_i converges to the most upstream value $\phi_0(t)$ as $t \to \infty$. Based on the definition of ϕ_i, this implies that T_i converges to T_0 that is maintained constant.

4 SIMULATION

In order to evaluate the performance of the proposed control method, we conducted a simulation wherein the proposed control law was compared with the conventional PD control one. Table 1 shows physical parameters of the printing press.

Table 1: Physical parameters of printing press.

Number of Units	3
Line speed: V	100 [m/min]
Web pass length between ith and (i+1)th unit: l_i	7.46 [m]
Radius of gravure cylinder: r	0.103 [m]
Entry tension to 1st unit: T_0	100 [N]
Web extension ratio: K	2.3×10^{-4}[1/N]
Angular velocity of 1st unit gravure cylinder: ω^*	16.1 [rad/s]

Moreover, we considered a parameter ignored in the mathematical model. A rapid change in the angular velocity of the gravure cylinder caused web breakages or printing defects due to a slippage between the web and the gravure cylinder. Therefore, the controlling time per revolution of the gravure cylinder was limited to 4ms.

The PD and the proposed control gains were adjusted as shown in Table 2 to yield the best possible performance.

Table 2: Control gains of PD and proposed method

Proportional gain : K_p	19.2
Differential gain : K_d	249.6
Nonlinear control gain: λ_i	0.5
Nonlinear control gain: μ_i	100

Figure 5 demonstrates the calculated results of the PD and the proposed control method in the case of a register error of 0.5 mm introduced to the upstream register e_1 as a disturbance input. Compared to the PD control method, the proposed control method yields considerably better results as regards to the fluctuation in the downstream register error e_2. The register specification of the conventional printed products is within ±0.1 mm, and the proposed method maintains the e_2 within ±0.1 mm. Therefore this result indicates that the proposed control method is significantly effective in preventing the upstream register fluctuation from traveling toward the downstream register.

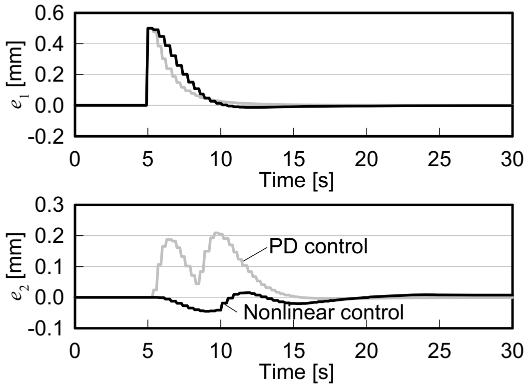


Figure 5: Calculated results of nonlinear and PD control method

5 EXPERIMENT

5.1 Experimental apparatus

To verify the effectiveness of the proposed control method, we conducted an experiment with a sectional drive printing press, as shown in Figure 6. This printing press was developed by Dai Nippon Printing Co., Ltd. and the main specifications of this system are as follows.

・The printing press consists of 7 units.

・The width of the gravure cylinder is 1100 mm.

・The maximum line speed is 350 m/min (5.83 m/s).

Figure 7 illustrates the experimental setup. The control input was calculated by using a digital signal processor (DSP). The register errors were measured with specialized optical sensors, and the analog measurement signal was converted to digital signals by using a commercially manufactured register controller (DT-2000, TAIYO Electric Industry Co., Ltd.). The DSP communicated with the register controller via a digital input board. The web tension between the units was measured with a tension sensor (MB25B, NIRECO Co., Ltd.). The tension sensor communicates with the DSP via an analog input board. The control signals calculated by the DSP were outputted via a digital output board.

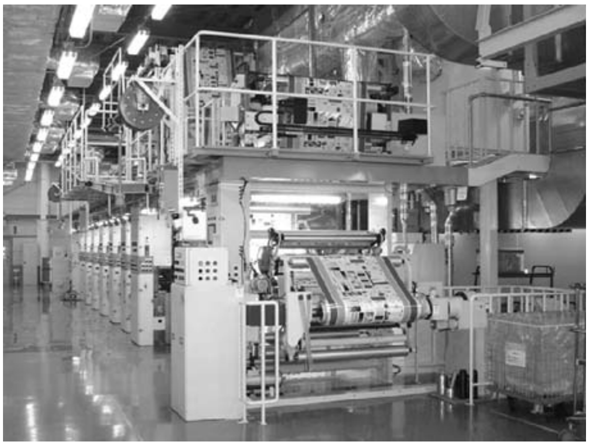


Figure 6: Experimental apparatus

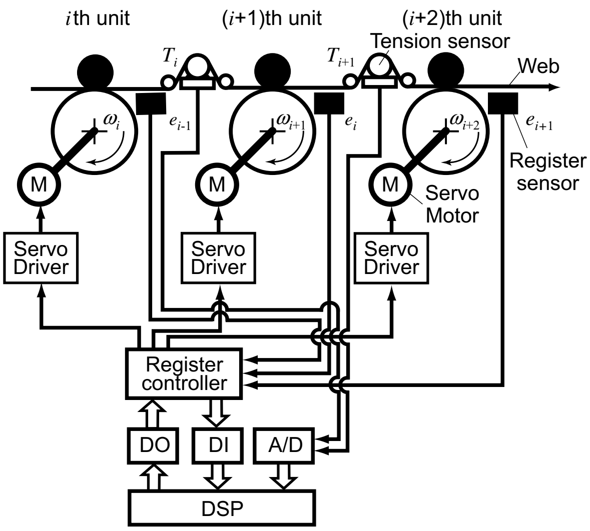


Figure 7: Experimental setup

5.2 Experimental result

Figure 8 illustrates the experimental results of the proposed and the conventional PD control method. Physical and control parameters are the same as what are used for the simulation as shown in table 1 and table 2. As well as the calculated results as shown in Figure 5, the fluctuations of the downstream register error e_2 of the proposed control law were constrained within ±0.1 mm. Therefore, the effectiveness of the proposed method could be confirmed with the actual printing press.

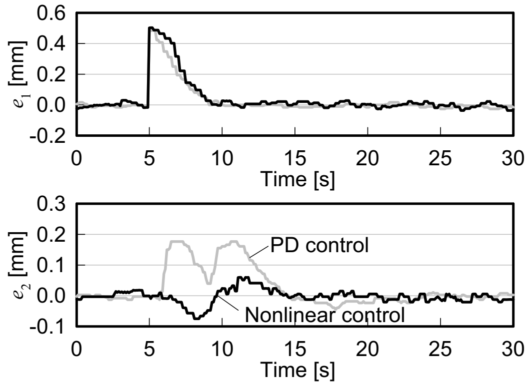


Figure 8: Experimental results of nonlinear and PD control method

6 CONCLUSION

This paper presents a nonlinear control method for the register of the sectional drive rotogravure printing press. The purpose of this control method is to prevent the generation of downstream register errors by controlling the upstream register. First, we proposed a nonlinear coordinate transformation for the state variables of the printing press dynamics. Then, based on the new coordinate system, a nonlinear control law was designed with the Lyapunov stability theorem such that the register errors converged to zero. Finally, the performance of the proposed method was demonstrated by simulations and experiments and compared with that of the conventional PD control method. As observed from the results, the proposed control method exhibited a considerably improved performance. Further, the proposed control method required not only the value of the register errors but also the tension between the units. However, sensors required for measuring the tension between units in a conventional rotogravure printing press are not available. Therefore, to apply the proposed control method to an actual printing press, an observer for estimating the tension value is required. Thus, our future work will focus on the observer design.

7 REFERENCES

[1] Yoshida, T., Takagi, S., Shen, T., Muto, Y., 2007, Improvement of Register Control in Gravure Printing Press, Preprint of the 118th Spring Conference of the Japanese Society of Printing Science and Technology, 11-16.

[2] Komatsu, H., Yoshida, T., Takagi, S., Shen, T., Muto, Y., 2007, Improvement of Printing Accuracy via Web Handling Control in Multi-colors Printing, International Conference on Control, Automation, and Systems 2007, 953-956.

[3] GURAVURE Process and Technology, 1991, Gravure association of America.

[4] Krstic, M., Kanellakopoulos, I., Kokotovic, P., 1995, Nonlinear and Adaptive Control Design, John Wily & Sons.

Incremental Sheet Forming with a Robot System for an Industrial Application

Arnold Puzik[1]

[1] Fraunhofer Institute for Manufacturing Engineering and Automation (IPA), Stuttgart, Germany

Abstract

This paper introduces new results of the single point incremental forming process of an industrial sheet metal component. Incremental sheet forming (ISF) is an innovative and compared to conventional forming technologies which does require forming tools like a specific die and punch, a cost-effective forming method for prototyping and limited-lot production. Moving a hammering tool on a generated path above the sheet metal one can manufacture a three dimensional complex sheet metal product. Contrary to deep drawing process and due to the low forming forces since using a high frequent hammering tool a common industrial robot can be applied. A robot with handling payloads of 400 kg was sufficient. Thus this ISF-process enables production of sheet metal parts and preliminary models at low costs.

Keywords:

Incremental Sheet Forming; Single Point Incremental Forming; Rapid Manufacturing

1 INTRODUCTION

The acceleration of the development and the entry of new products to the market, which mirrors the wishes of customers get a decisive key property to meet the international pressure of competition. In order to deal with the industrial competition one needs to increase the number of developments and therefore to reduce time and costs of manufacturing prototypes. Both the increasing number and speed of developing innovative products compared to the decreasing time of merchandise are reasons for an increasing dynamic behaviour of innovation [1]. Today there are decisive requirements for the technique of sheet metal forming. Not only the economically production of small lot sizes but also manufacturing of flexible and functional forms of high quality are necessary and a great challenge, [2] and [3]. The trend in automobile production leads to more complex part geometries and requires a faster introduction of new products by using real sheet metal prototypes. Furthermore the demand of more individual products produceable only in small lot sizes are increasing [4].

To try to increase the flexibility and to meet the challenge of today manufacturing prototypes with flexible form geometries tests were done, [5] and [6]. Nevertheless the geometry and thus the flexibility is restricted depending on the geometry of the tool. On the one hand the construction of a robust forming machine with a high forming capacity is very cost-intensive, but on the other hand the multiple use of the tool reduces the total costs only slightly [7].

In recent years new forming processes have been developed to help increasing the flexibility of conventional sheet metal forming methods. By repeating local deformation the desired geometries can be produced. As the local forming occurs step by step, the hammer forming process is already known as incremental forming process. For the incremental forming one need only a small plastic deformation area and thus the forming tool is smaller compared to conventional forming process like deep-drawing. A higher deformation degree, a possible deformation of materials with reduced ability of deformation and a higher strength can also achieved by applying the incremental forming [8]. When looking on other incremental forming processes which need a die plate for complex geometries conventional deep drawing is favoured because of less of process time and less of costs for the technical equipment of available incremental forming processes for producing prototypes and limited-lot production. In spite of remaining problems of flexibility and industrial complex requirements the aim is to establish the incremental forming process with all its advantages in industry. Thus a new incremental forming process was developed at the Fraunhofer-Institute for Manufacturing Engineering and Automation [9] to [13]. There are two main aspects to meet time and cost factors. First this process works without any die plates and second the investment costs for the mechanical equipment are low since an industrial robot is applied and able to handle the forming tool.

2 PROCESS

The process of ISF by using an industrial robot system has the following working process. The deformation of the sheet metal is caused by a hammer tool. Contrary to conventional deep drawing the dimensions of the hammer tool of ISF with a robot system is small compared to the blank sheet metal. The hammer tool itself is installed on the last axis of the robot by using a suitable flange. The high-frequency executed hammer punches occurs the incremental deformation of the sheet metal. While moving the robot system with the interacting hammering tool above the designated shape of the plate can be formed. The steps necessary for the entire incremental hammering process is shown in Figure 1. The principle of the new process is shown in Figure 2.

For the forming process the sheet metal is clamped in a frame. The position of the frame is calibrated and given to the numeric control of the robot. The designated geometry of the part is formed through the programmed path of the robot on the xy-plane. After each circle of forming the half-rounded hammer tool is positioned by Δz.

The development of the hammering tool itself has been worked out at Fraunhofer IPA. Since a high number of punches per second had shown an influence on the forming result tools using the spring energy were not applied any longer. Therefore a hammering tool consisting of an excentric connecting rod which excites the vertical movement of the tool was designed. The mass balanced excenter on the crankshaft generates an amplitude of 2 mm. The last generation of the hammering tool has a frequency of 200 Hz and is shown is Figure 3.

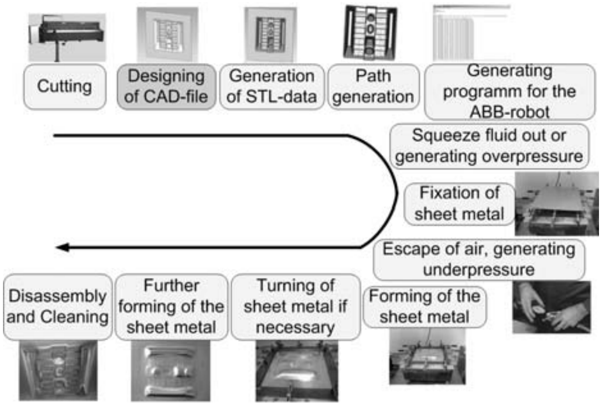


Figure 1: Current single steps needed for ISF by Hammering

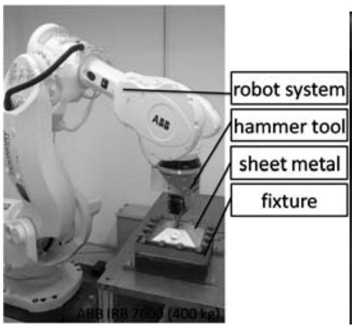


Figure 2: Operating setup Figure 3: Current forming tool

3 GENERAL ASPECTS AND PROBLEMS IN ISF

General aspects of incremental forming by hammering are affected through tool size, wall angles and the shape of fixture. The smaller the tool size the inferior the formability of sheet metal will get. This means that perforation of the sheet metal could occur. The used tool diameter for common forming experiments is 10 mm. That results in a minimum edge radius of 5 mm for formed parts. Designated wall angles are limited by material and thickness. Depending on the named influences and the forming strategies one can reach wall angles up to 70 degrees for stainless steel. Further practical tests have shown maximum wall angles of 75 degrees for deep drawing steel and up to 85 degrees for geometry features of depth smaller than 10 mm, [14]. The plastic deformation area of hammering is very small and limited to the contact area between hammering tool and sheet metal. So deformation is close to the plane strain condition, studied in [6] for incremental forming processes. Thus the sine law determines the steepest wall angle in terms of the material thickness:

$$s_1 = s_0 \cdot (1 - \sin\alpha) \quad (1)$$

where s_1 is the thickness of the formed shape, s_0 the original material thickness and α the designated wall angle. This was approved in [15] for incremental forming by hammering, but also generally for ISF methods, as shown in [16] and [17].

3.1 Material, Thickness, Speed, Step Size and Surface

Most influences on the forming process result from material properties, thickness, speed and step size. Tryouts have shown a possible formability for different kinds of steels, aluminium, brass, copper and titanium. Most common materials are shwon in Table I. The maximum shapeable thickness depends on material and forming capacity. In previous experimental studies a suitable rate of feed and step size was found. A suitable speed had to be found experimentally in order to stabilize the process. If the rate of feed is too high the process would get instable and a bumping tool occurs. For incremental hammering forming a speed in the range of 40 mm/min was found as suitable rate of feed for most materials and thicknesses. This number, the complexity of geometry and the step size are directly determining for production time.

Table 1: Most shapeable materials and thickness

Material	Possible Thickness [mm]
Stainless steel	0,5…2
Deep drawing steel	0,5…3
Aluminium	0,5…5

3.2 Lubrication

The lubrication in incremental forming cases is needed to reduce friction between tool and formed sheet metal in the forming area. A further reason is to transfer the heat away from the point of origin. Thus lubrication increases surface quality and forming limits.

3.3 Interaction of Process Parameters

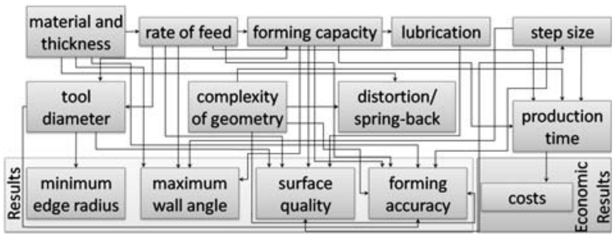


Figure 4: Process parameters and their interaction on each other with the resulting effects

All interacting factors are shown in Figure 4. All of the named factors affect each other in a strong way. As shown in this Figure the process parameters are listed in the upper part and the effects having an influence on the results are named below.

4 THE TOOL PATH GENERATION

When starting the generation of the tool path for the industrial robot one need to know the required forming steps to obtain the formed part as accurate as possible. Tool path generation in incremental forming technologies is as complicated as the 3D-information of the desired shape, which has to be expressed by CAD-modell or mathematical function. Principally there are two ways of generating the tool path with the help of the Tool Center Point (TCP). The first case defines the TCP in the middle of the spherical tool tip. In the second case the TCP could be defined by the point where the z-axis of the forming tool interferes with the surface of the desired formed part. In this study the first way was followed for tool path generation, since the vertical direction of the hammering tool occurs the main forming forces. This direction equals the z-axis of the workpiece.

Defining the TCP in the middle of the spherical tool tip enables two possibilities of calculating the tool track of the TCP. One possibility is defined by a horizontal plane of the contact points between the spherical tool tip and the desired surface and the other shows the track of TCP situated in a horizontal plane. In the first way the hammering tool could move in the vertical direction whereas in the second case the hammering tool remains on the same z-level during one circular rotation. To calculate the track of TCP the second way of horizontal tracks of TCP during one rotation is applied since the forming process in incremental hammering forming works in horizontal levels.

Equal to the conventional milling process with a spherical cutter standard CAD- and CAM-software is used for the tool path generation. The software automatically considers the tool dimensions like diameter and curvature of radius.. First of all a dedicated CAD-model of the designated shape has to be designed. During the designing process one has to obtain the tool dimensions. It is obvious that forming geometries smaller than the tool dimensions are impossible to form accurately. Thus forming geometries should be adjusted or not to be smaller than tooling dimensions. By opening the CAD-file with the CAM-modul NC-data can be generated. With the setting of constant z-step a suitable NC-program can be created. This data has to be loaded with a specially post-processor generates the program for the ABB robot.

5 INDUSTRIAL TEST WORKPIECES

Since research on ISF has been done tasks like material behavior and parameters of elongation, sheet thinning, forming limits and design rules, as shown in [14] and [18] of simple basic parts were investigated and discussed. Knowing the results and impacts of the properties mentioned in the section before enables to manufacture an industrial part as accurately as possible with the developed hammeringsystem of today. In order to show the possibility to form industrial parts several steps of Figure 1 were necessary to obtain the final and designated shape. Figure 5 shows the result of the formed part, which normally would be formed by conventional deep drawing.

A main task of this part was the accuracy. The depth of this part was 19 mm with great forming angles of approximately 60 degree. The z-step for both forming steps was 0.3 mm. During the forming process of the first side a noteworthy distortion and spring-back effect of the sheet metal occured. Thus both these effects had a great influence on positioning and clamping of the formed sheet metal and on z-positioning of the hammering tool by an offset value. After turning around the sheet metal and forming the second step both effects could be slightly removed and almost straighten the sheet metal for further processing. But still some tasks need to be optimized remained. One aspect results from the named effects during the first forming step, which was the 1.5 mm depth of the small and tight laying forms next to the long grooves. Both of them were not formed deep enough. Another aspect occurs on the walls of the grooves. Here one can see slight distortion of the sheet metal. This could occur because of an inaccurate clamping fixture.

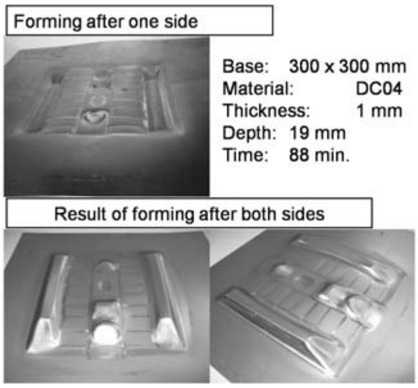


Figure 5: Two step forming of an industrial part

A solution to remove these effects could be one of the following factors. A great aspect of sheet metal forming is the surface roughness. A possibility for improvement could be to reduce the z-step to a value of 0.1 to 0.15 mm. So further experiments of z-step-optimization need to be done. For an accurate clamping and turning around of the sheet metal one need to have a fixture, which can turn the fixture without moving the sheet metal itself and to adjust the fixture to an mechanical bedstop. Such a fixture enables flexible turning depending on the forming strategy. A further solution for slight formings as seen next to the long grooves could be a greater value for those depth. Thus a changing of depth from 1.5 mm to 3 mm in the drawing-file could result in the designated shape. Therefore one additional experiment was done with a self created CAD-file, as shown in Figure 6.

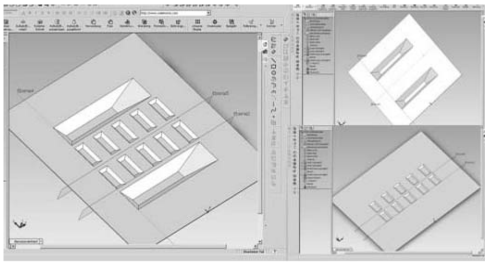


Figure 6: Adopted drawing of the original part

As shown in Figure 6 the main tasks were divided into two forming sections. As discribed before forming of the long grooves occured distortion of the sheet metal, which could be removed after forming of the second side a further experiment in order to improve forming needs to be done. Therefore the idea was to form the long grooves from both sides first and then to form the slight formings between those long grooves, as seen on the lower drawing of the right side of Figure 6. That means the result of the forming of the long grooves would not affect the forming of the slight formings. The results after the formed steps are shown in Figure 7. On the upper left side one can see the forming results after one forming step and after three forming steps on the upper right side. The final result after four forming steps is shown both on the lower left and right side. Compared to the first results in Figure 5 one can see improvements in accuracy in this process. In spite of distortion and spring-back effect of the sheet metal the depth could be reached and accuracy increased.

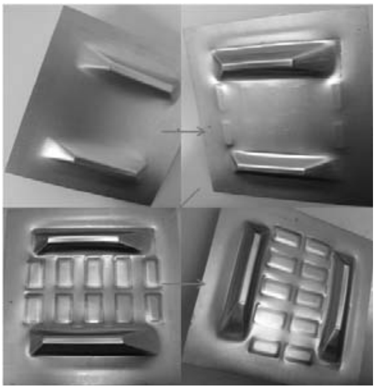


Figure 7: Single different steps of the adopted part

6 CONCLUSION AND PERSPECTIVES

This article describes first the different steps necessary to be able to form by incremental hammering forming with an industrial robot. One basic aim of research was the functionality and application of industrial manufactured parts. First of all the process has been developed. For the forming tool one figured out that forming with a hammering tool existing of one excentric punch and two mass balancing excenter rings has been shown the best results. The hammer tool has a frequency of 200 punches per second and generates an amplitude of 2 mm.

In order to enable this hammering manufacturing different aspects and problems in ISF have been figured out. That means to consider and notice the limitations of ISF with a robot. Specially it also means to keep in mind the interacting parameters and their effects on the forming results.

This developed incremental forming process makes it possible to form both convex and concave geometries. The test workpiece shows the forming possibilities and the challenge to meet the distortion and spring-back effect. Therefore the idea was to adapt the forming strategies and geometry dimensions to these problems. So the CAD was new designed by increasing the dimensions of small geometries in order to reach the original depth and also divided in four forming steps to remove distortion and to enable optimized forming in already distorted areas.

In future works this approach of dividing forming tasks of several geometries and the adaptation of geometry dimensions will be investigated to increase the accuracy of the forming process. Furthermore the combination of a database shall be found by examination of deformation of several materials and workpieces along with the proposed approach should be the best suitable method in order to compensate deformation, distortion and spring-back as much as possible.

7 REFERNCES

[1] N.N.: Problematik der heutigen Produktentwicklung. www.sfb374.unistuttgart. de/darstellungsfb.htmlanker2, Comput. J..

[2] T. Nakagawa: Advances in Prototype and Low Volume Sheet Forming and Tooling. Journal of Materials Processing Technology 98 (2000), PP. 244-250,.

[3] P. Groche, R. Schneider: Umformtechnik für die Produkte von morgen. Mat. -wiss. u. Werkstofftech. 31 (2000) Nr. 11, S.958 -960.

[4] E. Harsch, C.-P. Neumann: Flexible Pressen fuer die Fertigung kleiner Lose. DFB-Kolloquium "Flexible Blechbearbeitung", 1992, Boeblingen, Bd. T11, S. 1-17.

[5] T. Hennige: Flexible Formgebung von Blechen durch Laserstrahlumformen. Bamberg: Meisenbach, 2001 Erlangen-Nuernberg, Univ., Diss., 2000.

[6] S. Junk: Inkrementelle Blechumformung mit CNC-Werkzeugmaschinen: Verfahrensgrenzen und Umformstrategien. Universität des Saarlandes, Schriftenreihe Produktionstechnik Band 25, Diss., Saarbruecken 2003.

[7] T. Schaefer, R. Schraft: Incremental Sheet Metal Forming with industrial robots. Fraunhofer Institute Manufacturing Engineering and Automation, Stuttgart, 2004.

[8] N.N: Modellierung inkrementeller Umformverfahren. http://www.miu.uni-saarland.de/docs/downloads/begruendung.pdf, 2002.

[9] E. Westkaemper, R.D. Schraft, W. Schaaf, T. Schaefer: Roboshaping - Flexible inkrementelle Blechumformung mit Industrierobotern. wt Werkstatttechnik online Heft 1, 2003.

[10] T. Schaefer: Experimentelle Untersuchung eines Verfahrens zum inkrementellen Biegeumformen von Blech mit Industrieroboter. UTF science (2003), Nr. 3, S 1-5.

[11] T. Schaefer: Inkrementelle Umformung von Feinblech mit Industrierobotern. VDI/VDE-Gesellschaft: VDI-Berichte 1841, Robotik: Leistungstand, Anwendungen, Visionen, Trends, München, 2004.

[12] E. Westkaemper, R.D. Schraft, W. Schaaf, T. Schaefer: Roboshaping. Flexible inkrementelle Blechumformung mit Industrierobotern, In: wt Werstatttechnik 93 (2003), Nr. 1/2, S. 7-11.

[13] R. D. Schraft, T. Schaefer: Inkrementelles Biegeumformen von Blech mit Industrierobotern: Experimentelle Untersuchung. wt Werstatttechnik 93 (2003), Nr. 6, S 489-493, 2003.

[14] L. Lamminen, J. Tuomi: Creating Design Rules for Incremental Sheet Forming. 18th International Conference on Production Research (ICPR-18), Salerno, Italy, 31.7 -4.8.2005, 2005.

[15] T. Schaefer: Verfahren zur haemmernden Blechumformung mit Industrieroboter. Institut für Industrielle Fertigung und Fabrikbetrieb, Universitaet Stuttgart, Diss., 2007.

[16] L. Lamminen: Incremental Sheet Forming with an Industrial Robot - forming limits and their effect on component design. SheMet 2005, 5.-8.4.2005, Erlangen, Germany.

[17] G. Hirt, S. Junk, N. Witulski: Incremental Sheet Forming: Quality Evaluation and Process Simulation. proceedings from 7th ICTP, Vol. 1 (2002), pp. 925-930.

[18] L. Lamminen, T. Tuominen, S. Kivivuori: Incremental Sheet Forming with an Industrial Robot. Proceedings of the 3rd International Conference on Advanced Materials Processing, Materials Forum, Vol. 29, 331-335, 2005.

Multi-Technology Platform for Hybrid Metal Processing

Christian Brecher[1,2], Frank Hoffmann[1], Andreas Karlberger[1], Chris-Jörg Rosen[2]
[1] Laboratory for Machine Tools (WZL), Chair of Machine Tools, RWTH Aachen University, Germany
[2] Fraunhofer Institute for Production Technology (IPT), RWTH Aachen University, Germany

Abstract
This paper presents research concepts that concentrate on the functional hybridization of a milling machine. The main process milling is supplemented with several laser processes like laser deposition welding, laser hardening and laser drilling. Some laser processes and the workpiece handling are realized with a robot that is directly integrated into the machine structure and connecetd to the machine control, so that an integrated hybrid machining cell is built up. By flexible combination possibilities, a high average utilization of the modules machine, robot and laser is achieved. Thereby, the profitability of the production system is increased significantly.

Keywords:
Laser Integration, Robot Integration, Milling Machines

1 INTRODUCTION

Modern production systems have to guarantee high flexibility combined with high productivity and high performance in combination with high precision. Especially in high-wage countries the design of a machining system with respect to the machining tasks is essential to achieve a competitive production. Within global competition high-wage-countries are challenged by low-cost countries that generate their main advantage by minimal wage costs while closing the technology gap [1]. Therefore, the high-wage countries have to adapt their production technology in order to benefit from the globalization and other megatrends like the individualization of products [2].

The competition between high- and low-wage countries is dominated by two main dimensions. The first dimension is built up by the contradiction between the economies of scale and the economies of scope [3]. Mass production and batch size one are the extremes of this facet. The concept of the individualized production aims for production systems that can handle individual, unpredictable customer requirements. Thereby, small batch sizes of individualized products shall be manufactured at a price comparable to mass production. Furthermore, the contradiction between value oriented and planning oriented processes spans the second dimension. Virtual production methods intend to minimize the efforts for the planning and startup of production systems, e.g by extending CAD-CAM technologies [4]. Only progress in both dimensions by a loosening of the contradictions is sufficiently effective to strengthen the competitive advantages of high-wage countries [5].

In high-wage countries, the available high qualification and experience of workers can be used effectively for competitive advantages by autonomous adaptation and optimization. For this reason, technologies for the mass production can be integrated into more complex, hybrid machining systems with advanced flexibility and productivity.

2 HYBRIDIZATION OF MACHINE TOOLS

The hybridization of machine tools describes the parallel or serial usage of multiple process technologies in a single machine tool. If two different physical mechanisms are used in parallel the process technology itself is hybridized. For example, a combination of laser processing and turning can be used for a high productive machining of ceramics [6]. The application of such laser assisted machining is so far restricted to non-metallic materials. However, the serial hybridization or function integration of machine tools is an ongoing trend of the machine technology. The combination of turning and milling applications is state of the art. New concepts also implement a combination of turning and grinding, e.g. for the repair of turbine components. Furthermore, a combination of laser structuring and honing is realized on a specialized machine tool [7]. This process combination is used to minimize the friction in piston tracks. More general combinations of cutting and laser applications have been investigated in recent research projects. For example, a project in cooperation with industry partners focused on machining of rotation-symmetric workpieces before and after heat treatment [8]. The main turning process is supplemented with laser hardening and laser coating.

The main benefit of the process integration is that the machining can be done in one single clamping setup. Thereby the cycle time is reduced because the logistic efforts for transfers between several machine tools or external partners, clamping and calibration times are eliminated [8]. These advantages are especially valuable if they are applied to small lot sizes and products with many variants. In theory, a one-piece-flow, an individualized production without buffers, can be realized. At the same time the accuracy of the whole production process is increased. However, several challenges still limit the productivity of complex, hybrid machining systems. These limitations will be investigated in the scope of the presented work on the basis of an enhanced machining platform for multiple production technologies.

3 HYBRID PLATFORM CONCEPT

3.1 Main concept and mechanical setup

The multi-technology platform that has been developed is a hybrid machining cell for the machining of metal parts. Its concept is based on the idea of a modular system for a maximum of flexibility and cost effectiveness. The hybrid platform integrates the modules of a machine tool, a jointed-arm robot and two laser sources in order to realize a maximum number of cutting and laser applications in a single clamping position (Figure 1).

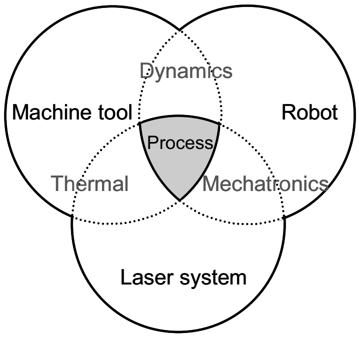


Figure 1: Hybrid platform concept.

By the modularization a high flexibility of the system is provided since existing standardized modules can be used with minimal changes. Therefore, the machining platform presented here is an exemplary realization of the hybrid concept which aims for a maximum flexibility in research applications. Another specific variety of workpieces will require an individualized machining platform with adapted flexibility. Since small lot sizes of individualized machining platforms are expected, the design of the robot and laser integration has to be compatible to existing construction kits of machine tools.

The main challenge of hybrid machining concepts is the cost effectiveness. The combination of the modules has to be more than the sum of the single parts. In the case of combined cutting and laser applications, the different requirements of the machining processes have to be accounted for. Most of the laser processes do not need the high kinematic accuracy and the stiffness of a cutting machine tool. Actually, a lot of laser applications are realized on jointed-arm robots. For this reason laser and robot integration require each other. In combination with the machine tool also high accuracy laser processes become possible. The combination of robot and machine tool not only allows serial processing but also a parallelization of cutting and laser processing so that a high average utilization of machine spindle and laser source is achieved.

Based on this concept, the selection of modules and the definition of efficient mechanic and electronic interfaces have been elaborated. The mechanical setup that is built on a 3-axis milling tool with long bed and a 20000 rpm main spindle is presented in Figure 2. The fixed standard workpiece table is exchanged for two NC swivel rotary tables. Thus two workspaces for 5-axis machining are generated. One workspace can be used for milling, drilling and turning while the other one can be used for the workpiece exchange or for robot processing. A 6-DOF jointed-arm robot is integrated into the middle of the machine bed and can interact with the swivel rotary tables to enhance the workpiece accessibility. Machine and robot are able to use both workspaces independently. This way, a machining in a single setting with multiple process technologies becomes possible.

Concerning the laser integration, the hybrid platform is featuring a 2.4 kW fiber laser with very high beam quality. Between the swivel rotary tables a magazine for the storage of several laser processing units is located. Thus, the magazine doors separate the milling workspace from the laser workspace to avoid errors in the laser process due to coolant lubricant and chips. Additionally, the inlet for the energy and media chains is located in the ceiling of the magazine. These chains have to be permanently attached to the processing units which can be picked up by the machine or the robot with a HSK-interface. Because of the fragile fiber for the laser transmission a complex chain guiding and tension unit is installed on top of the machine tool.

Another very important aspect of the laser integration is the concept which has to ensure the safety for the operator and all other staff nearby. Laser safety is highly dependent on the laser source and the laser power. Therefore a complete redesign of the machine tool's housing is not cost effective. In the case of modern high power fiber lasers, passive safety housings get very thick and expensive. However, active systems can be used to detect laser malfunctions and directly shut off the laser source [9]. The hybrid platform has got an additional laser safety housing with active walls and ceiling around the machine tool and the laser source. During laser operations an outside operator panel with video surveillance of the workspaces is available.

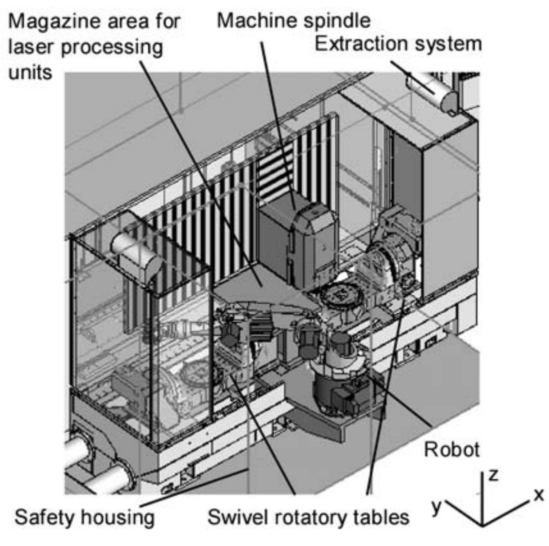


Figure 2: Multi-technology platform.

3.2 Laser processing units

This part of the project concerns the development, qualification and optimization of modular and highly integrated laser processing units and supporting components. All laser system components are designed with respect to maximum processing flexibility at minimal size and high system robustness. The proposed concept consists of two processing units for wire-based laser deposition welding including laser hardening and laser ablation (Figure 3) as well as a short pulse processing unit for laser structuring.

Aiming at the improvement of existing solutions such as [8] basic investigations will focus on a system concept for laser deposition welding with double hot-wire technology and fast 2-D high power scanning with respect to maximum process flexibility. Beside the deposition of different materials by two-wire feeding, the beam deflection, integrated in the beam guiding system, will allow the selective manipulation of the resulting coating geometry by influencing the surface tension and the local temperature distribution in the molten wire material. Thus the resulting seam width can be varied within specific ranges.

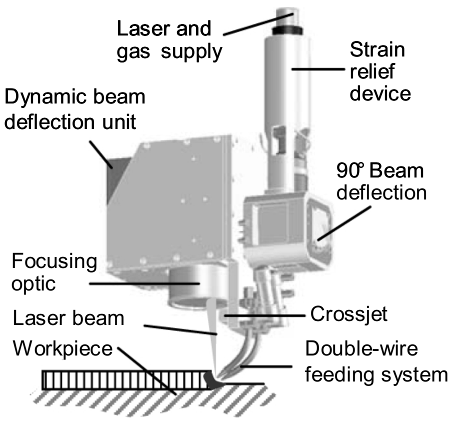


Figure 3: Concept of a hybrid processing unit for laser deposition welding, laser hardening and laser ablation.

A rapid beam deflection in the multi kW range will create further benefit. Operating without wire feeding, the oscillating laser spot will be used for the flexible generation of specific heat distribution on the workpiece surface during laser hardening. Thereby the oscillating laser spot creates a nearly uniform and highly adjustable temperature field which can be adapted to different workpiece geometries. In addition to the scanning functionality, the integrated deflection system can also be used for high-precision alignment between focal spot and wire tips by an appropriate deflection offset. Thus, a significantly simplified process set-up and reduced set-up times can be achieved. In combination with the high beam quality of the fiber laser source, the design of the optical system will be customized to enable ablation processes with a continuous wave laser system by controlled pulsing with pulse frequencies longer than 0.2 ms.

The increased system flexibility requires space optimized handling of the processing units. Therefore the integration concept features vertical energy chain supply including flexible laser beam guiding via optical fiber and optimized laser system components like crossjet for protection of optics against foreign particles, media supply with coolant and shielding gas, pyrometer for temperature measurement and adaptive beam power control as well as double hot-wire feeding system for the deposition welding process.

With the described system the potential for combining single functionalities to a hybrid laser processing unit will be investigated in order to evaluate hybridization compatibility of different laser systems and processes. Based on existing discrete system solutions for wire based laser deposition welding and laser hardening the new integration concept aims at the improvement of process flexibility in both deposition welding and hardening. Thereby it will be analyzed to what extend this hybrid laser processing unit can further broaden the scope of processible workpiece geometries by generating flexible deposition seams, hardening with specifically adjustable temperature fields, achievable surface hardness and depth, defined material ablation and thus the application of multiple and novel process chains.

3.3 Integrated machine and robot control

The proposed machine concept is totally based on standard components which can be integrated by existing software and hardware interfaces. One major goal is to realize an automation platform which allows the deterministic interaction of heterogeneous automation components including motion synchronization and real time data exchange with a minimum of engineering costs. In this application the considered components are numerical control (NC), robot control (RC) and laser control (LC).

One main challenge is the synchronization between the motion of the robot and the machine tool, e.g. for laser hardening or controlled metal build up. Depending on the geometry of the parts the accessibility for the robot can be limited. Another restriction on the mobility is given by the media supply of the modular laser heads. These problems can be addressed by the application of an external rotary axis. In already existing stand-alone applications of robot guided laser processes external axes are interpolated by the RC so that a precise path movement can be achieved. For the proposed machine concept this method of a centralized motion control is not suitable. High precision 5-axis milling must be executable on both working positions. Therefore, it is essential that the 5-axis-interpolation is performed entirely by the NC. In order to combine the movement of the NC-controlled swivel rotary table with the RC-controlled robot, a synchronization strategy will be developed. Since either NC or RC analyze several reference points of a programmed trajectory beforehand in order to achieve a targeted optimum between path accuracy and feederate, an online providing of desired position values by one of the controllers does not tap the full potential of currently available motion controllers. A master-slave architecture therefore does not match the requirements of the proposed hybrid machine center.

Only a decentralized architecture which allows the online synchronization of trajectories, being planned and generated offline, leads to a maximum of performance in this case.

Although hybrid machining centers offer a much higher flexibility they are more complex than conventional machines. An economic operation supposes a high use of capacity and a fast ramp-up on the one hand in combination with low costs for planning and scheduling on the other hand. Therefore, an integrated planning chain will be developed, ranging from the definition of the part as input to the task execution on the machine, including program verification and sequence planning (Figure 4). The unit for program generation is based on standard CAM-solutions, which are extended by modules for the graphical representation of laser processes as well as for the planning of process parameters. This approach leads to a maximum use of already existing path planning strategies for 5-axis machining and therefore cuts the initial costs. The generated nominal path of the laser motion relative to the surface of the workpiece has to be split into two separate programs either for the NC or the RC. For this purpose a postprocessor, able to transform the output data of the CAM system into a RC-specific format, will be used.

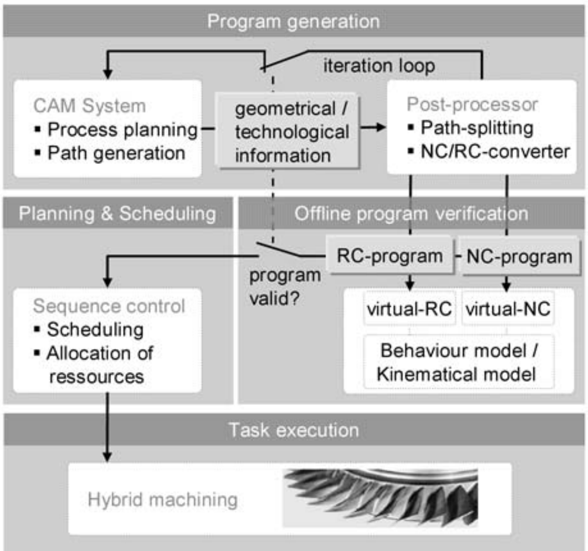


Figure 4: integrated planning and verification

Assuming that the orientation of the robot guided laser head remains constant during the laser process a 5-axis-machining strategy can be emulated, solving the problem of axis redundancies and restrictions on the mobility of the laser heads. Before a machining program is released, an integrated verification procedure based on virtual controllers in conjunction with a behaviour model and a kinematical model is applied. Collisions or violations of the working space can thus be detected beforehand. The exact machining time will be determined and passed to a planning and scheduling unit. This unit aims at fulfilling the targeted use of capacity by an optimized distribution of parallel processes with respect to the specified adherence on delivery dates.

3.4 Hybrid process technology

For the processing of an as wide as possible range of workpieces an intelligent combination of mechanical processes with laser technologies is necessary. So far only product specific process chains have been integrated into machine tools. The new concept will allow determining empirically the mechanical, thermal and chemical interactions between processes in detail. Moreover, the interactions of the manufacturing system and the single process step are important limiting factors for the productivity and flexibility of the one-piece-flow through the machining system. The tight temporal and spatial schedule of the process steps can emphasize these effects.

For the design specification of the multi-technology platform several hybrid process chains are used. One of these scenarios is the high productive manufacturing of molds and dies with contour cooling by Controlled Metal Build Up (CMB) [10]. CMB is a rapid manufacturing technology with a material deposition in layers. It combines welding for the material deposition and milling for the precise shaping of the layers. CMB can be used to achieve very high aspect ratios. On the multi-technology platform the contour of a layer is prepared with high precision by the machine tool. However, the main deposition welding inside the contour can be performed efficiently by the robot and a swivel rotary table. Afterwards the contour and the surface of the layer are machined by high speed milling.

4 SUMMARY

This paper presented concepts for the integration of laser processing and milling in a hybrid machining platform. Based on the analysis of requirements for an individualized production technology in high-wage countries a modular concept for the efficient machining of small batch sizes is developed. The hybrid system is built by the integration of a jointed-arm robot and specialized laser sources. Afterwards the requirements for the modules and the mechanical and electronic interfaces are described.

The multi-technology platform will be used to gain empirical knowledge about mechanical and thermal interactions of the hybrid process chains and single system components. Thereby, adequate compensation strategies for machine, robot and laser control will be derived. Finally a requirement specification for hybrid manufacturing systems that allows the cost effective configuration of the modules for specific workpiece ranges will be defined.

5 ACKNOWLEDGMENTS

The authors would like to thank the German Research Foundation DFG for the support of the depicted research within the Cluster of Excellence “Integrative Production Technology for High-Wage Countries”.

6 REFERENCES

[1] Tseng, M. M., 2003, Industry Development Perspectives: Global Distribution of Work and Market. CIRP 53rd General Assembly, Montreal, Canada.

[2] Duray, R., 2002, Mass Customization Origins – Mass or Custom Manufacturing? International Journal of Operations and Production, 3/22: 314-328.

[3] Chandler, A. D., 2004, Scale and Scope. The Dynamics of Industrial Capitalization, Belknap Press of Harvard University Press, Cambridge (Mass.), p. 17ff.

[4] Yeung, C.H., Altintas, Y., Erkorkmaz, K., 2006, Virtual CNC system – Part I and II. International Journal of Machine Tools & Manufacture, 46: 1107-1138.

[5] Westkämper, E., 2007, Strategic Development of Factories under the Influence of Emergent Technologies, CIRP Annals, 56/1: 419-422.

[6] Klocke, F., Bausch, S., Bergs, T., 2001, Laser-assisted turning of silicon-nitride ceramics. 7th International Symposium Ceramic Materials and Components for Engines, Goslar, Germany.

[7] Abeln, T., Klink, U., Laseroberflächenstrukturierung – Verbesserung der tribologischen Eigenschaften, Proceedings Stuttgarter Lasertage, 2003, Stuttgart, Germany.

[8] Brecher, C., Emonts, M., 2007, Integrierte Produktion - Fertigungszelle für Dreh-, Fräs und Laserbearbeitung. VDI-Z, 2007/09.

[9] Goebel, K. R., 2006, Patent DE 102006026555A1.

[10] Fuchs, E., 2006, Controlled metal build up process for generation, modification and repair of molds and dies, Moldmaking Technology Magazine, 2006/4.

Super-Resolution Imaging for Automatic Visual Inspection Systems

Hiroyuki Nakai,[1] Yasuhiro Ueda[1] and Yoshihide Shigeyama[1]
[1]Production Technology Development Center, Production Technology Development Group,
Sharp Corporation, Japan

Abstract
The purpose of this research is to apply super-resolution processing to defect inspections in order to automate such inspections and increase their precision. Super-resolution processing, which is a method that creates a high-resolution image from multiple low-resolution images, could be an effective solution in environments where images of sufficient resolution cannot be obtained. However, reduction of noise in the generated images is necessary in order to conduct inspections reliably. Therefore, in this paper, we propose a noise reduction method for super-resolution processing, and then report on the defect inspection performance when this super-resolution processing is implemented.

Keywords:
Super-resolution imaging, Visual inspection system, Noise reduction

1 INTRODUCTION

In recent years, the automation of factory work has advanced rapidly even in the electronics industry. Two factors contribute to this. One is that labor shortages are expected in the future due to declining birth rates and the aging of the population in Japan. The other is that the increasing scale and precision of products have begun to make it difficult for people to handle the work.

In the electronic products field, along with the increased performance of production technology, manufacturing larger products with even higher precision has become possible. As the popularization of large-scale, high-precision products advances, the integration of semiconductors in the internal circuits of products is also increasing and electronic circuits of even higher density are coming into use.

For this reason, as these types of products have become even more advanced, it has become difficult to realize inspections that are both rapid and very precise using human visual inspection, the main method used for inspection in manufacturing in the past. For this reason, the development of automatic inspection equipment that will allow precise inspection of even larger subjects has become desirable.

Processing with automatic inspection equipment can be broadly classified as image-capture processing and defect detection processing. However, if sufficient resolution is not achieved in image-capture processing, defect detection processing will frequently fail to detect defects. For this reason, high-resolution images must be captured in image-capture processing.

One image-capturing method used to achieve high-resolution images is to increase the number of image-capturing cameras and have each camera capture only a small area, but this method is difficult to implement due to price and maintenance considerations.

Considering this, one method to increase image resolution for inspection equipment is pixel shift processing. In this method, multiple low-resolution images are captured while changing the position of the image-capturing elements relative to the subject and then these pixels are reallocated to create a high-resolution image. Another method is super-resolution processing, which considers the optical degradation function of the image-capturing system and extrapolates a high-resolution image from multiple captured images.[1]

The latter approaches include the ML approach to super-resolution processing that extrapolates a high-resolution image by comparing it with captured images,[2] back projection approaches that use back projection kernels when extrapolating high-resolution images,[3][4][5] MAP approaches that use preexisting image information to extrapolate high-resolution images,[6][7][8] and POCS approaches.[9][10] In addition, research has been conducted on the theoretical limits of super-resolution processing.[11][12][13]

Since super-resolution processing is image extrapolation processing, unlike pixel shift processing, it has the advantage that a high degree of control over the image-capture positions for the captured images is not necessary. On the other hand, repeat operations are necessary, so calculated costs are high and processing of large images is difficult. For these reasons, methods have been reported that reduce operation time, including one that decreases the number of repetitions[14] and one that reduces the PSF varieties that require the approximation of the pixel positions of captured images in discrete positions to also curtail operation costs.[15]

In the inspection equipment for these high-resolution image-processing methods, since the cameras are fixed, the position of image capture can be controlled with precision. Therefore, the relative image capturing position is already known, making registration processing unnecessary. On the other hand, high precision in obtaining image luminance information is desirable.

Moreover, in high-resolution image processing used for inspections, evaluation indices other than clarity and reconstruction, which are used in visual image evaluation, are necessary. In other words, the criterion used in inspection should be the ability to distinguish signals from the parts with defects and the good parts.

Furthermore, good performance is desirable in the detection of minute structural defects in the subjects of inspection, which are difficult to identify visually in inspections. For this reason, one characteristic is that the frequency of the processing subject becomes higher when compared with the natural image.

In this paper, we propose a method in which noise added to the captured image is simultaneously extrapolated and reduced when high-resolution images using super-resolution processing from multiple images are extrapolated. This method establishes a noise extrapolation item to express and extrapolate noise simultaneously with the high-resolution image extrapolation in a conventional ML approach. Furthermore, by removing this extrapolated noise from the captured image, the effect of noise in the generated high-resolution image is reduced.

Next, we prepare a simulation image that has more of the high-frequency components that are expected to be the subjects of inspection in automatic inspection equipment and discuss evaluation of the noise reduction effect in the super-resolution processing using that image.

2 SUPER-RESOLUTION PROCESSING IN AUTOMATIC INSPECTION EQUIPMENT

High-precision inspection is possible if the resolution of captured images used in automatic inspection equipment is high. Captured image resolution can be raised by increasing the number of cameras used for image capture, but keeping the number of cameras as low as possible is desirable considering equipment costs and maintenance.

In addition, captured image resolution can be raised in high-resolution processing by increasing the number of captured images, but the use of numerous images is difficult considering constraints on inspection time. For this reason, in automatic inspection equipment, it is necessary to create high-resolution images that can be used for inspection from small numbers of low-resolution captured images.

Moreover, the frequency characteristics of inspection subjects also differ greatly from natural image-capture subjects. This is because the target of automatic inspection equipment is minute structural defects that are difficult to detect visually. In addition, inspection subjects are often structures with minute patterns such as display equipment and electronic circuits.

In display equipment, every pixel forms a pattern. Furthermore, since color displays have picture elements that are divided into RGB, the inspection subjects have frequency components that are two and three times greater than the pattern. In comparison, in electronic circuits, since the configuration elements form patterns and those configuration elements have structures of connected minute circuits, they have frequency components that are several times greater than their patterns.

For this reason, when image capture is done with a single camera, frequently only a few image-capture elements are allocated for one pattern. So, when sufficient resolution

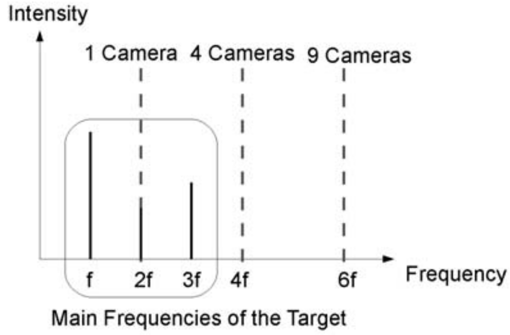


Figure 1: Relationship between the main frequency components of the inspection subject and the sampling frequency.

cannot be achieved with a single camera, an inspection processing method that employs multiple cameras is used to capture images of the subject in parts.

For example, for an inspection subject with a 2x2 [pixel/pattern] image-capture element allocation with a single camera, the relationship between the frequency components of the inspection subject and the sampling frequency when using multiple cameras in image capture is shown in Figure 1. Moreover, here, f indicates the pattern frequency of the inspection subject, and we assumed that the inspection subject includes 2–3 times higher frequency components in consideration of cases when the pattern is further segmented.

Avoiding interference fringes from moiré in the inspection image is also desirable, however, because they have a negative effect on inspection performance. For this reason, from the sampling theorem, image capture with a sampling frequency that is at least twice the frequency of the inspection subject is necessary. Therefore, in this case, the image capture frequency of a high-resolution image must be at least six times (6f) the sampling frequency of the pattern of the inspection subject.

In this case, imaging with sufficient sampling frequency is difficult even using a configuration with 9 cameras. For this reason, super-resolution processing becomes necessary. The generated high-resolution image has a sampling frequency that is proportional to the resolutions of the images. For example, an extrapolated high-resolution image that is twice the size of the captured images will also have a sampling frequency that is twice that of the captured images.

On the other hand, research on the theoretical limits of super-resolution processing reports that the possible extrapolation scale of high-resolution images is from 3x3 to 4x4 times the size of the captured image.[15]

So, in this paper, we evaluate super-resolution processing using this image-capture element allocation. For a subject of image capture with a pattern that has a largest main frequency component of 3f, we designed a simulation image with a sampling frequency of 4f. Furthermore, we conducted super-resolution processing on this image and applied a method to generate a 3x3 high-resolution image. As a result of this, the sampling frequency became 12f, and we achieved an image with a resolution that avoids the moiré effect.

3 SUPER-RESOLUTION PROCESSING WITH SIMULTANEOUS NOISE EXTRAPOLATION

3.1 Reconstruction based super-resolution processing

Reconstruction based super-resolution processing include ML and MAP. In reconstruction based super-resolution processing, an evaluation function is set with the sought pixel value for the high-resolution image as a variable, and a high-resolution image is reconstructed through iterative operations in the way that minimizes the evaluation function.

Figure 2 shows the processing loop for reconstruction based super-resolution processing. The evaluation function when using a MAP approach can be expressed with the following equation (Eq.1).

$$E = \left| C\mathbf{h} - \mathbf{f} \right|_2^2 + \alpha \left| B\mathbf{h} \right|_2^2 \quad (1)$$

The evaluation function is shown as the sum of first item on the right side, which is the error item, and the second item on the right side, which is the constraint item. The ML approach, however, extrapolates high-resolution images with only an error item in the evaluation function.

The error item in Eq.1 is a matrix comprised of the variables **h** and **f**, which are vectors that represent the high-resolution image and captured images, respectively, while C expresses the optical degradation in the image-capture system. This optical degradation matrix is sought in the following way.

Usually, when CCD elements are used to sample and capture an image, processing occurs in the order shown in Figure 3. The captured image of the subject passes through the camera's optical system and image formation occurs in the CCD elements. This image depends on the positional relationship between the captured image subject and the CCD. Furthermore, the image formed is degraded by the lens and other parts of the optical system as well as by blur. Moreover, sampling occurs at the CCD element position, multiplying noise before output of the captured image.

Captured images of the subject are replaced with a high-resolution image, and, consequently, sampling from the natural image is changed to down-sampling from the high-resolution image.

Accordingly, C is defined as the conversion matrix that is a combination of the position adjustment of the high-resolution image with the image-capture positions of each captured image, image degradation using the PSF in the camera, and sampling by the pixel size of the captured image.

In the Eq.1 constraint item, α is the significance component of the constraint item in relation to the error item, while B is the matrix that expresses preexisting information for the high-resolution image. To express this preexisting information, a method that applies a secondary differential norm using a Laplacian filter has been reported,[16] for example.

The following equation (Eq.2) differentiated with **h** is used when reconstructing.

$$\frac{1}{2}\frac{\partial E}{\partial \mathbf{h}} = C^T\left(C\mathbf{h} - \mathbf{f}\right) + \alpha B^T B\mathbf{h} \quad \text{(Eq.2)}$$

The steepest descent method can be used in reconstruction processing, but the contragradient method[7] and the contragradient method with pre-processing[14] are also used. In addition, a high-speed method that conducts reconstruction processing using frequency area has also been reported.[16]

This reconstruction processing ends when it achieves fixed final conditions. Final conditions can be set as, for example, (1) when the evaluation function goes below a threshold, (2) when the high-resolution image change rate goes below a threshold, or (3) when the number of reconstruction iterations goes above a threshold.

3.2 Super-resolution processing with simultaneous noise extrapolation

The super-resolution processing proposed in this paper adds processing to extrapolate noise to a conventional ML approach. Inspection equipment requires images with accurate information about the subject. The reason we used a method based on the ML approach is because there is the possibility of degrading observation information in the error item due to the constraint item in the MAP approach.

The processing flow can be expressed in the form of Figure 4. In this figure, the left side loop is processing to extrapolate a high-resolution image in the same way as a conventional

ML approach, while the right side loop is processing to extrapolate noise components.

In order to conduct these extrapolation processes simultaneously in super-resolution processing, an evaluation function must be used that includes both high-resolution image and noise components. In this paper, we propose the evaluation function shown in Eq.3 for this purpose.

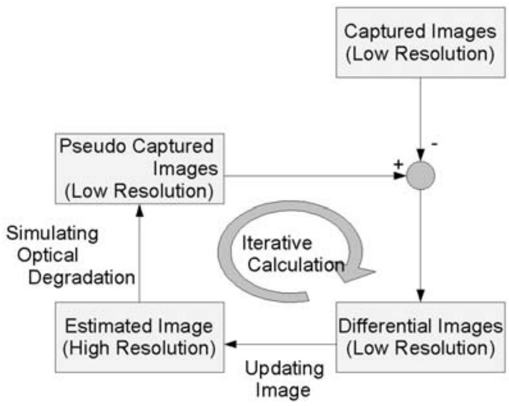


Figure 2: Processing loop for reconstruction based super-resolution processing.

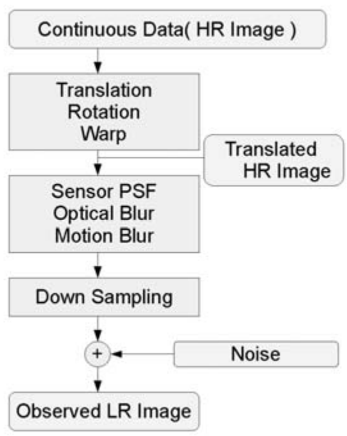


Figure 3: Image sampling by CCD elements.

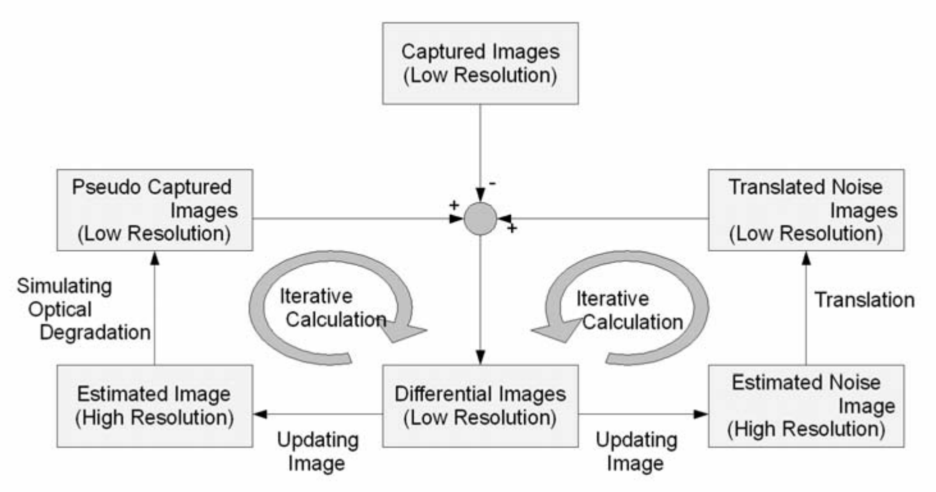


Figure 4: Processing block diagram for super-resolution processing with simultaneous noise extrapolation.

$$E = \left| C\mathbf{h} - (\mathbf{f} - wK\mathbf{x}) \right|_2^2 \qquad \text{(Eq.3)}$$

In Eq.3, **h** and **f** express the vectors for the high-resolution image and the captured images, respectively, while C expresses the optical degradation that occurs in the image-capture system. In addition, K**x** is the noise component in the captured image. This noise component is expressed as the matrix K, which expresses the constituent noise components, and the product of noise strength vector x, which expresses the strength of each component. Moreover, w is the significance component that is used as a multiplier when the noise components are subtracted from the captured image.

Noise constituent components are the noise that can be thought to occur in captured images. We considered their characteristics and set these in advance. This noise can be thought to be, for example, (1) sensitivity variations among CCD elements and (2) white noise that occurs randomly in every element.

Thus, we set as noise constituent components (1) the sensitivity variation that affects each pixel in the captured image, and (2) white noise, which is a high-frequency component that should be cut by PSF. These are expressed by Eq.4 and Eq.5, respectively.

$$K = \left(\mathbf{e}_{00}, \mathbf{e}_{01} \cdots \mathbf{e}_{ij} \cdots \mathbf{e}_{mn} \right) \qquad \text{(Eq.4)}$$

$$CK\mathbf{x} < \mathbf{1} \qquad \text{(Eq.5)}$$

The sensitivity variation appears uniformly in all images captured by the same CCD element, so when there are multiple captured images, it occurs the same way in every image.

Accordingly, since **f** is expressed as the vector of the elements of [the number of pixels in the captured image] x [the number of captured images], $\mathbf{e}_{ij}$ is the vector in which the only element that corresponds to the (i, j) coordinate pixel of each captured image is 1 and the other components are 0. In addition, noise constituent components are expressed in a matrix with these vectors as row components.

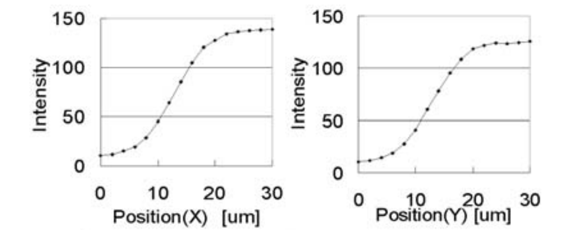
On the other hand, for random noise, the elements are expressed as Eq.5. In this equation, after multiplying the optical degradation, components that are of small numerical value as a result of this processing are rounded down and not expressed in the captured image. Consequently, components that become small numerical value as a result of multiplication by C are regarded as random noise.

High-resolution image and noise components are reconstructed to minimize the value of this evaluation function. The derivative functions for the high-resolution image h and noise strength vector x of these evaluation functions are found to be as in Eq.6 and Eq.7, respectively.

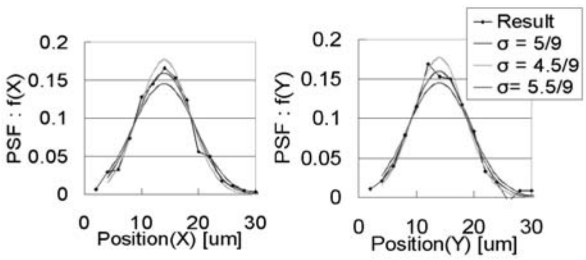
$$\frac{1}{2}\frac{\partial E}{\partial \mathbf{h}} = C^T \left(C\mathbf{h} - (\mathbf{f} - wK\mathbf{x}) \right) \qquad \text{(Eq.6)}$$

$$\frac{1}{2}\frac{\partial E}{\partial \mathbf{x}} = wK^T \left(C\mathbf{h} - (\mathbf{f} - wK\mathbf{x}) \right) \qquad \text{(Eq.7)}$$

Using these derivative functions, high-resolution image and noise components are reconstructed using the steepest descent method and other approaches.



(a)



(b)

Figure 5: Camera PSF measurement: (a) edge luminance value change, (b) PSF form derived from luminance differences.

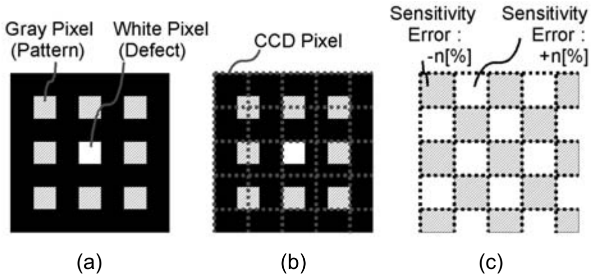


Figure 6: Preparation of simulation image: (a) original image, (b) image-capture element allocation compared to the original image, (c) CCD sensitivity variation.

4 NOISE REDUCTION EVALUATION USING SIMULATION IMAGES

4.1 Measurement of image capture conditions

Super-resolution processing includes processing of simulated optical degradation of the captured images through iterative operations. This optical degradation processing is a combination of the next three processes.

1. Movement processing to the image capture position.
2. Processing of image blur by the camera PSF.
3. Down-sampling to the resolution of the captured image.

Usually, the conversion matrix in No.1 is sought from differences in the pixel values of each captured image using the registration method.[17] However, since the camera image-capture position can be controlled with high-precision in inspection equipment, the conversion matrix can be assumed to be already known. Moreover, the conversion matrix in No.3 can be determined if the numbers of pixels in the high-resolution and captured images are determined.

However, measurement of the conversion matrix for No.2 from the image-capture equipment is still necessary. The camera PSF can be measured by capturing images of a subject with an edge while microscopically changing the image-capture position.

Figure 5(a) shows pixel value changes when the edge area is captured as the image-capture position is moved only in the X or Y direction. In addition, Figure 5(b) shows a calculation of the PSF form from luminance value differences between each position, as well as several different PSF forms that account for standard deviation, assuming that the PSF form has a normal distribution. From this graph, it can be seen that the camera PSF can approximate a normal distribution with a standard deviation of 0.56(=5.5/9) [CCD Pix].

4.2 Noise reduction evaluation using simulation images

Next, we will discuss evaluation of noise reduction using super-resolution processing with simultaneous noise extrapolation. We prepared a grid pattern image assuming that inspection subjects would have minute patterns. We sampled this at a resolution equivalent to that used for image-capture in automatic inspection equipment. Furthermore, we also prepared images with CCD element sensitivity variation and random noise added and used these for image capture.

Figure 6 shows the method used for preparing the images used in the evaluation. Figure 6 (a) with a pattern of grey squares arranged on a black background is the original image that corresponds to the inspection subject. Some of the grey squares are drawn brighter than the other grey squares, corresponding to defects in the inspection subject.

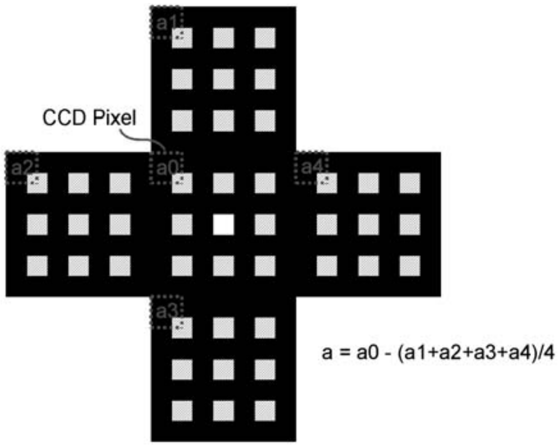


Figure 7: Signal value calculation through comparison of the luminance values of image-capture elements.

This was sampled as in Figure 6 (b) with the rectangles showing the size of the CCD elements. A Gauss function with a standard deviation of 0.56 [CCD Pix] was used for the PSF.

Next, noise from CCD element sensitivity variation was multiplied to the image. We assumed that CCD element sensitivity variation was distributed in a chessboard form as in Figure 6 (c) and multiplied these values. In short, if sensitivity variation is ±n [%], the value of elements that correspond to light squares on the chessboard are increased while the value of elements that correspond to dark squares are reduced by n [%]. Finally, we added random noise with a Gauss function with a dispersion of which standard deviation was 1.43 to create captured images that would be the sources for super-resolution processing.

The number of captured images that can be used in inspection equipment is constrained by processing time in the manufacturing process. We conducted super-resolution processing using nine images that correspond to image-capture at a total of nine positions with the CCD elements shifted exactly 1/3 in the X and Y directions.

In addition, in evaluation of this super-resolution processing, we used the ratio of the detected values in the areas with defects to the size of the noise existing in the good parts as the S/N ratio. Since areas with defects are not buried in surrounding noise when this S/N ratio value is high, it can be said that it is possible to conduct detection easily.

In order to find the detection values of the noise and defect areas, it is necessary to remove the grid pattern in the simulation image. The method that we used for this was to calculate the average of 4 pixels in identical adjacent patterns, as shown in Figure 7, and subtract that average value from the central pixel value.

Figure 8 shows the CCD element variation and S/N ratio relationships in high-resolution images extrapolated using the ML approach, the ML approach with post-processing smoothing and super-resolution processing with simultaneous noise extrapolation.

Moreover, in this super-resolution processing, the noise constituent components are expressed by Eq.5 and the

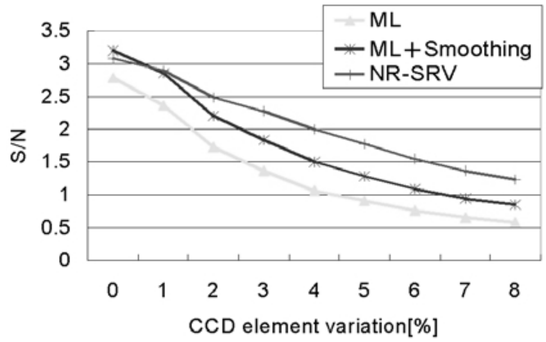


Figure 8: CCD element variation and S/N ratio relationships in each type of super-resolution processing.

significance when subtracting the noise image from the original image is 2.0.

As a result, the method described in this paper has an improved overall S/N ratio compared to the conventional ML approach. In addition, we can see that the ML approach using smoothing to achieve a high-resolution image also shows a high S/N ratio value when there is no CCD sensitivity variation, but the method described in this paper has a higher S/N ratio when there is sensitivity variation.

5 SUMMARY

In this paper, we proposed a method with the goal of improving the S/N ratio in super-resolution processing in automatic inspection equipment. This method simultaneously extrapolates noise components and removes them from the captured image when extrapolating a high-resolution image.

In order to extrapolate noise components, regarding them as linear combinations of the constituent components, we first defined the noise constituent components that occur in the captured image and extrapolated the strength of each one.

In addition, we considered CCD element allocation in relation to inspection subjects with minute patterns, and we created a simulation image that corresponds to that element allocation at the time of image capture to evaluate the proposed method. We compared these results with a conventional ML approach and an ML approach with post-processing smoothing, confirming that the S/N ratio in this method is better.

In the future, along with making this method faster, we will develop an algorithm that allows application to even more diverse images by advancing the method for expressing the noise components.

[1] Sung C.P., Min K.P. et al, 2003, Super-Resolution Image Reconstruction: A Technical Overview, IEEE Signal Proc. Magazine, Vol.26, no. 3, pp.21-36.

[2] B.C. Tom and A.K. Katsaggelos, 1995, Reconstruction of a high-resolution image by simultaneous registration, restoration, and interpolation of low-resolution images, Proc. IEEE Int Conf. Image Processing, Vol.2, pp.539—542.

[3] M. Irani and S. Peleg, 1990, Super resolution from image sequences, Proc. of the 10th International Conference on Pattern Recognition, Vol.2, pp.115-120.

[4] S. Mann and R.W. Picard, 1994, Virtual bellows: Constructing high quality stills from video, Proc. IEEE Int. Conf. Image Processing, Austin, TX, pp. 13-16.

[5] M. Irani and S. Peleg, 1993, Motion analysis for image enhancement resolution, occlusion, and transparency, J. Visual Commun. Image Represent., vol.4, pp. 324-335.

[6] R.R. Schulz and R.L. Stevenson, 1996, Extraction of high-resolution frames from video sequences, IEEE Trans. Image Processing, Vol.5, pp.996-1011.

[7] R.C. Hardie, K.J. Barnard, and E.E. Armstrong, 1997, Joint MAP registration and high-resolution image estimation using a sequence of undersampled images, IEEE Trans. Image Processing., vol. 6, pp. 1621-1633.

[8] P. Cheeseman, B. Kanefsky, R. Kraft, J. Stutz, and R. Hanson, 1994, Super-resolved surface reconstruction from multiple images NASA Ames Research Center, Moffett Field, CA, Tech. Rep. FIA-94-12.

[9] H. Stark and P. Oskoui, 1989, High resolution image recovery from image-plane arrays, using convex projections, J. Opt. Soc. Am. A, vol. 6, pp.1715-1726.

[10] A.M. Tekalp, M.K. Ozkan, and M.I. Sezan, 1992, High-resolution image reconstruction from lower-resolution image sequences and space varying image restoration, Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP), San Francisco, CA., vol. 3, Mar. 1992, pp. 169-172.

[11] S. Baker and T. Kanade, 2002, Limits on Super-Resolution and How to Break Them, IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 24, no. 9, pp. 1167-1183, 2002.

[12] Z. Lin and H.Y. Shum, 2004, Fundamental Limits of Reconstruction-Based Superreslution Algorithms under Local Translation, IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 26, no. 1, pp. 83-97, 2004.

[13] M. Tanaka and M.Okutomi, 2005, Theoretical Analysis on Reconstruction-Based Super-Resolution for an Arbitrary PSF, International Conference on Computer Vision and Pattern Recognition (CVPR2005).

[14] N. Nguyen, P. Milanfar et al, 2001, A Computationally Efficient Superresolution Image Reconstruction Algorithm, IEEE Trans. Image Processing, Vol.10, No.4, pp.573-583.

[15] M. Tanaka and M. Okutomi, 2004, Fast Algorithm for Reconstruction-based Super-resolution, 2004-CVIM-146-(14), Vol.2004, No.113, pp.97-104.

[16] M. Tanaka and M. Okutomi, 2005, A Fast Algorithm MAP Super-Resolution by Optimization in Frequancy Domain, MIRU2005, pp.64-71.

[17] V.N. Dvorchenko, 1983, Bounds on (deterministic) correlation applications to registration, IEEE Trans. Pattern Anal. Machine Intell., vol.5, no.2, pp.206-213.

Instrumented fixtures for on-line correction of welding paths

T.K.Lien[1] and M. Lind[1]

[1] Department of Production and Quality Engineering,

Norwegian University of Science and Technology, Trondheim, Norway

Abstract

Robotic welding of thin-walled aluminium structures is very sensitive to deviations in parts' positions. Position deviations give inferior weld quality. Measurement devices integrated in the clamping elements of the welding fixture can provide data to the welding robot for dynamic correction of the welding path. The paper describes a system of integrated measurement in clamping systems. Various approaches to measurement of critical dimension deviations are evaluated and optimal solutions have been selected. The obtainable accuracy is analyzed and verified through laboratory tests. A microcontroller-based control system has been developed to process and transmit the measurement data to the robot controller.

Keywords:

Robot; Welding; Measurement; Fixture

1 INTRODUCTION

All manufacturing operations that are performed automatically by some sort of machinery need a firm positioning of the work piece. This is one of the most fundamental truths of manufacturing. In many cases this is not a big problem either. But in modern highly automated processes where both the primary operation and the product or component has been automated challenges appear also in this area. The removal of human interference in the positioning and fixation of the work piece sometimes leads to challenges in securing a proper positioning to assure the wanted outcome of the manufacturing process.

2 HANDLING LARGE PART DIMENSION VARIATIONS

2.1 The problem of part tolerances outside the process window

Even products with wide tolerances can present big challenges to automation. One typical problem is that neither the final product nor the ingoing parts in an operation have particularly narrow tolerances. But the process in itself may require a much tighter tolerance to yield a good quality result. Typical examples in this category are grinding, polishing and welding. The case examined in this paper concerns welding.

The parts in many weldments can have quite wide tolerances if the weld is manually performed. Even if the parts have slight variation in relative positions from product to product the yield of the process is good since the operator will adjust the welding path according to observation in each case. It is not so in automatic welding. The robot or automatic welding machine will assume that the parts are identically placed in each operation. But it may well be that the parts are offset more than 1 mm from the ideal position even on relatively small parts. Such large offsets will create problems in the welding of thin walled components. Because of the thin material only small welding currents are allowed. The filling capacity is thus limited and the burn-in into the base material can be too low to assure proper welding strength.

2.2 Performance requirement in the case study

In this case study the challenge was to weld thin walled extruded box profiles to a base plate. The product is shown in figure 1. This is a simplified example of a component in automobile bumper system, the so called crash box. The wall thickness of the box typically lies in the region around 2mm, while the outer dimension tolerance of the box is in the order of ±0.9 mm in the worst case. The expected variation in outer dimension of the box is too large for a good quality automated welding process.

One possible solution to the problem is to use the well established method of edge detection by means of the welding gun as a touch probe [1,2]. But this method has limited accuracy, and it is time consuming since it has to be performed after the parts in the weldments has been locked in the fixture and before the actual welding starts. The time consumed for this pre-weld check was considered to be too high. It would severely reduce the productivity of the automated welding operation.

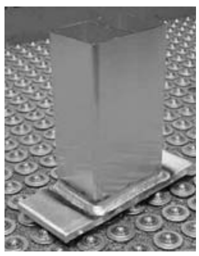


Figure 1: The study object, a simplified crash box

The alternative that was studied was to include measurement of the position of the extruded box as it is clamped in the fixture. By using this method the exact position of each box would be known at the moment it was clamped. This position information could then be used to adjust the welding path for the automatic welding operation.

For best quality welding the welding path should be corrected to within ±0.1mm of the ideal path. This is a conservative requirement to make sure that local variations along the weld path will not give risk of inferior local quality.

3 METHODS FOR MEASUREMENT COMBINED WITH CLAMPING

3.1 Evaluation of possible methods

The measurement of the position of the workpiece while clamping can be done either directly or indirectly through the clamping elements. Both possibilities have advantages and disadvantages as shown in table 1 [3,4].

Table 1: Comparison of measuring methods

Measuring method comparison	
Direct measurement on object	
Advantages	Disadvantages
• Most precise method	• Measurement device will interfere with clamping device or welding gun • Expensive measurement devices
Indirect measurement through clamping element	
Advantages	Disadvantages
• Measurement device away from clamp and weld area • Low costs measuring devices can be used • Mechanical protection of measurement device easy	• Less accurate due to coupling through clamp • Non-linearity in some of the applicable measurement methods

Table 2: Characteristics of clamp position measurement methods

Method	Merits
Rotational encoder on the clamp arm shaft	Low cost robust solution. Non linear measurement
Laser distance measurement of clamp arm position	Linear measurement Robust solution High cost
Pneumatic linear distance measurement of clamp arm position	Linear measurement Robust low cost solution Short measurement range
Inductive linear distance measurement of clamp arm position	Linear measurement Robust low cost solution Short measurement range
Touch probe linear distance measurement of clamp arm position	Linear measurement Less robust High cost

The evaluation of measuring methods placed strong emphasis on robustness, simplicity and reliability. Thus the qualitative advantages and disadvantages listed in table 1 lead to the clear conclusion that indirect measurement through the clamping elements was the better method.

For clamping both linear and rotational clamps were considered. Again the space requirements lead to the conclusion that rotational arm clamps would be the best solution. It then remains to decide which measurement system that would be most suitable for measurement of the clamp movement. Table 2 shows the candidates and their merits [3,4].

Again the requirement for robustness and low cost were dominating. In addition it was realized that only the rotational encoder could be fitted without any major mechanical interference problems with any other element of the system. The final design ended up with rotational clamps that included incremental rotary encodes for position measurement of the clamped surface.

4 THE COMPLETE INSTRUMENTED WELDING FIXTURE

Figure 2 shows the laboratory version of the instrumented welding jig. The work piece is clamped against two vertical fixed walls with three contact points on one wall and two on the other to ensure orthogonality and stable positioning. The variation in dimension of the workpiece can then be measured directly on the two surfaces pressed on by the clamps.

The clamps are pneumatically operated. The operation of the clamps and reading of the measured signals is controlled by a local microcontroller that communicates with a PC that acts as an overall process controller. This PC generates the path correction data for the welding robot controller.

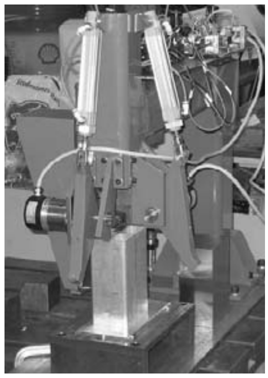


Figure 2: Laboratory version of the complete welding jig with one work piece in position ready for clamping.

4.1 The precision of the selected clamping method

Figure 3 shows a sketch of the claming-measurement arm principle. In an ideal design the distance x in figure 3 should be 0 for the calibrating position. But for mechanical reasons the clamping arm had to be L-shaped, thus there will be a substantial x value in the measurement setup. In this design the clamp arm had the dimensions X=20 mm, Y=20 mm giving a total length L=101.98 mm.

This gives an amplification of the non-linear behaviour of this measuring principle. It is of interest to evaluate how large the error due to non-linearity will be. The measurement system is calibrated against a master block with dimensional tolerance better than ±0.01 mm. This assures a reference for the measurement system with an accuracy one order of magnitude smaller than the required correction accuracy.

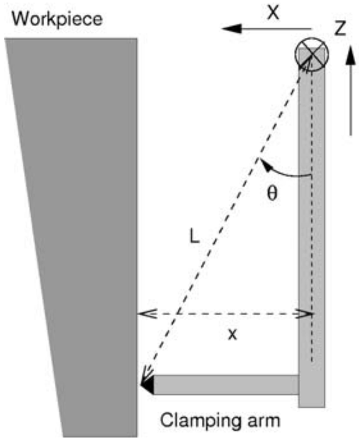


Figure 3: Sketch of the principle of the rotating arm clamp and measure device.

Let the angle θ_0 represent the reference angle corresponding to the reference position x_0 as is given in equation (1)

$$x_0 = L\sin\theta_0 \quad (1)$$

A small measurement distance Δx_s relative to x_0 can be expressed as:

$$\Delta x_s = L(\sin(\theta_0 + \Delta\theta_s) - \sin\theta_0) \quad (2)$$

For simplicity the measurement uses a linearization of Δx_s around x_0 as given in (3):

$$\Delta\tilde{x}_s = -\Delta\theta_s L\cos\theta_0 \quad (3)$$

The error due to this linearization is the difference between equations (2) and (3):

$$\Delta x_{err} = L\left[\sin(\theta_0 + \Delta\theta_s) - \sin\theta_0 + \Delta\theta_s\cos\theta_0\right] \quad (4)$$

Figure 4 shows the variation in Δx_{err} as function of Δx_s for different values of x_0. This error is one order of magnitude less than the resolution of the measuring system.

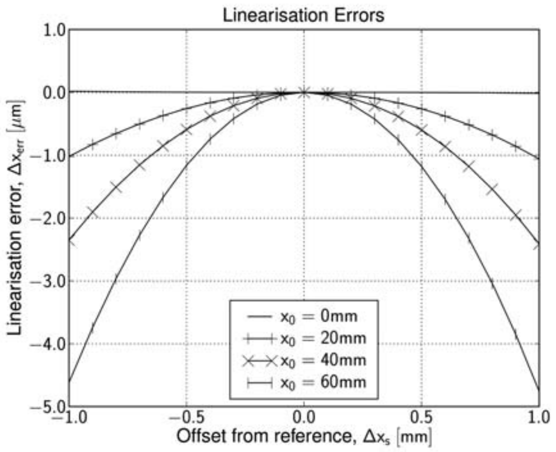


Figure 4: Theoretical measurement error Δx_{err} as function of measurement distance Δx_s for various x_0 values.

5 CONTROL SYSTEM FOR ROBOT PATH CORRECTIONS

The control PC is the central unit, through which a user initiates all action chains. The user interface developed has three modes: communication setup, calibration control, and production control. In communication setup mode, communication to the micro controller and the robot controller is specified and verified. In calibration mode, a semi-automatic calibration routine can be performed and stored. Production mode presents to the user a chain of check points for the whole operation cycle. The check points serve only safety purposes for the prototype. The necessary external signals for fully automated operation are that the welding is done and that a new crash box has been positioned in the fixture.

The robot controller is set up with a program where the welding path is taught from the nominal crash box. The robot controller has program code for receiving offsets over the RS232 interface from the control PC. Further, the robot controller listens over the RS232 interface for a command from the control PC for executing the offset corrected weld program.

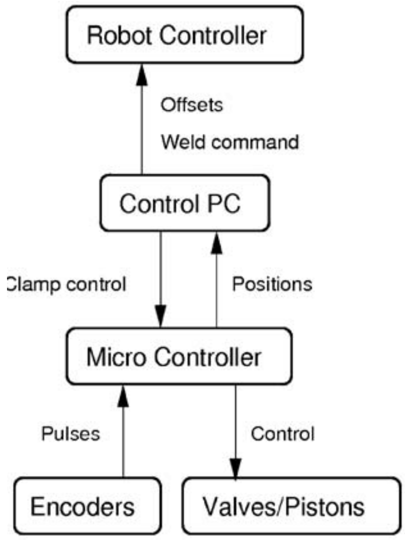


Figure 5: Block structure of the control system.

The micro controller traps pin change interrupts on one parallel port, where the 3 channels from each encoder are connected. Two of the lines from an encoder forms the quadrature pulse train, on basis of which the encoder position counter is updated. The third line for an encoder gives the index pulse, generated at a fixed position in the encoder. The index pulse interrupt from an encoder resets the encoder pulse counter to zero, thus eliminating possible drift by loss of pulses. The micro controller will reply with instantaneous encoder positions over the RS232 interface by request.

The pneumatic system of 3 double acting pistons is controlled by 3 solenoid controlled valves. One parallel port of the micro controller has 3 pins dedicated to control the gates of 3 MOSFETs, which controls the valve solenoids. The RS232 interface exposes direct control of the valves to the control PC.

6 LABORATORY VERIFICATION OF THE PRINCIPLE OF MEASUREMENT

The selected encoder for the clamp arm has a resolution of 5000 pulses/rev. With quadrature counting principle this gives 20000 counts/rev and the selected arm design gives a resolution of 32 μm/count. This should give sufficient accuracy.

A laboratory test was performed to verify that the expected measurement resolution 30 pulses/mm did not create problems for the measurement linearity. The test was performed by using the calibration block as base. Successively laboratory grade gauge blocks were stacked to create an offset from the calibration block, starting with a 1 mm gauge block as the offset 0 point. Figure 6 shows the measured result. It is in perfect agreement with the expected performance, giving 3 pulses per 0.1 mm offset change. The change in pulse steps is also very regular switching between 1 and 2 pulses per 0.05 mm. This regular pattern assures that no extra inaccuracy is introduced by the granularity in measurements near the reference point.

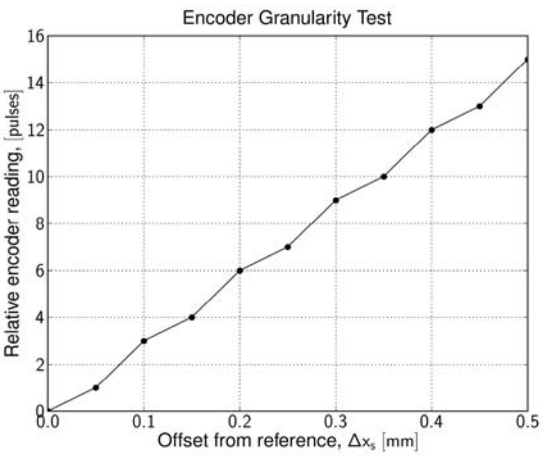


Figure 6: Granularity test for small linear offsets

7 LABORATORY WELDING TESTS

Real welding tests were performed to verify the performance of the correction system. Figure 7 shows a cross-section of a perfect weld performed with optimal path programming on a part with nominal dimensions. Figure 8 shows a weld on a part with 2 mm offset of the box wall. This weld has inferior quality because of too little melt-in into the box wall. Use of the path correction system results in the weld shown in figure 9 for a wall offset of 2mm. This is a weld of similar quality as the reference weld.

A series of similar welding tests were performed. The all gave similar results. The industrial partner in the project feels confident that the system performs as required to compensate for dimension tolerance on ingoing parts.

8 SUMMARY AND CONCLUSIONS

- A system has been built for making measurement of part dimension offset combined with the clamping action in a welding fixture.
- A theoretical model for the measurement error has been developed.
- A control system to operate both clamping and offset measurement has been developed.

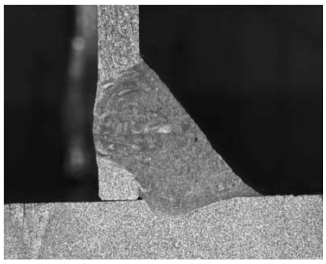


Figure 7: Reference weld of satisfying quality.

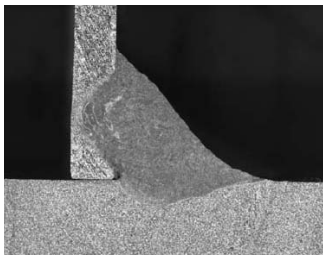


Figure 8: Weld test with 2 mm offset uncorrected, insufficient melt-in on box wall

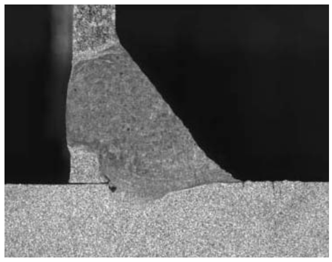


Figure 9: Weld test with 2mm offset corrected through feed forward correction from the clamping measurement system

The complete system has been tested in a laboratory setup. The laboratory tests verified that the system performed as expected. It enables the measurement of individual parts offset during the clamping cycle to generate offset command to the welding robot before the welding operation starts. The corrected weld satisfies the quality requirement for automatic welding of the investigated and similar parts.

9 ACKNOWLEDGMENTS

This research work was sponsored by Hydro Aluminium and the Norwegian Research council

10 REFERENCES

[1] Ågren, B.G., 1995, Sensor integration for robotic arc welding, Department of Production and Materials Engineering, Lund University.

[2] Cary, H.B., 1995, Arc Welding Automation, Marcel Dekker, NY.

[3] Bolton, W., 2003, Mechatronics – Electronic control systems in mechanical and electrical engineering, Pearson Education Ltd, Harlow, England.

[4] Cetinkunt, S., 2007, Mechatronics, John Wiley &Sons, Hoboken, NJ.

Micro-Nano Technology and Surfaces

Surface Roughness and Microstructure in Ultrasonically Assisted Turning of W-Fe-Ni Alloy

Jie CHEN, Guangxue TIAN, Yonggang CHI, Mei LIU, Dongwei SHAN and Yuelin LIU
[1]China Academy of Engineering Physics

Abstract
The experiment is carried out of turning W-Fe-Ni alloy by conventional turning and ultrasonically assisted turning. Effects of cutting parameters and vibration parameters on surface roughness are studied. The significant factors influence surface roughness by the orders of feed rate, tool amplitude and tool frequency in ultrasonically assisted turning, whereas the main factors influence surface roughness is feed rate in conventional turning. Using ultrasonically assisted turning can reduce surface roughness by 7.05%30.06% compared with conventional turning in same cutting parameters in experiments. To analyze the characteristics of surface machined, X ray diffractometer, white light interferometer and scanning electronic microscope are used to observe micro surface profile. It is confirmed that apart from the regular feed marks with an equal spacing, there are also regular parallel stripes on the finished surface in ultrasonically assisted turning. It is intense ironing effects that reduce the surface roughness and increase compressive stress. A high quality surface can be obtained by selecting appropriate parameters in ultrasonically assisted turning.

Keywords:
ultrasonically assisted turning; surface roughness; scanning electronic microscope

1 INTRODUCTION

Ultrasonically assisted turning (UAT) is an advanced machining technique. Usually, to break chips use low-frequency (several hundreds Hz) and large amplitude(maximum to several millimeter),whereas to improve machining use high-frequency (≈ 20kHz) and small amplitude(maximum to tens micron). The high-frequency vibration is superimposed on the movement of the cutting tool. Most of the investigators have studied on ultrasonically assisted turning. This technology demonstrates a range of benefits in machining hard material alloys: a decrease in cutting forces of up to several times [1][2]. The effects of cutting parameters on surface roughness have been analyzed on various materials. Improvement in surface finish by up to 50% compare to CT[3][4][5]. In this paper, compare the effects of cutting parameters and vibration parameters on surface roughness in ultrasonically assisted turning with conventional turning, a series of tests were made on machined parts and the results were analyzed to understand UAT process.

2 EXPERIMENTAL PROCEDURE

2.1 Materials and Equipment

The work piece material was W-Fe-Ni alloys, sintered powder materials. Chemical composition of W-Fe-Ni alloys was 93W-4.5Ni-2Fe-0.5Co. The mechanical properties of W-Fe-Ni alloy are given in table 1.

Table1: Mechanical properties of W-Fe-Ni alloys.

density/g cm^3	R_m/MPa	$R_{0.2}$ /MPa	$A_{5.65}$ /%	Z/%	E/× 10^5MPa
17.68	980	595	18	14.0	3.4

The cutting tool used was made by SANDVIK Company with an ISO designation of VBMT110304-KF, and materials is H13A.The inserts were clamped onto vibration tool holder. The major tool geometry is defined by a positive rake, a clearance angle 5º, a nose radius of 0.4mm and a cutting edge angle 107º. The experimental setup used to study UAT is shown in Figure 1. The work piece is clamped in the chuck of the CK6160 CNC lathe and rotates with a constant speed. High frequency electric impulses, fed to the input of the ultrasonic transducer, excite vibration. The vibration amplitude is intensified in the concentrator and transmitted to the tool holder at the end of the concentrator. Resultant vibration of the cutting tip fixed in the tool holder reaches 4μm (i.e. 8 μm peak-to-peak) at a frequency of about 20 kHz. The vibration direction is in the direction tangential to the surface of the work piece. The turning tests were conducted in dry conditions.

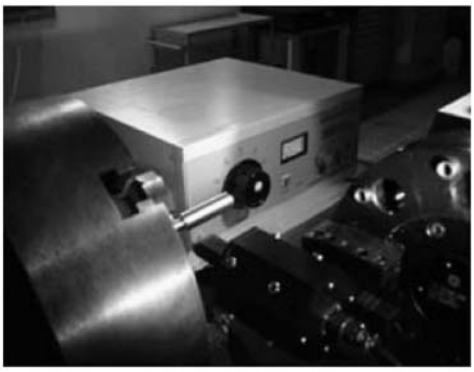


Figure 1: Experimental setup for ultrasonic assisted turning.

2.2 Experimental method and measurement

In this study the same parameters were used for making the comparative tests between UAT and CT. At the both end of specimen were turned by UAT and CT method respectively. Switch off the ultrasonic generator to proceed with CT and switch on the ultrasonic generator to proceed with ultrasonic turning.

An Orthogonal Test Method was selected to investigate the influence of independent variable on surface machined. Independent variables were amplitude, vibration frequency, feed rate, depth of cut and cutting speed in this experiment. The level of the independent variables were selected to cover the normal cutting operations and listed as follows: cutting speed 100,250,400(r/min), feed rate 0.04,0.08,0.12(mm/r), deep of cut 0.04,0.06,0.08(mm), amplitude 1,2.5,4(μm) and vibration frequency high(20897~21022Hz),middle (20509~20630Hz) and low(20030~20109Hz). The cutting parameters selected and corresponding measurements after machining were shown in table 2.

3 ANALYSES AND DISCUSSION

3.1 ANOVA of independent variables in UAT

The plan of the experiment was developed for assessing the influence of the five independent variables on the surface roughness. The experimental results were analyzed with analysis of variance (ANOVA), which is used for identifying the factors significantly affecting the performance measures.

Hypothesizing the levels of independent variables is m, repeated experimental times under each level is k, the sum of experimental results is T. The formula of calculating Sf was as follows:

$$s_f = \frac{1}{k} \times \sum_{i=1}^{m} (\sum_{j=1}^{k} y_{ij})^2 - \frac{T^2}{km} \quad (1)$$

The degrees of freedom of factors (ff) and errors (fe) both are 2, thus F-ratio equation calculated was as follows:

$$F = \frac{S_f / f_f}{S_e / f_e} \quad (2)$$

This analysis was carried out for a significance level of α=0.10, i.e. for a confidence level of 90%. According to measurement in table 2,the results of analysis of variance were shown in table 3. The value of F-critical was considered the realized significance levels. Which F-ratio value is larger than F-critical value shows that which factor has a highly significant contribution to the performance measurements.

Table 3 shows that the only significant factor for the surface roughness in UAT is feed rate. The next largest contributions are by the order of amplitude, depth of cut, frequency and cutting speed. While in the CT, the factors having influence in the Ra parameter are feed rate and depth of cut, not being the influence of cutting speed significant. These results show that vibration parameters play an important role in UAT.

In conventional turning, the surface roughness of the machined part is known to be affected mainly by the feed and tool nose radius. The geometric contribution of tool nose geometry and tool feed is called theoretical surface roughness and is given approximately by the following equation:

$$R_{th} \cong \frac{f^2}{8r_n} \quad (3)$$

Where f is the feed rate and r_n is the tool nose radius, represents the theoretical peak-to-valley surface roughness. However, manufacturing processes do not allow achieving the theoretical surface roughness due to defects appearing on machined surfaces and mainly generated by deficiencies and imbalances in the process. Surface roughness depends on cutting parameters and the irregularities of machining operations such as tool wear, chatter, tool deflections, and work pieces properties. In conventional turning, with f increasing, the over-lap effects weaken in the feed direction and Ra increases at an increasing rate. With depth of cut increasing, deformation of chip in shear zone increases, which increases the risk of excessive chipping or fracture and surface roughness. The cutting speed mainly affects the cutting force. With cutting forces increasing, cutting forces decreases and chatter increases. It makes a negative influence on surface roughness and easy for turning. The mechanism of turning process in UAT is different from that of CT. With high frequency vibration of small amplitude on the cutting tool, the interaction conditions between the contacting surfaces of the tool and the work piece are different from those in conventional cutting process. In ultrasonic vibration cutting, instead of static friction between the tool and the work piece, dynamic friction will be generated and the friction between the tool and the work-material will be reduced. The additional reciprocating movement between the contacting surfaces of the tool and the work piece results in the periodical change of the friction force on the rake and flank face of the tool. The resistance to chip flow is decreased and the chips can be generated under better conditions. There will cause a remarkable reduction in the cutting force and in the cutting temperature. As a result, the possibility of crack initiation and its propagation can be reduced. Vibration parameters have a more significant influence on surface roughness than the other two parameters (depth of cut and cutting speed).

3.2 Comparison of surface roughness in UAT and CT

The average surface roughness is taken for this study. Taking account of random errors of measurements, five measurements were proceeded at different part of machined surface. The surface roughness machined by UAT and CT is shown in figure 2. Comparing with CT, the UAT has a smoother surface at each comparative test. Comprehensive effects between cutting parameter and vibration parameter make differences in each comparative test. Comparing with CT, the surface roughness machined by UAT reduced by the range of 7.05%~30.06% for specimens in all comparative tests.

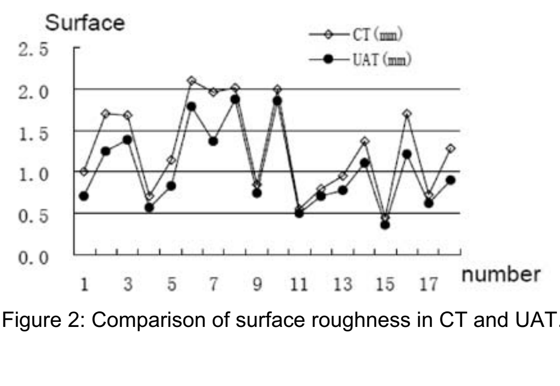


Figure 2: Comparison of surface roughness in CT and UAT.

3.3 The surface microstructure machined by CT and UAT

Figure 3 shows profiles of the machined surface to be used white light interferometer. Figure 3(a) being the surface profile machined without ultrasonic vibration and Figure 3(b) that machined with ultrasonic vibration. The distance between every two peaks of the wave reflects the profile of surface machined. From Figure3, it is observed that the surface can easily produce some defects such as burrs, tearing and so on in CT, thus the wave length has deeper peak-to-valley. While the profile of surface machined by UAT turns to even, so as to make the cutting process more stable.

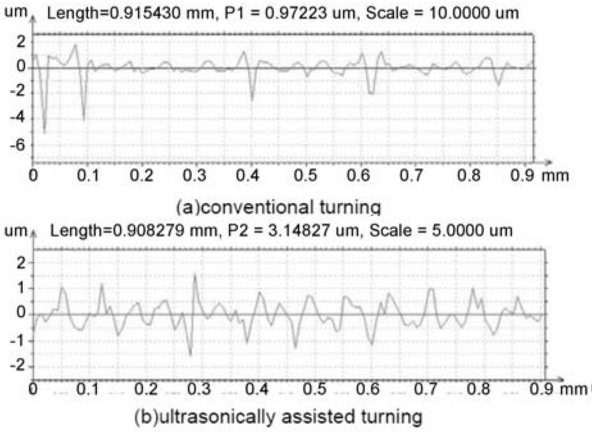


Figure 3: The profile of the machined surface measured by white light interferometer.
v=225r/min,ap=0.08mm,f=0.04mm/r,f'=20827Hz,A=4um

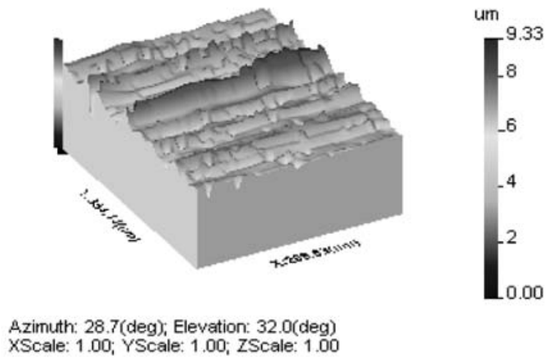


Figure 4:Microstructure of the surface machined by UAT.

Figure 4 shows the 3D surface machined by UAT, plot data coming from measurement of white light interferometer. It can be seen that on the bottom of the grooves and traces of tools. To investigate the characteristics of the surface in detail, the machined surface was observed using SEM. Figure 5(a), 5(b) shows the details of machining surface in CT and UAT respectively. Regular feed marks can be seen in Figure 5(a). While in Figure 5(b), apart from the regular feed marks with an equal spacing, there are also the smaller transverse marks, which are formed by tool vibration in the thrust force direction. The spacing of the stripes is determined by the ratio of the cutting speed to the vibration frequency. From photograph, it can be observed that the cutting grooves are smoother in UAT, while the surface machined by CT illustrates numerous cracks and a great deal of fragmentation at the fringe of the grooves, so the surface quality is poor.

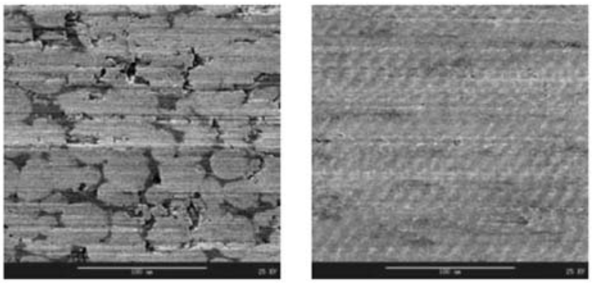


Figure 5: SEM images of conventional and ultrasonically assisted turning.
v=300r/min,ap=0.06mm,f=0.04mm/r,f'=21030Hz,A=1um

3.4 Residual stress of the machined surface

In order to study the difference of the surface residual stress between UAT and CT. the tangential and axial residual stress were tested by X ray diffractometer. According to the measurements, the residual stress in vibration cutting is bigger than that in CT by 400700Mp a. Because the rake surface of the vibration cutter has an intense ironing effect on the work piece surface. When cutting speed increased, the separating time of cutter and chip becomes short, the ironing effects of the cutter weaken, and the residual stress reduces.

4 CONCLUSIONS

The following conclusions can be drawn based on the results of the experimental study on turning W-Fe-Ni Alloy by conventional turning and ultrasonic assisted turning.

- In ultrasonic assisted turning, feed rate still has significant influence on surface roughness, next important factors are vibration parameters, while the influence of DOC and cutting speed on surface roughness become less profound than others parameters.
- Comparing with CT, the surface roughness machined by UAT reduced by the range of 7.05%30.06% for specimens in all comparative tests. It proves that UAT can obtain smoother surface.
- Because of the unstable turning process in CT, the surface can easily produce some defects such as burrs, tearing and so on, so the quality of surface becomes poor. While the UAT can reduce the influence of deformation and restrain flutter because of high frequency reciprocating movement between the contacting surfaces of the tool and the work piece, so as to make the turning process more stable.
- The residual stress in vibration cutting is bigger than that in CT by 400700Mpa. It confirms that the rake surface of the vibration cutter has an intense ironing effect on the work piece surface.

5 ACKNOWLEDGEMENTS

The authors would like to thank Mr. Xiaobin Xian for his valuable support, and Mr. Pin Dong, Mr. Dinmu Lang and Mr. Xuechao Lv for their testing of specimens.

6 REFERENCES

[1] V.I. Babitsky, A.N. Kalashnikov, A. Meadows etal. Ultrasonically assisted turning of aviation materials[J]. J. Mater. Process. Technol. 132(1-3)(2003):157-167.

[2] A.R. Sharman. P. Bowen,etal. Ultrasonic Assisted Turning of Gamma Titanium Aluminide, Rolls-Royceplc,2001.

[3] A.V. Mitrofanov, N. Ahmed, etal. Effect of lubrication and cutting parameters on ultrasonically assisted turning of Inconel 718[J]. J. Mater. Process. Technol. 162-163(2005):649-654.

[4] V.I.Babitsky,A.V.Mitrofanov,V.V.Silberschmidt. Ultrasonically assisted turning of aviation materials: simulations and experimental study[J]. Ultrasonics42(2004):81-86.

[5] Jeong-Du Kim,In-Hyu Choi. Micro surface phenomenon of ductile cutting in the ultrasonic vibration cutting of optical plastics[J]. Journal of Materials Processing Technology. 68(1997):89-98.

[6] X. Wang, M. Zhou, J.K.G. Gan, B. Ngoi, Theoretical and experimental studies of ultra precision machining of brittle materials with ultrasonic vibration[J], International Journal of Advance Manufacturing Technology 20 (2002) 99~102.

Table 2: Measurement results of conventional turning and ultrasonically assisted turning.

cutting speed(r/min)	depth of cut(mm)	Feed rate(mm/r)	frequency（Hz）	amplitude（μm）	CT(μm) Average	UAT(μm) Average
100	0.04	0.04	20030	1	1.003	0.706
100	0.06	0.08	20613	2.5	1.697	1.257
100	0.08	0.12	20931	4	1.690	1.382
250	0.04	0.04	20630	2.5	0.708	0.575
250	0.06	0.08	20986	4	1.145	0.839
250	0.08	0.12	20075	1	2.105	1.783
400	0.04	0.08	20109	4	1.956	1.368
400	0.06	0.12	20594	1	2.010	1.868
400	0.08	0.04	20967	2.5	0.842	0.752
100	0.04	0.12	21022	2.5	1.998	1.850
100	0.06	0.04	20058	4	0.553	0.496
100	0.08	0.08	20509	1	0.797	0.707
250	0.04	0.08	20938	1	0.953	0.774
250	0.06	0.12	20084	2.5	1.370	1.119
250	0.08	0.04	20531	4	0.459	0.362
400	0.04	0.12	20552	4	1.706	1.217
400	0.06	0.04	20953	1	0.728	0.633
400	0.08	0.08	20078	2.5	1.281	0.901

Table3: Variance analysis for ultrasonically assisted turning.

factor	sum of squares	d.f.	F-ratio	F-critical	notability
cutting speed	2.989	2	1.734	19.0	
depth of cut	4.330	2	2.512	19.0	
feed rate	33.856	2	19.64	19.0	high
frequency	3.964	2	2.299	19.0	
amplitude	4.720	2	2.783	19.0	
error	6.034	2			

Finishing Complex Surfaces with Zonal Polishing Tools

Fritz Klocke[1], Christian Brecher[2], Richard Zunke[1] and Roland Tuecks[2]

[1] Department of Process Technology, Fraunhofer Institute for Production Technology, Aachen, Germany

[2] Department of Production Machines, Fraunhofer Institute for Production Technology, Aachen, Germany

Abstract

High quality surfaces in terms of low roughness and high form accuracy are achieved by polishing as the essential finishing step. The machining is usually limited to planar or spherical geometries. Finishing of complex geometries require high efforts concerning machine tool and process development. The objective of this paper is to present first results to overcome these limitations indicated by processing advanced ceramics. In order to shorten the process development, a technology transfer of known parameters and conditions from 2-dimensional to zonal polishing is used. This first approach can be used for corrective polishing of zonal form deviations of ceramic samples in the future.

Keywords:

Polishing; Manufacturing; Transfer; Ceramic

1 INTRODUCTION

Advanced ceramics feature high hardness, remarkable corrosion resistance, low thermal expansion, high thermal shock resistance and low weight. These material properties allow their usage in adverse environments and applications. Popular examples are silicon nitride balls in bearing components and silicon carbide mirrors for astronomical optics devices. Furthermore, these non-oxide ceramics can be used by mold manufacturers for molding complex optics components out of glass. In order to achieve the low roughness requirements of the optical systems, a defect free surface and an accurate form is indispensable.

In spite of diverse scientific investigations and developments regarding machines for computer controlled polishing [1, 2, 3] the variety of robust process strategies is still insufficient. This deficit is due to the high amount of time and effort needed for the design of robust strategies, especially for small samples and complex geometries up to free forms.

In this paper, results of investigations of the transfer of polishing strategies (parameters and polishing system) from a conventional process (referred to as 2-dimensional) to a computer controlled process for zonal polishing are presented. The main objective is an efficient design of polishing strategies and a successful technology transfer in order to determine stable process conditions and high material removal rates (MRR) for the zonal correction of complex geometries.

2 EXPERIMENTAL SETUP

2.1 Materials

The transfer of process strategies from "2-dimensional to zonal" is exemplified by processing of silicon nitride (hot isostatically pressed silicon nitride) as one of the most popular non-oxide ceramics. The specimen were ground by using diamond grinding wheels (2-dimensional) and polished for the investigations of influence functions (zonal), respectively. The polishing systems were specified by using polyurethane foils as tools and ceria slurry with a deionized water base.

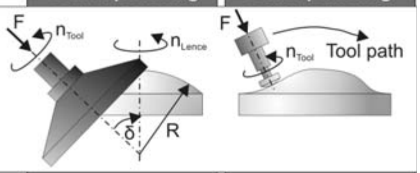
The efficiency of the polishing systems were evaluated by using different machining parameters in respect of process stability and reproducibility while maintaining high MRR and surface quality.

2.2 Methods

In comparison to zonal polishing (figure 1) the 2- dimensional process has a number of advantages in terms of process development and scientific investigations, such as high MRR, clearly measurable or even visible effects with respect to technological interactions and straightforward use of metrology of the processing results. Hence, the scientific understanding of technological aspects can be elaborated with reasonable efforts and used to develop polishing strategies for commonly used materials.

However, this specific process is not capable of polishing free form surfaces which becomes more and more important in a variety of innovative applications. To overcome these limitations, the zonal polishing process can be used. Unfortunately, an apparent lack of understanding and the intricate procedure of investigations hinder the holistic development of polishing strategies for the zonal process.

Therefore, the approach of this paper is to use the polishing strategies developed with an established 2-dimensional polishing process and transfer these to a zonal polishing machine tool. Based on theoretical considerations the optimization of the zonal process include practical investigations concerning the variation of pressure and relative speed as well as specific zonal process parameters such as eccentric frequency and radius on the formation and stability of the influence function. These preliminary investigations are of vital importance for establishing a stable correction process of zonal form deviations on complex geometries based on an adapted polishing system as former investigations on the polishing of steel already proved [4].



	2-dim. polishing	Zonal polishing
Process kinematic		
Material removal rate	High	Low
Measurement of MRR	Easy	Difficult
Process time	Fast	Slow
Process stability	Mostly state-of-the -art	High research efforts
Metrology	Easy metrology	Limited metrology
Form correction	Limited	Dwell time polishing
Geometry complexity	Low (planar, sperical)	High (asperical)
Conclusion	High efficiency in scientific research of technological interactions	Decisive for every complex component with optical surface quality

Figure 1: Advantages and disadvantages of different polishing processes at a glance.

2.3 Machine

The studies for the development of an efficient polishing system were carried out on a CNC-controlled polishing machine (2-dimensional). The subsequent investigations took place on an adaptable 5-axes polishing module (figure 2a) which was developed at the Fraunhofer IPT [4, 5]. The setup and its related control system for synchronizing the module to a 3-axes machine provides the required flexibility for investigations of polishing complex geometries (figure 2a).

Due to the double V-kinematic structure featuring five degrees of freedom the module is capable of adjusting its alignment to continuously ensure a perpendicular angle of contact between tool and surface of the workpiece. Furthermore, an individual eccentric movement can be superimposed to the rotating polishing tool at different angles (figure 2b), decisively influencing the characteristic removal profile of the tool. The polishing module is capable of different eccentric movements specified by radius and frequency (table 1). This specific tool movement can be superimposed to the rotation of the spindle (figure 2b). Thus, heterodyne velocity profiles can be realized. In addition to the position-controlled mode of the z-axis, the system can change to force-controlled mode in order to permanently assure a steady allocation of pressure in the process zone (table 1).

Table 1: Parameter range for the DoE.

Polishing force F	0.5 - 20 N
Force increment	0.2 N
Eccentric frequency f_{ecc}	1 - 10 Hz
Eccentric radius r_{ecc}	0.2 - 4 mm
Spindle revolutions n	500 - 5000 min^{-1}

In order to realize a corrective zonal polishing process with this machine tool setup a dwell-time based algorithm is used prior to the actual polishing task. As input data for this algorithm, the supposed geometry of the surface is needed. In addition, a form measurement taken either by optical or tactile metrology is used to determine the local error map of the sample. The third required input is the influence function of the zonal polishing tool, describing the material removal profile dependent on the adjustable process parameters such as pressure and relative speed. The subsequent calculation of the algorithm results in a dwell-time controlled path of the polishing tool over the sample and particular process parameters of the polishing process. In order to allow correction of form deviations on complex surfaces a stable, reproducible and deterministically adjustable influence function is of great importance.

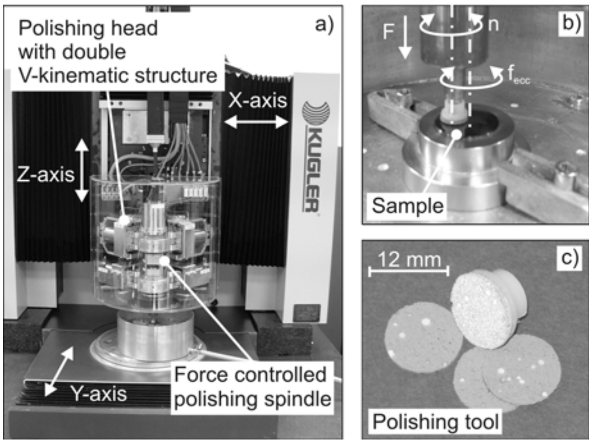


Figure 2: a) Machine tool setup used for investigations of the zonal polishing process, b) closer view on the polishing tool, c) polishing pad with polyurethane foil.

The applied polishing tools (figure 2c) consist of a flat plastic base body with a diameter of 12 mm, covered with different polishing foils. The polishing head allows the usage of further polishing tools like rubber based body materials or even tools with an bulb formed membrane, which will be taken into account in further scientific work.

2.4 Metrology

The MRR for developing efficient polishing strategies are quantified by weight measurements of the samples (2-dimensional process) and mathematically estimated by measuring the geometry of the influence function (zonal process), respectively. The latter is realized by using a Form TalySurf (tactile) measurement device. The surface quality was determined by a white light interferometer.

3 RESULTS AND DISCUSSION

3.1 Removal mechanisms

Former investigations [6] reveal satisfying results using ceria and diamond based slurries in a conventional polishing process (2-dimensional). Material removal rates of up 0.3 µm/min can be achieved. The surface qualities of both polishing systems are highly reproducible and yield optical surfaces in the range of 1-2 nm Ra and 20-30 nm Rt. However, ceria slurry tends to result in a more efficient polishing step than diamond slurry due to less machining efforts, i.e. same applied pressure though less revolutions, and certainly less monetary effort.

It is assumed that the advantage of polishing with ceria slurry is based on an interaction of chemical and mechanical removal mechanisms as outlined in figure 3. The beneficial process conditions can be explained by hydrolysis of the silicon nitride surface in aqueous solutions in combination with the so called "chemical tooth" ability of ceria [3, 6, 7, 8]. An indicator of the oxidation of silicon nitride by forming sodium hydroxide is shown by an increase of pH (figure 3).

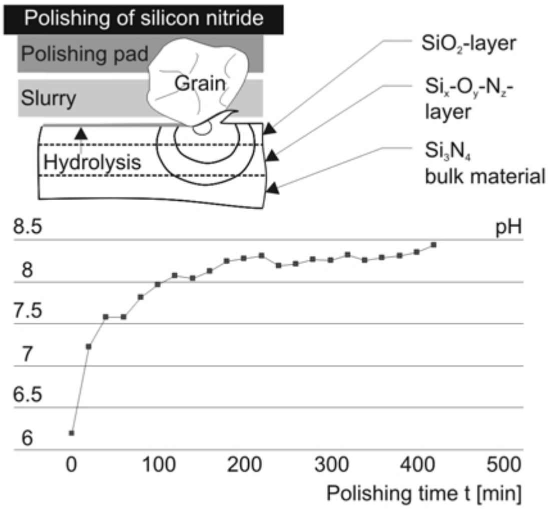


Figure 3: Assumed removal mechanisms of polishing silicon nitride with ceria slurry (2-dimensional process).

3.2 Tool influence function by zonal polishing

The previously identified polishing strategies which determine the process conditions and hence the result of the finishing are used for first experiments regarding the influence function of the zonal process. Taking the Preston Hypothesis for granted (formula 1), a linear increase of the MRR should be realized starting from the center of the rotating tool. Figure 4a shows the accruing characteristic W-profile by the adjusted process parameters pressure p and relative velocity v_r which are based on theoretical conversion of former specified values. The Preston coefficient K_P represents various other influences on the process.

$$MRR = K_P \cdot p \cdot v_r \quad (1)$$

However, the maximum material removal is not found on the outer diameter but at one half to two thirds of the radius of the polishing tool, forming the influence function. Looking further towards the boundary of the influence function, the removal depth declines steadily. This simple trial emphasizes the fact that either an inhomogeneous pressure profile or an influence of the relative velocity in the gap between the tool and the sample is influencing the geometrical formation.

The influence function is characterized by depth d_{IF} and radius R_{IF} which in this case can be identified as 1.6 mm. Preliminary investigations on polishing of steel have shown that this dimension can be used for setting up the eccentricity of the polishing tool movement in order to achieve a Gaussian profile of the influence function [2]. Further investigations lead to a ratio of R_{IF} and eccentric radius r_{ecc} of about 1:2. By using this value and an experimentally identified value for the eccentric frequency f_{ecc} of 4 Hz the required Gaussian profile can be achieved as shown in figure 4b. The appropriate volume of removed material V_{IF} was computed by importing the 2D-graph of a Form TalySurf in a specially developed Matlab program and amounts to 0.005 mm³ on average (after polishing time of 240 s).

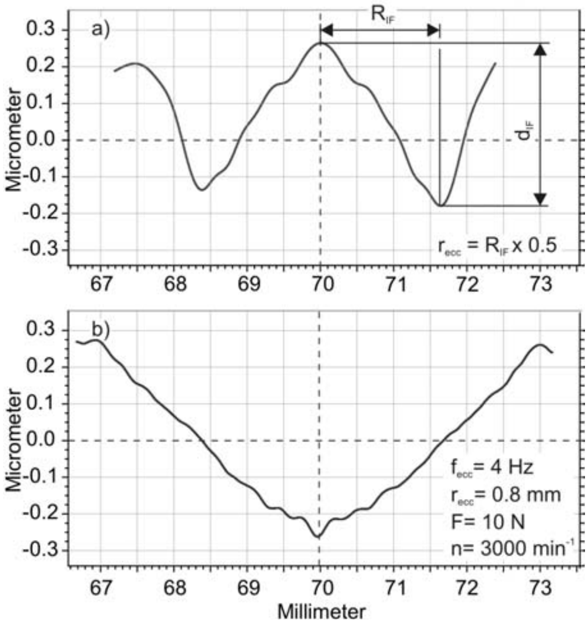


Figure 4: a) Determination of eccentric radius as the main parameter for achieving the required (b) Gaussian profile of the influence function.

3.3 Investigation of typical process parameters

As the process parameters lead to reproducible Gaussian profiles of the influence function the time related stability and corresponding MRR are verified. Figure 5 shows a non-linearity between the removed material V_{IF} after 120 s and 240 s, which can be explained by incomplete saturation of the polishing foil with ceria slurry after a process duration of 120 s. Therefore, a preparation of the polishing tool before starting the intrinsic polishing process was established. A linear increase of the material removal with increasing dwell-time (figure 5) is a main requirement for a form correction.

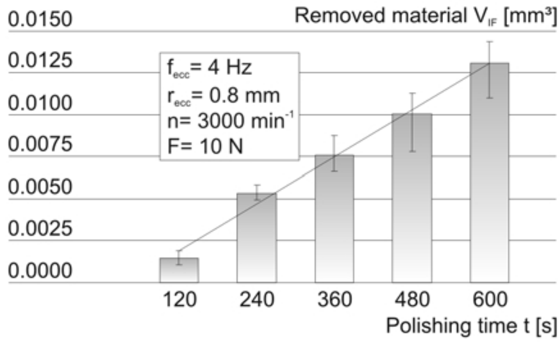


Figure 5: Evaluation of process stability by different processing durations and volume of removed material.

The overall process stability was proven by a long-term run of the polishing tool under real process conditions and can be estimated by up to 3 hours. After that duration the influence function matches the one of the beginning of the process.

Seeking a further understanding, a variation of the tool revolutions over constant applied pressure and vice versa have been conducted. Unlike the results of investigations of the steel polishing process [4], the influence of changes of relative velocity on MRR is far more sensitive. The increase of revolutions strictly involves the increase of the MRR (figure 6a). Furthermore, figure 6 shows that an increase of normal force and applied pressure, respectively, leads to rising MRR (figure 6b). The behavior is approximately linear.

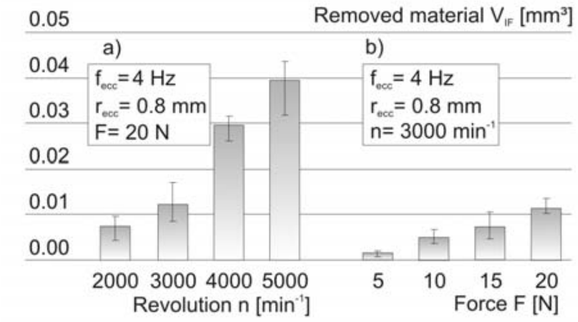


Figure 6: Dependency of material removal on parameters (a) relative velocity v_r and (b) force F (Preston hypothesis).

3.4 Enhancement of process efficiency

Due to its high stability, e.g. compared to felt as another popular polishing tool, polyurethane foil revealed the potential to further increase process normal forces. However, the stability of the process has to be ensured first.

The influence function in figure 4b changes from Gaussian to W-shaped profile with the increase of pressure. Moreover, the radius R_{IF} is increasing as shown in figure 4a and figure 7a. This phenomena can be explained by a growing deformation of the polishing tool. An adjustment of the eccentric radius r_{ecc} to 80% of the R_{IF} allowed the realization of reproducible influence functions with Gaussian profiles.

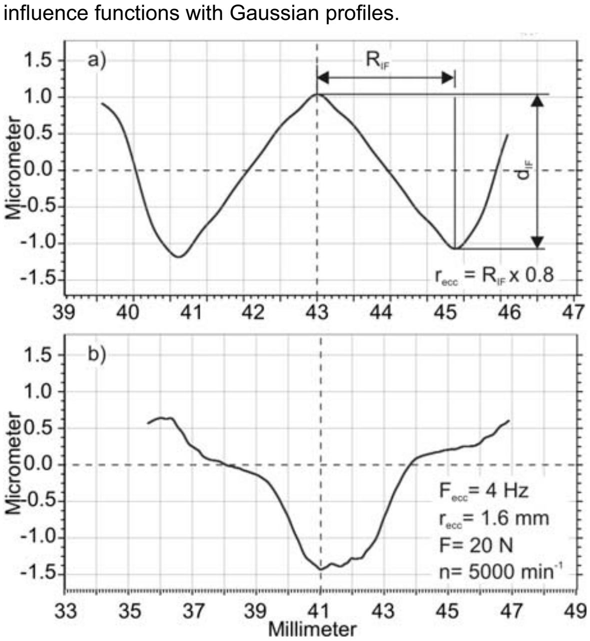


Figure 7: Influence function (a) before and (b) after adjusting specific corrective values for maximized removal rate.

It is assumed that the linearity of the MRR (figure 6b) and applied normal force allows the doubling of the MRR with the zonal process. Due to the different contact area the pressure of 2 bar equals roughly 360 N (2-dimensional process) and 10 N (zonal process), respectively. Hence, the use of the maximum of 20 N (table 1) would lead to doubled MRR for the zonal polishing compared to the 2-dimensional. A verification with specific parameters (figure 7b) results in a maximized MRR of $0.62 \cdot 10^{-3}$ mm^3/min (normalized to 1 mm^2) which indeed marks the double of the normalized MRR of $0.31 \cdot 10^{-3}$ mm^3/min for the 2-dimensional process.

4 CONCLUSION

This paper showed the results of scientific studies with the goal of shortening the development of a stable and reproducible polishing process for corrective machining of complex ceramics samples. The transfer was successful by rendering unnecessary intensive preparatory work on polishing strategies and resulted in satisfying overall results with a high and reproducible MRR and surface quality. In terms of the machine parameters, the eccentric frequency shows only a small impact for the profile formation. However, the eccentric radius, the applied normal force and the relative velocity do have significant impact on the formation of the influence function determining the process efficiency. These investigations establish the basis for a correction of form deviations on complex geometries of ceramic samples, which is the objective of ongoing scientific work and will be presented in future publications.

5 ACKNOWLEDGMENTS

We sincerely thank the German Research Foundation (DFG) for financial support of the Collaborative Research Center SFB/TR4 (RWTH Aachen and University Bremen (Germany), Oklahoma State University at Stillwater (USA)).

6 REFERENCES

[1] Bingham, R.G., Walker, D.D., Kim, D.-H., Brooks, D., Freemann, R., Riley, D., 2000, A Novel Automated Process for Aspheric Surfaces, Proc. SPIE, 4093/10:445-450

[2] Yi, A., Hezlep, M., Pol, T., 2004, A Computer Controlled Optical Pin Polishing Machine, J. Materials Processing Technology 146/2:156-162.

[3] Evans, C.J., Paul, E., Dornfeld, D., Lucca, D.A., Byrne, G., Tricard, M., Klocke, F., Dambon, O., Mullany, B.A., 2003, Material Removal Mechanisms in Lapping and Polishing, CIRP Annals, 52/2:611–633.

[4] Brecher, C., Wenzel, C., 2006, Kinematic Influences on the Formation of the Footprint During Local Polishing of Steel, WGP Annals, XIII/1:23-26.

[5] Weck, M., Wenzel, C., 2004, Adaptable 5-axes Polishing Machine-head, Production Engineering, 11/1:95-98.

[6] Klocke, F., Zunke, R., Dambon, O., 2007, Chemo-mechanical Polishing of Silicon Nitride, Proc. euSPEN, 2:356-359.

[7] Jiang, M., Komanduri, R., 1998, On the finishing of Si_3N_4 Balls for Bearing Applications, Wear, 215/1:267-278.

[8] Hu, Y.Z., Gutmann, R.J., Chow, T.P., 1998, Silicon Nitride Chemical-mechanical Polishing Mechanisms, J. Electrochem. Soc., 145/11: 3919-3925.

Sub-wavelength Pitched Cubic Mosaic Multi-layer Precisely Pressed by Nano-features Mold for Multi-functional Optical Elements

M. Nakao (2)[1], K. Nagato[1], H. Suzuki[1], T. Nishino[1], H. Yonemoto[1],
H. Kaito[1], T. Hamaguchi[1] and K. Tsuchiya[2]
[1] Dept. of Engineering Synthesis, School of Engineering, The University of Tokyo, Tokyo, Japan
[2] Institute of Industrial Science, The University of Tokyo, Tokyo, Japan

Abstract
The new manufacturing process for the sub-wavelength, 200-400 nm pitched cubic mosaic multi-layer structure was designed. The two materials with the high/low refractive index form the mosaic structure for multi-functional optical elements such as polarizers or color separators. The structure could be precisely pressed by a nano-features mold with the convex mosaic pattern. This paper proposed two processes even at early development stage: an all-at-once press of the mold on the multi-layered sheet at several meters per second, and a repetitious press on the about 1 μm thick one-layered sheets, providing the other material sheet alternately.

Keywords:
Press, Optical Element, Nano-feature

1 INTRODUCTION

The authors aim to form sub-micron sized structures on a square meter wide substrate with a mechanical press, not with a semiconductor process using deposit, lithography and etch. The mechanical press has an advantage for quicker reproduction than the semiconductor process, stamping some nano-features on the meter-wide substrate. [1] [2] [3]

The sub-micron sized structure, for example, is a sub-wavelength, several hundreds nm pitch structure. The structure functions some optical characters. The authors had already designed several optical elements as described in Figure 1: (i) a color separator on the light guide plate, which inserts a white light into its side surface and emits a separated single color light normally from its upper surface, which looks a diffraction grating with the 200-400 nm pitch; (ii) a polarizer for a transmission light, also looking a grating with metal wires; (iii) a multi-functional optical element such as color separator, polarizer or light accumulator, constructing a cubic mosaic multi-layer structure using two materials with a high or low refractive index.

The sub-wavelength, 200-400 nm pitch structure will be useful for a liquid crystal display (LCD). The structure can be reproduced on the light guide plastic plate or the liquid crystal glass substrate; the structure will replace some conventional, thick optical sheets for normal emission, polarizer, reflector, brightness accumulator or view angle adjuster.

The authors had prototyped the structures as (i) the color separator and (ii) the polarizer above mentioned in Figure 1.

Figure 2 shows the results of the preliminary design of (i) the color separator; (a) the SEM view of the grating surface with a 400 nm pitch. The grooves were reproduced by hot-pressing a glass substrate with a silicon carbide mold; (b) the simulation result on the cross section. The light is inserted into the side, emitted from the upper surface; (c) the color separated view to red, green and blue emitted from the upper surface. Each color light was emitted from each pitch grating surface as designed.

Figure 3 shows the results of the preliminary design of (ii) the polarizer; (a) the SEM views of the grating surface and cross section with 300 nm pitch gold stripes; (b) the process illustration of hot-press on a glass substrate with the silicon carbide mold to reproduce the grooves with the 300 nm pitch, and damascene of gold stripes on the grooves [4]; (c) the evaluation of polarized ratio of the polarizer. The measured polarized ratio of the fabricated gold grating (blue line) resembles the simulated polarized ratio of a trapezoid cross section of the gold stripes (red), which is more similar to the measured one than that of a rectangle cross section (green). The grooves of the mold had the trapezoid cross section because the side walls of the grooves on the silicon carbide mold were etched obliquely.

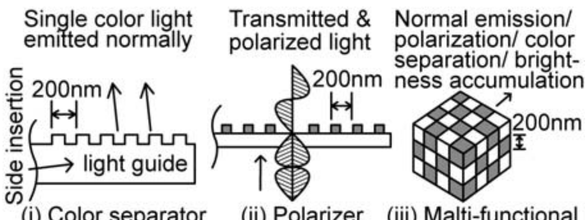
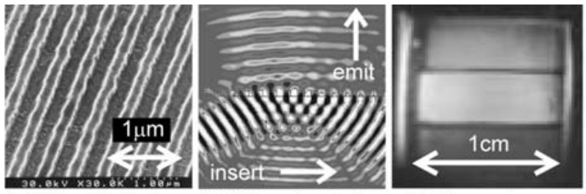


Figure 1: Optical elements with sub-wavelength pitched structures.



(a) The SEM view of the grating surface with a 400nm pitch (b) The simulation result on the cross section (c) The color separation to red, green and blue emitted from the upper surface

Figure 2: Results of the preliminary design of the color separator.

These results concluded that the mechanical press could precisely form the sub-wavelength, 200-400 nm pitched structure.

This paper challenges to prototype (iii) the multi-functional optical element described in Figure 1 using the mechanical press. The element has the cubic mosaic multi-layer structure with the sub-wavelength pitch.

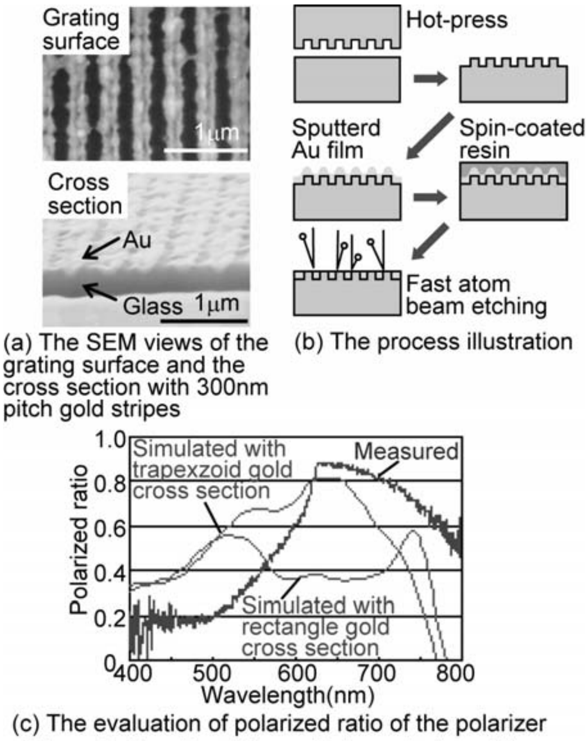


Figure 3: Results of the preliminary design of the polarizer.

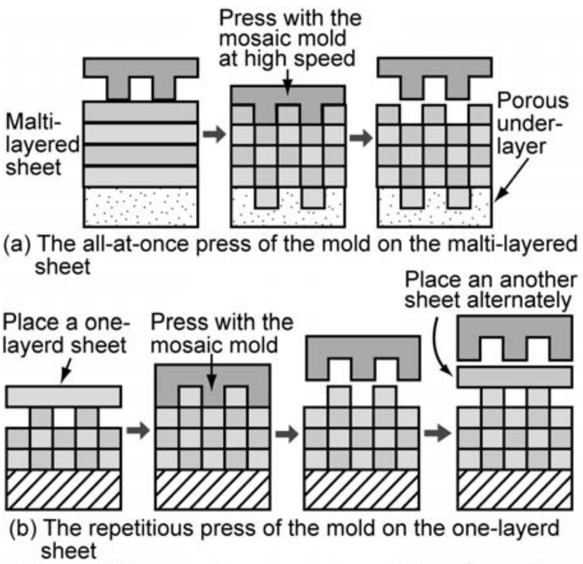


Figure 4: Proposed processes to prototype the cubic mosaic malti-layer structure.

2 DESIGN OF MECHANICAL PRESS PROCESS FOR THE SUB-WAVELENGTH PITCHED CUBIC MOSAIC STRUCTURE

The authors propose two novel processes for prototyping the cubic mosaic multi-layer structure: (a) an all-at-once press of the mold on the multi-layered sheet and (b) a repetitious press of it to the one-layered sheet, providing the different material sheet alternately.

Figure 4 illustrates schematics of the two processes.

e all-at-once press of Figure 4(a) is executed on the multi-layered sheet. The structure should provide two materials with a high (H) or low (L) refractive index; so the multi-layered sheet alternately stacked the films like H, L, H, L, H, L perpendicularly. A convex mosaic patterned mold shears the films only with a one-layer thick displacement, so arranging the mosaic like H, L, H, L, H, L horizontally and perpendicularly. The sharp shear tends to happen at high speed press, in the brittle films, and on the porous under-layer. Using the porous under-layer, the convex mold could straightly step down to the substrate, making the normal side wall of the concave groove as if the shoe steps down to the snow.

The repetitious press of Figure 4(b) is executed on the one-layered sheet. At first an H material sheet is pressed by the convex mosaic patterned mold, and next an L sheet is placed on the pressed H sheet. Then the L sheet should be pressed by the same mold without any horizontal position error, on the H sheet without any deformation of the formerly reproduced structure. Finally, the stacked sheets like H, L, H, L, H, L looks a cubic mosaic multi-layer structure. When the L material is an air, only the H plastic sheet will be pressed and stacked as explained in Figure 7 later. The sub-micron thick sheet might be produced by sheet expanding, spin coating, or vapor depositing. Through their prototyping, we empirically analyze the processes, disassembling them into the several elemental requirement functions.

(a) The all-at-once press requires: (a1) to shear the sheet only with a one-layer displacement, (a2) to press the sheet at high speed for normal shearing, (a3) to set the porous under-layer for normal straight shearing, and (a4) to make a lot of normal shear lines through the films with a sub-wavelength pitch.

(b) The repetitious press requires: (b1) to reproduce many grooves/ridges/dots with a sub-wavelength pitch, (b2) to place an about 1 μm thick sheet, (b3) to reproduce grooves only on an uppermost film, and (b4) to repeatedly press the mold on the next sheet at the same position.

3 PROTOTYPE OF THE SUB-WAVELENGTH PITCH CUBIC MOSAIC MULTI-LAYER STRUCTURE

3.1 Development of the all-at-once press process

Figure 5 shows the prototyped structure at the cross section. Four layers of silicon and four of silicon oxide were alternately sputtered with a 250 nm thickness each. Porous aluminum under-layer was deposited with a 5 μm thickness on the steel plate. Its density was about 70% of the bulk. The used mosaic patterned mold etched by FIB on the steel plate, however, had a large 10 μm pitch to reduce the plastic deformation. Through results in Figure 5, the three functional requirements were satisfied: (a1) the films was sheared only with a one-layer displacement, 250 nm (Figure 5(a)) due to the 300 nm deep convex mold; (a2) the sheet was pressed at high speed for normal shearing. Free fall made a 0.9 meter/second velocity as illustrated in Figure 6(a)(c), representing about 2000 strain velocity, when we defined strain=1 for shared at 45 degrees, and divided the strain by a contact time (0.6 msec) measured by electrical resistance (Figure 6(c)); (a3) the porous under-layer for normal straight shearing was set, shearing at 80-90 degrees to the surface(Figure 5(a)(b)).

Next the finer convex mosaic patterned mold with a 1 μm pitch was prepared as shown in Figure 7(a). It was a Ni plated mold patterned by electron lithography. The sheet had a 70 nm thick silicon oxide sputtered on a 10 μm thick gold under-layer. Figure 7(b) shows the SEM view of the surface and the cross section of the reproduced layers. The silicon oxide film, though it was difficult to see by SEM and partly peeled over, was reproduced to the mosaic pattern at 6 meters/second press velocity. This lighter mold (18 gram) could be pressed by the spring at a 14 m/sec at maximum as illustrated in Figure 6(b). The displacement and contact time could be adjusted by the damping factor of the die cushion.

The Ni plated mold (Hv=200) was not enough hard to press a multi-layered sheet because the 1 μm pitch pattern generated larger plastic deformation. Future work for a harder annealed Ni or silicon carbide mold should satisfy the last requirement, (a4) to make a lot of normal shear lines through the films with a sub-wavelength pitch.

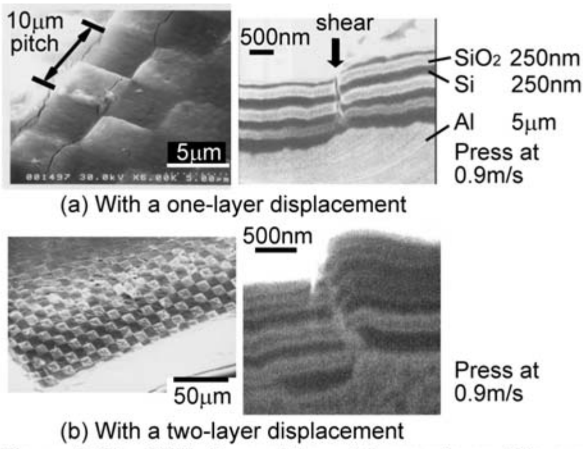


Figure 5: The SEM views of the cubic mosaic multi-layer structure reproduced by the all-at-once press as shown in Figure 4(a).

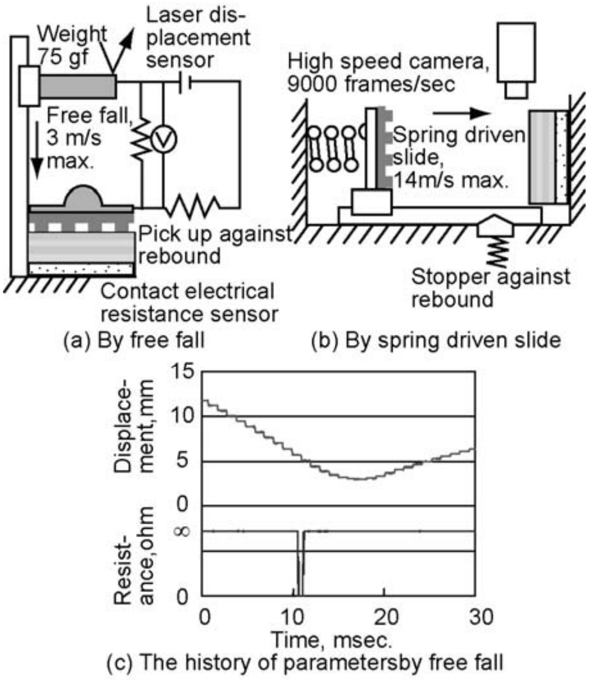


Figure 6: Processes of high speed press for the all-at-once press.

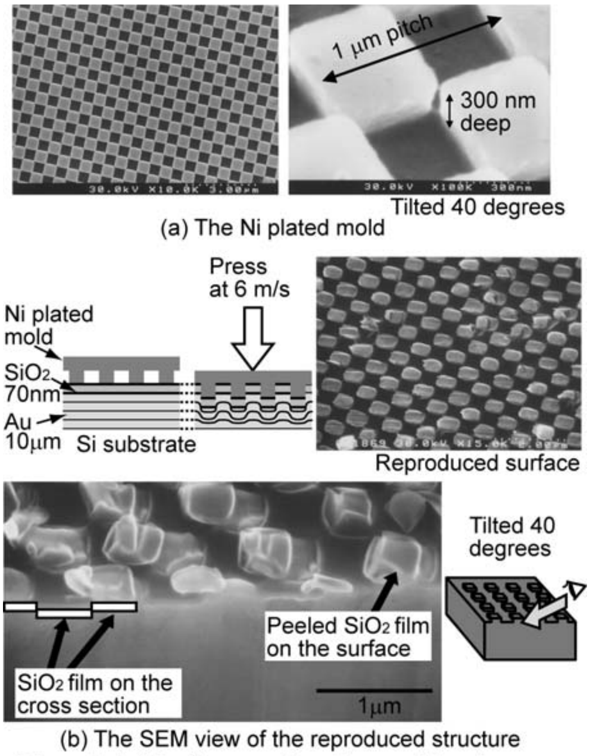


Figure 7: (a)The fine mosaic pattern mold and (b)Reproduced cubic mosaic structure using the all-at-once press.

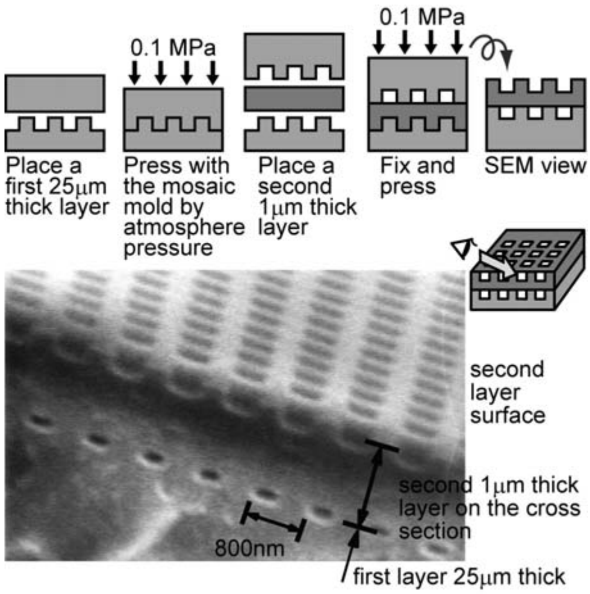


Figure 8: The SEM view of the reproduced cross mosaic structure using the repetitious press.

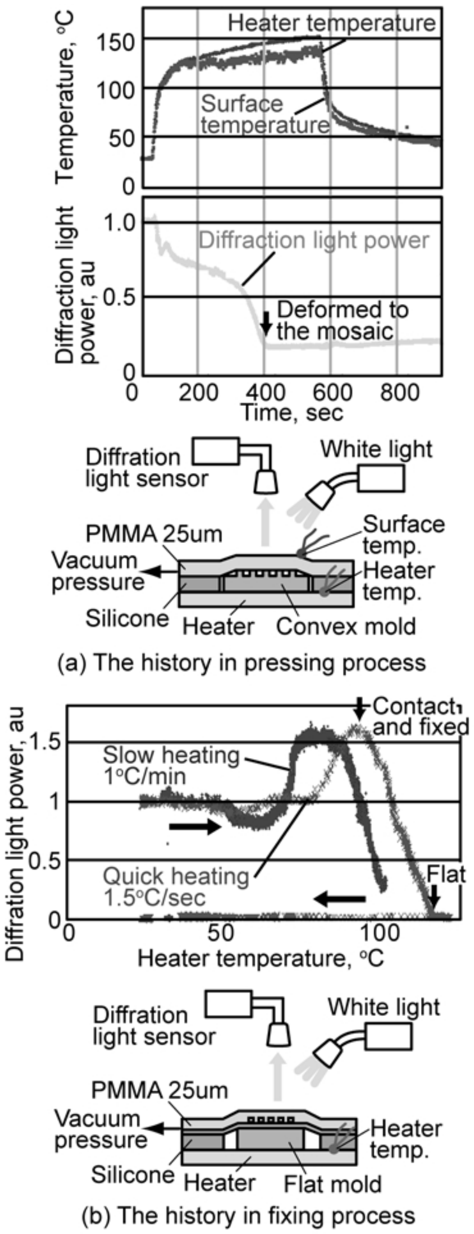


Figure 9: The history of process parameters during the repetitous press.

3.2 Development of the repetitious press

Figure 8 shows the prototyped structure at the cross section. The H material (n=1.5) is PMMA; the L (n=1) is the air. The three requirements were satisfied. (b1) to reproduce many grooves with a sub-wavelength pitch, 800 nm was realized; (b2) to place a sub-wavelength thick sheet was also realized, making a 1000 nm thick film by spin-coating and peeling from the flat plate; (b3) to reproduce grooves in second pressing only on an uppermost film was satisfied, representing that the mosaic grooves of the first layer still reminded as in Figure 8.

Figure 9(a) is the history of temperature of the sheet and diffractive light power in hot-pressing process to a 25 μm thick PMMA sheet as shown in Figure 8. Plastic deformation into the grooves changes the diffractive light power. At 400 sec., the grooves were reproduced. The gap between the mold and the PMMA sheet was vacuumed; press pressure was 1 atmosphere, 0.1 MPa.

Figure 9(b) is the history in next fixing of a 1 μm thick sheet using the flat mold on the 25 μm thick reproduced sheet. Faster heating could increase the fixing temperature by 10 °C, getting higher fixing strength. The reproduced structure in Figure 8 was also fixed in second hot-pressing with the mosaic mold at 100 °C under fast heating of 1 °C/sec. But two patterns had with a 100 nm positioning error. Future development of a mechanism to adjust the mold position is necessary to realize the last functional requirement, (b4) to repeatedly press the mold on the next sheet at the same position.

4 DISCUSSION

The section 3 confirmed that the mechanical press could form the cubic mosaic structure. Especially, the repetitious press is promising because it can be realized under low pressure. The multi-functional optical element for LCD, however, needs the new press process for a meter-wide substrate. The future reproduction will employ a rolling cylinder mold or a sliding broach mold instead of the 10 cm-wide punch mold. Just after reproduction, quickly solidifying by air flow cooling or UV light hardening will be effective to hold the groove shape.

5 CONCLUSION

The mechanical press process for the sub-wavelength, 200-400 nm pitched cubic mosaic multi-layer structure is proposed. The structure could be precisely pressed by a nano-features mold with the convex mosaic pattern. The structure was made of the two materials with the high/low refractive index, for example, silicon and silicon oxide, or PMMA and the air. The prototyped processes even at early development stage were reported: an all-at-once press of the mold on the 8 layered sheet at several meters per second, and a repetitious press on a 25 μm thick one-layered sheets, providing the other 1 μm thick sheet alternately. Both processes could form the 800-1000 nm pitched cubic mosaic structure as designed, and will produce multi-functional optical elements such as polarizers or color separators.

6 ACKNOWLEDGEMENT

This research was supported by a Grant-in-Aid for Scientific Research (S), 19106003 from the Ministry of Education, Culture, Sports, Science and Technology, Japan.

7 REFERENCE

[1] S. Y. Chou, P. R. Krauss, 1997, Imprint Lithography with Sub-10 nm Feature Size and High Throughput, Microelectronic Engineering, 35, 237-240.

[2] L. R. Bao, X. Cheng, X. D. Huang, L. J. Guo, S. W. Pang, and A. F. Yee, 2000, Nanoimprinting over topography and multilayer three-dimensional printing, J. Vac. Science Technology, 20/6, 2881-2886.

[3] M. Nakajima, T. Yoshikawa, K. Sogo, and Y. Hirai, 2006, Fabrication of multi-layered nano-channels by reversal imprint lithography, Microelectronic Engineering, 83/4-9, 876-879.

[4] M. Nakao, J. Kumaki, K. Matsumoto and Y. Hatamura, 2004, Multi-layered circuit board precisely pressed/damascened on glass plates, Precision Engineering, 28, 181-185.

Grinding

Simulation of Workpiece Kinematics in Centreless Throughfeed Grinding

Peter Krajnik[1], Radovan Drazumeric[1], Ivan Vengust[1], Rok Vrabic[1], Bernd Meyer[2], Janez Kopac[1], Fritz Klocke[2]

[1] Faculty of Mechanical Engineering, University of Ljubljana, Slovenia

[2] Laboratory for Machine Tools and Production Engineering WZL, RWTH Aachen University, Germany

Abstract

This paper deals with the simulation of the workpiece kinematics in centreless throughfeed grinding. The objective of this research was to develop a simulation tool that can be used to create an interactive virtual environment, to place the grinding gap elements in the defined set-up and to visualise process kinematics. The simulation is founded on an analytical 3D model, which includes a parametrical description of all grinding gap elements and their kinematics. The simulation software was programmed in C#. Its platform consists of Windows OS, .NET framework and OpenGL graphics library.

Keywords:

Centreless grinding; Modelling; Simulation

1 INTRODUCTION

Centreless throughfeed grinding is commonly employed for the production of rotationally symmetrical parts, particularly in automotive industry. It is a process that often suffers from inaccuracy problems due to the centerline of the workpiece not being fixed. The cause of these problems is usually due to incorrect set-up of the grinding gap, too low workpiece speed, and improperly dressed regulating wheel [1,2]. On the shop floor, where different part geometries have to be ground with the same machine tool, the employed parameters do not always yield desired results. This requires adjustments that are based on trial-and-error, which lead to a decrease in the machine tool usability of up to 50%. The goal of the research is to provide the possibility of performing adjustments that are independent of an operator, thus significantly reducing the set-up time and in turn increasing productivity. Productivity has been examined across several key leverage points such as technology deployment and process efficiency. A shop floor productivity diagnostics revealed that apart from the rounding mechanism [3], one of the main problems in centreless throughfeed grinding is related to variation in workpiece speed along the grinding gap. In order to attain a steady feedrate along the grinding gap the workpiece has to be in simultaneous line contact with the grinding wheel, the regulating wheel and the workrest blade. This yields a noncylindrical shape of the inclined regulating wheel that leads to variations in axial and rotational workpiece speed, especially when grinding long workpieces. This can lead to undesired process disturbances and vibrations of the workpiece. The simulation tool incorporates an analytical 3D model of the grinding gap that enables visualisation of the workpiece movement. Based on this, the exact grinding time and specific material removal rate can be calculated for any point along the grinding gap. The software is programmed in C#, using the .NET framework and the OpenGL 3D graphics library, which is used because the simulation tool is to be ported to the CAVE virtual reality system in near future [4].

2 GEOMETRY OF THE GRINDING GAP

In order to simulate workpiece kinematics the geometry of the grinding gap has to be analytically described. For this purpose several basic demands for the analytical display of the grinding gap have to be defined:

1. simultaneous line contact of the workpiece with the grinding wheel, regulating wheel and the workrest blade;
2. cylindrical grinding wheel;
3. noninclined workrest blade;
4. regulating wheel is inclined and swivelled in its centre.

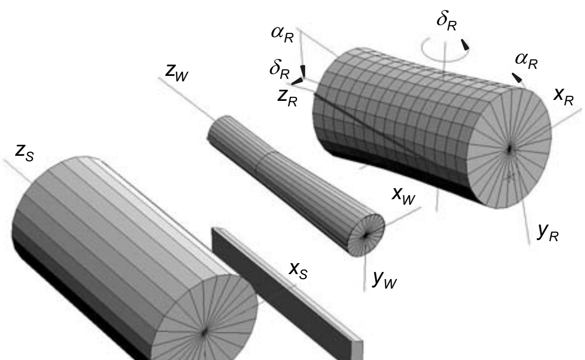


Figure 1: Grinding gap and local coordinate systems.

The simultaneous line contact of the workpiece with the grinding wheel, regulating wheel and the workrest blade enables steady workpiece kinematics along the grinding gap. In this way a contact loss between the workpiece and the regulating wheel can be avoided [5]. The simulation presumes the employment of a cBN grinding wheel that is hard to dress and is therefore practically cylindrical. In view of the machine tool design, the workrest blade is never inclined in the direction of the machine tool plane. The workpiece axial speed is obtained by the inclination of the regulating wheel. Most machine tools are designed so that the regulating wheel is inclined and swivelled in its centre, as shown in Figure 1. Analytical modelling of the grinding gap requires a parametrical description of the grinding wheel, the workpiece, the workrest blade and the regulating wheel. All four elements are described as two parametrical surfaces in a global coordinate system which lies in the front plane of the workrest blade, shown in Figure 2.

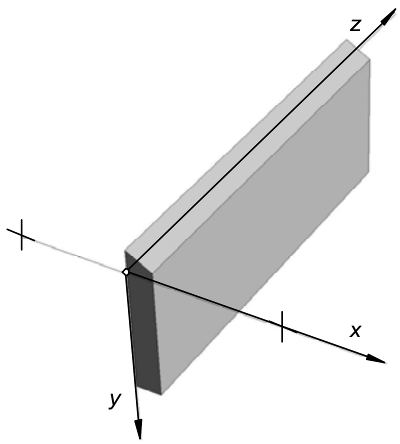


Figure 2: Global coordinate system.

For the grinding gap parameterisation, a definition of geometrical parameters is required. Some basic parameters that are included in the models are shown in Figure 3. Additional parameters are further described in subsections below.

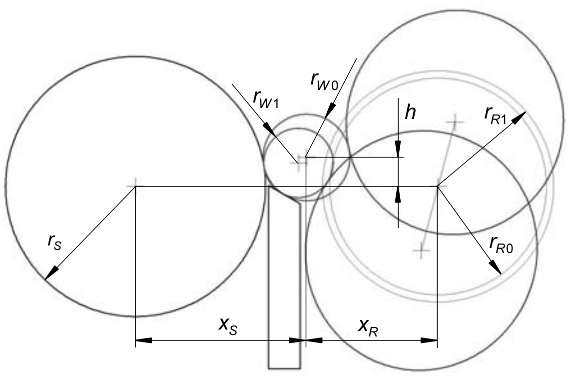


Figure 3: Basic geometrical parameters of a grinding gap.

2.1 Grinding wheel

The geometry of a cylindrical grinding wheel is determined by its radius r_S and width b_S. Its analytical model in the global coordinate system gives:

$$\begin{aligned} x &= r_S(\cos u - 1) - x_A \\ y &= r_S \sin u \\ z &= v \end{aligned}, \quad \begin{aligned} 0 \le u \le 2\pi \\ 0 \le v \le b_S \end{aligned} \tag{1}$$

where parameter x_A denotes the distance between the grinding wheel and the workrest blade. This distance is later shown in Figure 6.

2.2 Workpiece

The geometry of a workpiece with length l_W, centre height h, input radius r_{W0} and output radius r_{W1} along the grinding gap is determined with the following model:

$$\begin{aligned} x &= x_S - r_S - x_A + x_W(v) + r_W(v)\cos u \\ y &= -h + y_W(v) + r_W(v)\sin u \\ z &= v \end{aligned} \tag{2}$$

in the intervals $0 \le u \le 2\pi$ and $0 \le v \le b_S$. The parameter x_S describes the horizontal distance between the grinding wheel centre and the workpiece centre and is calculated as:

$$x_S = \sqrt{(r_S + r_{W0})^2 - h^2} \tag{3}$$

The workpiece radius function, shown in Figure 4, is a major technological input parameter that determines the material removal rate along the grinding gap. The centreless grinding end users require inputting this function directly to the machine tool controller. For the simulation tool we introduced a polynomial workpiece radius function along the grinding gap $0 \le v \le b_S - b_{sa}$ that enables continuous material removal and undisturbed workpiece passage to the spark-out zone:

$$r_W(v) = r_{W1} + (r_{W0} - r_{W1})\frac{(v - b_S + b_{sa})^2 (v + 2b_S - 2b_{sa})}{2(b_S - b_{sa})^3} \tag{4}$$

where b_{sa} describes the length of a spark-out zone.

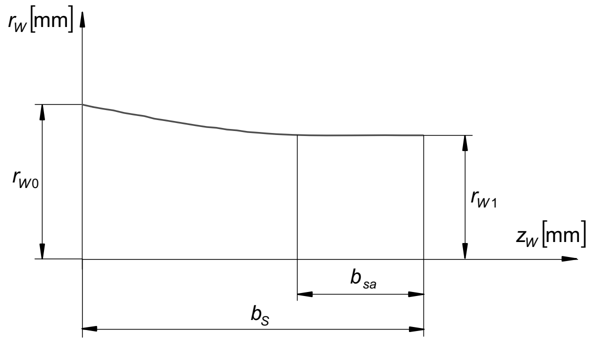


Figure 4: Workpiece radius function.

2.3 Workrest blade

The geometry of a workrest blade depends on its width b_A and angle β, shown in Figure 5. Its analytical model in the global coordinate system gives:

$$\begin{aligned} x &= u \\ y &= u \cdot \tan\beta - h_A \\ z &= v \end{aligned}, \quad \begin{aligned} 0 \le u \le b_A \\ 0 \le v \le b_S \end{aligned} \tag{5}$$

where parameter h_A is calculated as follows:

$$h_A = \left(\sqrt{(r_S + r_{W0})^2 - h^2} - r_S - x_A\right)\tan\beta + h - \frac{r_{W0}}{\cos\beta} \tag{6}$$

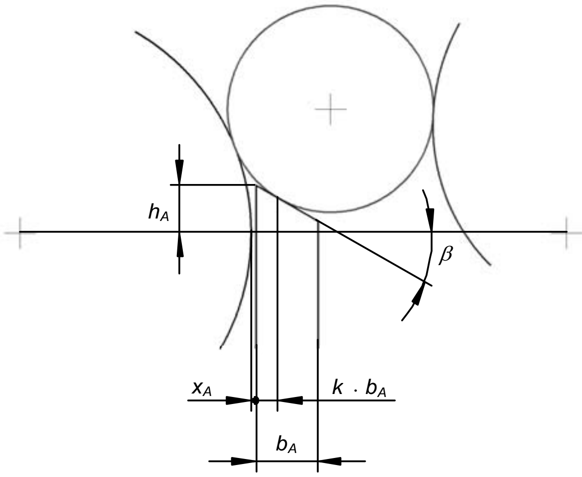


Figure 5: Geometrical parameters of a workrest blade.

The contact point between the workpiece and the workrest blade should lie within a certain interval on the support plane. The input workpiece position is controlled by parameter k. The size of the gap between the grinding wheel and the workrest blade x_A can also be entered as input. In order to prevent collision it is important to keep its value positive.

2.4 Regulating wheel

To achieve workpiece throughfeed, the regulating wheel has to be inclined for an appointed angle α_R. Because of this inclination, the contact point between the workpiece and the regulating wheel is displaced along the grinding gap. The regulating wheel form has to compensate this displacement in order to achieve the required line contact. If the regulating wheel form deviates from the optimal form, the contact gets interrupted and continuous throughfeed of the workpiece is not guaranteed. In order to avoid unsteady workpiece kinematics, the regulating wheel form has to be adapted to the already determined workpiece radius function.

The geometry of the regulating wheel depends on its width b_R, inclination angle α_R, input diameter d_{R0} and output diameter d_{R1}. The latter two values can be easily acquired from the machine tool controller. Its parameterisation gives:

$$\begin{aligned} x &= x_S + x_R - r_S - x_A - (v - z_0)\cos\alpha_R \sin\delta_R - \\ &\quad r_R(v)(\cos\delta_R \cos u + \sin\alpha_R \sin\delta_R \sin u) \\ y &= (v - z_0)\sin\alpha_R - r_R(v)\cos\alpha_R \sin u \\ z &= z_0 + (v - z_0)\cos\alpha_R \cos\delta_R - \\ &\quad r_R(v)(\sin\delta_R \cos u - \sin\alpha_R \cos\delta_R \sin u) \end{aligned} \tag{7}$$

The application of a system of two conditional equations of the contact enables us to calculate the regulating wheel radius r_R and the contact angle φ_R, shown in Figure 6, in every point of the grinding gap. The solution to these two equations further enables us to determine the distance between the regulating wheel centre, the workpiece centre x_R and the swivelling angle δ_R, which is an important parameter for the machine tool set-up. The discussed system of two equations, which describe the contact between the workpiece and the regulating wheel, is given in [6].

3 WORKPIECE KINEMATICS

In order to determine workpiece kinematics, a non-sliding contact with the regulating wheel in every point of the grinding gap is presupposed. The regulating wheel speed depends on its radius and revolutions. Since the regulating wheel rotation is constant, the workpiece speed is directly related to the calculated regulating wheel form.

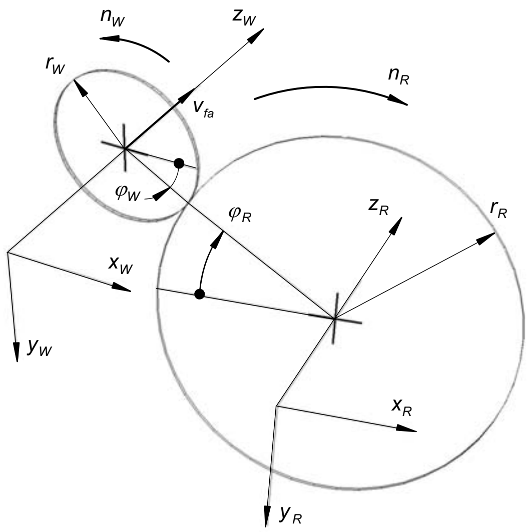


Figure 6: Kinematical parameters.

Taking into account the workpiece and the regulating wheel geometries we can determine the rotational component n_W and the axial component v_{fa} of a workpiece speed:

$$\begin{aligned} n_W = \frac{r_R}{r_W} n_R \big(&\cos\alpha_R \cos\varphi_R \cos\varphi_W + \\ &\cos\delta_R \sin\varphi_R \sin\varphi_W - \sin\alpha_R \sin\delta_R \cos\varphi_R \sin\varphi_W\big) \end{aligned} \tag{8}$$

$$v_{fa} = \frac{\pi \cdot r_R}{30} n_R (\sin\alpha_R \cos\delta_R \cos\varphi_R + \sin\delta_R \sin\varphi_R) \tag{9}$$

The workpiece speed in a contact is changing throughout the grinding gap, whereas a nonsliding contact is only possible in one point at a time. Therefore actual speeds are calculated as average values at a certain workpiece position. These values are prerequisites for the calculation of the grinding time and the specific material removal rate that are major estimates of process productivity [7].

4 SIMULATION TOOL

A software simulation tool was developed to support the underlying analytical model with visualisation. The grinding gap elements can be placed into the defined set-up. The process kinematics and workpiece forming are simulated in an interactive virtual environment.

4.1 Software platform

The simulation tool is currently developed for the Windows platform and is written in the C# programming language using the .NET framework and the OpenGL library. The .NET framework is used for the graphical user interface (GUI) and for certain operations that support the software features, such as parameter loading, saving and screenshot acquisition. The OpenGL library is used to visualise the geometry and kinematics. All of the technologies are widely used and allow the simulation tool to be ported to a variety of operating systems using the Mono platform, which is a runtime environment similar to the .NET framework.

4.2 Visualisation

The simulation tool structure is shown in Figure 7.

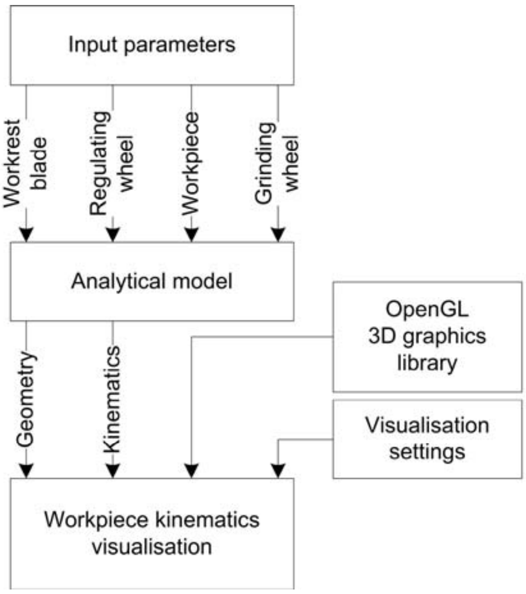


Figure 7: Simulation tool structure.

The analytical model of the grinding gap inputs the parametrical data that describe all grinding gap elements. When the parameters are determined, the analytical model outputs the geometry and kinematics of the elements. In case any errors are found, they are reported to the user and to the visualisation classes, which are able to represent them graphically. All images are built with the OpenGL function calls. Figure 8 shows a screen image of the simulation tool.

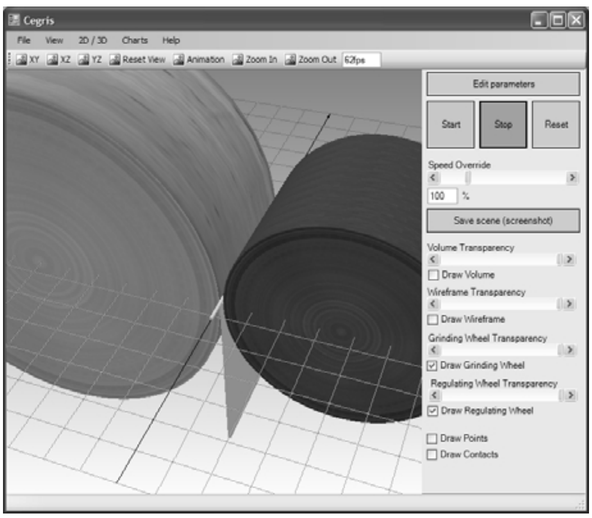


Figure 8: Simulation tool screenshot.

5 SUMMARY

A simulation tool was developed for the visualisation of workpiece kinematics in centreless throughfeed grinding. The simulation requires a parametric description of the geometry and the position of the grinding gap elements and an analysis of their kinematics.

In the next step, the simulation will be experimentally verified. The workpiece rotational and axial speed will be measured with noncontact eddy current sensors mounted into coolant needle nozzles.

The simulation of workpiece kinematics enables an operator to input set-up and process parameters in such a way as to achieve the required productivity and to avoid unsteady throughfeed. It can also warn an operator if error conditions exist for the given set-up. In this way it is possible to test different parameter combinations without trial-and-error on the machine tools, which in turn reduces the downtime and avoids the risks of crashes in reality.

The presented PC based simulation tool provides the basis for a state-of-the-art CAVE embedded software tool for training activities that will enhance knowledge and skills in centreless grinding technology.

6 ACKNOWLEDGMENTS

This research was conducted within the CEGRIS II project, which is sponsored by the European Commission as part of the Sixth Framework Programme for Research and Technological Development (FP6).

7 REFERENCES

[1] Klocke, F., Friedrich, D., Linke, B., Nachmani, Z., 2004, Basics for In-Process Roundness Error Improvement by a Functional Workrest Blade, Annals of the CIRP, 53/1: 275-280.

[2] Krajnik, P., Drazumeric, R., Meyer, B., Kopac, J., Zeppenfeld, C., 2008, Simulation of Workpiece Forming and Centre Displacement in Plunge Centreless Grinding, International Journal of Machine Tools & Manufacture, *In Press*, *Accepted Manuscript*.

[3] Rowe, W.B., Miyashita, M., König, W., 1989, Centreless Grinding Research and its Application in Advanced Manufacturing Technology, Annals of the CIRP, 38/2: 617-625.

[4] Cave Automatic Virtual Environment, http://en.wikipedia.org/wiki/CAVE, Accessed 28.1.2008.

[5] Meis, F.U., 1980, Geometrische und kinematische Grundlagen für das spitzenlose Durchlaufschleifen, Dissertation, RWTH Aachen University, Germany.

[6] Krajnik, P., Drazumeric, R., Meyer, B., Kopac, J., Klocke, F., 2007, Advanced Regulating Wheel Dressing in Throughfeed Centreless Grinding, 40th CIRP International Seminar on Manufacturing Systems, Liverpool, UK.

[7] Klocke, F., König, W., 2005, Fertigungsverfahren 2, Springer-Verlag, Berlin, Heidelberg, Germany.

Development of an electroplated polishing tape applying electrodeposited nickel foil method

Kentaro Miyazaki[1], Yasuhiro Tani[2], Kensuke Tsuchiya[1] and Yasuyuki Kamimura[1]

[1] Institute of Industrial Science, The University of Tokyo, Japan

[2] Department of Mechanical Engineering, College of Science and Engineering, Ritsumeikan University, Japan

Abstract

Diamond polishing tapes are commonly used for edge finishing of zirconia ferrules. There remains a problem that the diamond tape has too short tool life due to weak bonding strength. Then we propose a new method to make a diamond tape applying electrodeposited nickel foil method. This method is very useful to make a long metal foil continuously and commercially used for manufacturing of a copper foil. In order to improve the tensile strength of the foil, we performed composite electroforming utilizing glass fibers. It was confirmed that the concentration of abrasives could be increased by using Ni-electroplated diamond abrasives. A series of experiments to finish zirconia ferrules proved long tool life of the developed polishing tape.

Keywords:

Polishing; Fixed abrasive tape; Electroforming

1 INTRODUCTION

Pushing glass fibers through hard ceramic zirconia ferrule and then polishing the ends produce optical fibers for optical communication. This polishing process generally applies polishing tapes coated with fine grains of diamond so that materials with different hardness are effectively polished. Polishing tapes available in the market have abrasive grains adhered, with special binder, to the base material of PET film or unwoven fabric, however, tapes made in this manner have weak strength in holding grains.

Reports have been published, on the other hand, about electroplated diamond wire saws with abrasive grains fixed to the saw by electroplating[1][2]. This method allows continuous composite electroplating to produce high holding strength of the grains and is effective in producing polishing tapes with long tool life. Furthermore, composite electroplating of diamond grains during electrolysis, a process generally used for producing metal foil, allows continuously producing electroplated diamond tapes.

This paper proposes a continuous process of producing electroplated diamond polishing tapes with long life by applying electrolytic foil method.

2 PRODUCING ELECTROLYTIC NICKEL FOIL TAPES

2.1 Equipment for producing electrolytic nickel foil

This research uses nickel for the plating metal with adequate hardness as a tool material. There are a number of plating baths for nickel; a nickel chloride bath, sulfamic acid bath, Watts bath, and so on, and for its easy peeling after electroplating, we used a sulfamic acid bath also with relatively small stress in the electrodeposits and high hardness and toughness.

Based on electrolytic foil method of copper, we developed the continuous nickel foiling machine shown in Figure 1. The machine was equipped with a negative pole drum (100mm DIA.) for electrodeposition and peeling, an insoluble positive pole, a temperature controller to maintain constant temperature, an electrolysis bath, and a power supply. For the positive pole, we used titanium that is insoluble in an electrolysis bath and easy to maintain. Electrolytic foil method differs from general plating and a peeling process follows the electrodeposition of the foil. We adopted stainless steel for the negative pole drum because stainless steel easily generates a passivation film on its surface to make the peeling easy [3].

2.2 Property of electrolytic nickel foil

The machine described in the previous section successfully produced nickel foils, and we studied the electrodeposition parameters and the foil properties. Higher electrical current density makes the film forming faster in general, however, with our machine, a current density over 15A/dm2 caused burnt deposits. We, thus, set the electrical current density to 10 A/dm2 to avoid burnt deposits. The film forming speed with this current was about 5μm/min. The size of the negative pole drum produced a 50 μm thick nickel foil tape at the speed of

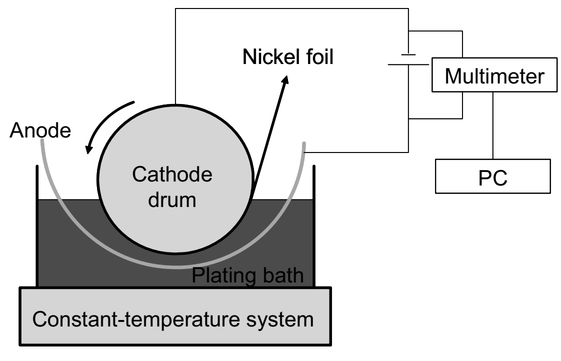


Figure 1: Producing equipment for nickel foils.

10 mm/min. A larger size negative pole drum will make the production faster. The nickel foil tape produced had a Vickers hardness of 160 (HV300kgf) and the tensile strength for a unit width was 32.0 N/cm. The stress corresponding to this value was 64MPa.

2.3 Adding filler

We needed to further enhance the strength of the nickel foil to give the polishing tape a longer life and processing performance. We strengthened the foil by adding a fibrous filler to produce a composite electroplated film.

The fibrous material to add had to be acid-resistant due to the acidic plating bath with pH 2.7. We also had to select a material with small diameter because our machine required to apply suspension eutectoid method for the fibers. Moreover, the expendable nature of polishing tapes requires the fibrous material to be inexpensive. Based on the above factors, we compared 3 types of fibrous material shown in Table 1. We also added cationic surfactant to disperse the filler within the plating solution.

First, carbon fiber A showed poor affinity with the plating solution, and although we observed eutectoid of the fiber, we also found burnt deposits when the addition was 1g/L or more. The probable cause was reduction in the nickel ions due to obstruction of boundary activities by the fibers attached on the pole. So we produced nickel foil tapes with the remaining 2 types of fillers and compared the Vickers hardness and tensile strength values. Figure 2 shows the results. Vickers hardness improved proportional to the amount of eutectoid fiber with both the carbon fiber and glass fiber. The tensile strength, on the other hand, improved with both fibers, however, the increase in strength was much bigger with glass fiber.

To identify the cause, we observed the electrodeposition of the fiber by etching the surface with a nickel solution. Figure 3(a) shows the eutectoid fibers in case of carbon fiber A and Figure 3(b) of carbon fiber B. The electrically conductive carbon fiber B was affected by the electrical field and many of the fibers were aligned in the vertical direction. The nonconductive glass fiber probably behaved similarly to nonconductive carbon fiber A to disperse in the horizontal direction giving the foil a stronger tensile strength than carbon fiber B.

The above results led us to use fine glass fiber as the reinforcing filler for the nickel foil tapes. The tensile strength of the polishing tape increased from 64MPa without the filler to 152MPa with the filler of 2g/L.

Table 1: Properties of fillers.

Type of fillers	Surface treatment	Diameter × Length (Average) mm	Specific gravity g/cm^3
Carbon fiber A	No	ϕ7×30	1.76
Carbon fiber B	Electroconductive	ϕ7×60	1.79
Glass fiber	No	ϕ1×100	2.53

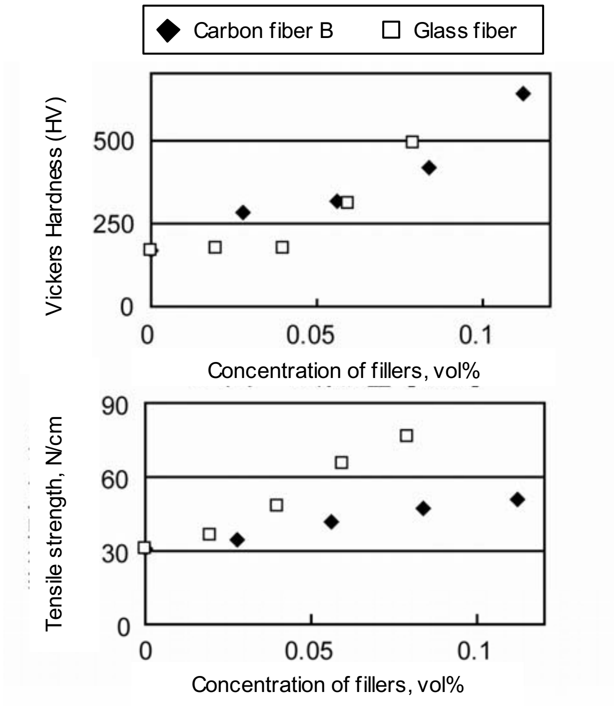
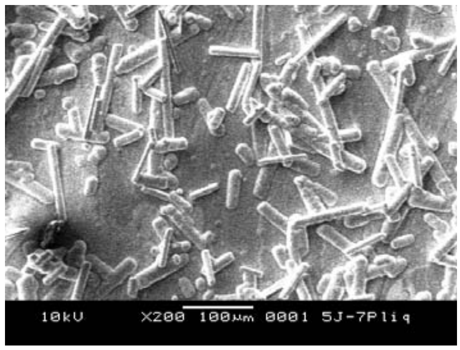
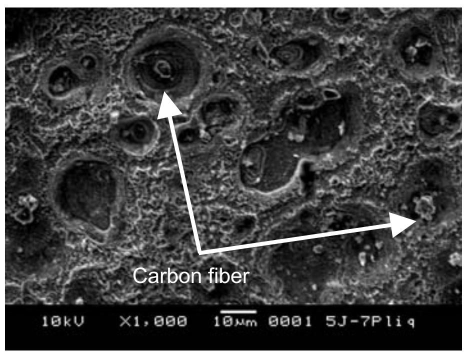


Figure 2: Vickers hardness and tensile strength.



(a) Case of nonconductive fiber A



(b) Case of conductive fiber B

Figure 3: Dispersion of carbon fibers.

3 PRODUCING ELECTRODEPOSITION DIAMOND TAPE

We first tried producing electrodeposition diamond tape by composite plating in an electrolysis bath dispersed with diamond grains.

We prepared 3 types of abrasive grains to bond; diamond grains for resin bonding with diameters of 8 to 16 μm (IRM), diamond grains for metal bonding with diameters of 3 to 8 μm (IMM), and nickel coated diamond grains with diameters of 10 to 20 μm (IRM-NP).

The nonconductive IMM and IRM showed small amount of adhesion to the tapes even when we increased the amount to add (Figure 4-a). The conductive IRM-NP, on the other hand, showed a large amount of grains bonded to the surface (Figure 4-b).

From the above results, we moved on to continuous production of electroplated diamond polishing tapes. Figure 5 shows SEM images of the surface of the diamond tape we prototyped. The figure shows plenty of eutectoid diamond grains and glass fibers.

4 POLISHING EXPERIMENTS WITH ELECTROPLATED DIAMOND TAPES

To evaluate the polishing performance of our electroplated diamond tapes we prototyped, we polished the side surface of zirconia ferrule (the front face processing showed 1.12μmRy with centerless grinding) with the process parameters shown in Table 3. For comparison, we also used an electroplated diamond sheet (base material was copper) available in the market with about the same sized diamond grains (#1000, average diameter 11.4μm).

Figure 6 shows the durability and change in surface roughness as we repeated the polishing process with the same area on the prototyped tape. The figure shows improvement in the surface roughness as the number of repeated processes increase. The probable cause of this improvement over time is because more cutting edges of the diamond grains got exposed with repeated processes. We also found the prototyped tapes highly durable because they did not lose their cutting performance even after 10 repetitions of the polishing process. Figure 7 compares the results of polishing with of our prototyped polishing tape and the electroplated diamond sheet from the market. The figure shows that the polishing tape from the market has a higher polishing performance (speed), however, the surface finish was worse than the one produced by our prototype. We then compared the two polishing tapes by the rate of performance over surface finish, P/R. Both the polishing performance and the surface roughness relate positively to the unit of polishing process. These values indicate that the prototype is not at all inferior to the tape from the market. Figure 8 shows the surface of our electroplated diamond tape after 10 repetitions of the polishing process. The prototyped tape showed no buried or dropped grains to demonstrate its high durability in polishing.

5 CONCLUSIONS

We applied electrodeposited foil method to continuous production of electroplated polishing tape for the purpose of producing diamond tapes with long life, and we evaluated their features and polishing performance. The results show:

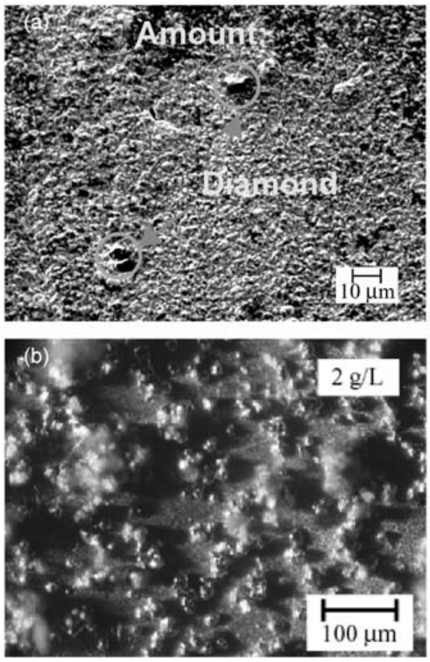


Figure 4: (a) Electroplated surface using IMM grains. (b) Electroplated surface using Ni coated diamond grains.

Table 2: Electroplating conditions.

Plating bath composition	$Ni(SO_3NH_2)_2 \cdot 4H_2O$:450g/l $NiCl_2 \cdot 6H_2O$:30g/l H_3BO_3 :30g/l Glass fiber :2g/l Diamond grain :1g/l
Current density	10A/dm^2
Manufacturing speed	10mm/min
Bath temperature	60^{o}C
Bath pH	2.7

Table 3: Machining conditions.

Work revolution	500 min^{-1}
Table feed rate	20 mm/min
Machining time	2.5 min
Oscillation speed	25 Hz
Oscillation amplitude	4 mm
Suppress strength of backup roller	0.3 MPa
Hardness of backup roller	50 Hs

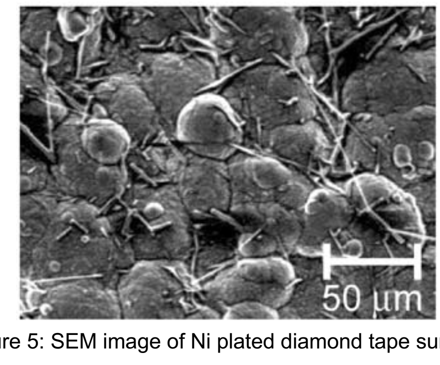


Figure 5: SEM image of Ni plated diamond tape surface.

1. We can apply electrodeposited foil method to continuously produce electrodeposited diamond tapes without base material.
2. Good material to add to electrodeposited nickel foil tape to enhance its mechanical strength is nonconductive fine glass fibers with small diameter.
3. The ideal amount of addition of nickel coated diamond is 1g/L. Excessive addition over this amount causes the plating to grow from the electrodeposited gains and weakens the force holding the grains.
4. The prototyped polishing tape improved its polishing performance every time it completed a polishing process to demonstrate that is has high durability against wear. As the number of polishing processes it completed grew, the prototyped polishing tape showed stable polishing performance.

6 REFERENCES

[1] Suwabe, H. et al., 1994, Slicing by Wire Tool Electrodeposited Diamond Grains, Journal of the Japan Society for Precision Engineering, 60/2:188-191. (in Japanese)

[2] Chiba, Y. et al., 2003, Development of a High-Speed Manufacturing Method for Electroplated Diamond Wire Tools, Annals of the CIRP, 52/1:281-284.

[3] Haruyama, S., 2001, Electrochemistry for Surface Engineers, Maruzen: 146-153.

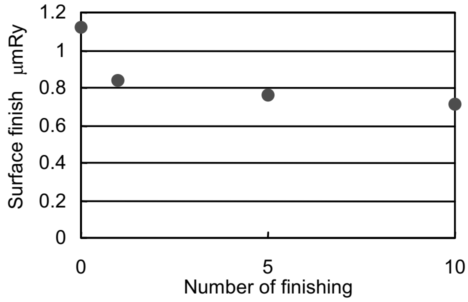


Figure 6: Comparison of surface roughness.

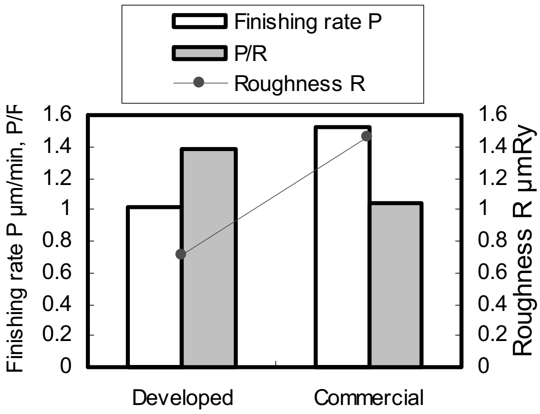


Figure 7: Comparison of finishing characteristics.

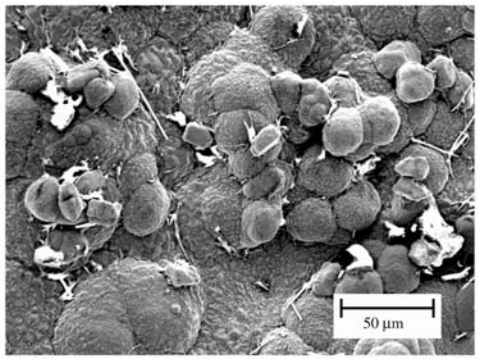


Figure 8: Surface of diamond tape after 10 times of polishing.

Development of Ultrafine-Crystalline cBN Abrasive Grains for Innovative Grinding Technology

Yoshio Ichida[1], Ryunosuke Sato[1] and Hitoshi Kajino[2]

[1] Department of Mechanical Systems Engineering, Faculty of Engineering, Utsunomiya University, Japan
[2] Corporate R&D Center, Mitsui Mining & Smelting Co. Ltd, Japan

Abstract
This paper presents some mechanical properties and grinding characteristics of newly developed ultrafine-crystalline cBN (UcBN) abrasive grains. This new UcBN grain possesses an ultrafine crystal structure composed of sub-micron sized primary crystal grains. The fracture strength of the UcBN grain is about 1.5 times higher than that of conventional polycrystalline cBN grain. The grinding ratio in grinding with the UcBN grain is 4-10 times higher than those in grinding with conventional typical monocrystalline and polycrystalline cBN grains. Therefore, the UcBN abrasive grain has an enough ability to innovate the grinding technology.

Keywords:
Ultrafine-crystalline cBN abrasive grain; Fracture strength; Grinding characteristics

1 INTRODUCTION

Cubic boron nitride (cBN) grinding wheels are becoming widely used for the machining of various engineering materials, such as steel, cast iron and superalloys [1][2]. However, recently, as high-level machining systems using cBN wheels, such as multifunctional grinding centers are developed, demands for a higher performance of cBN wheels are increasing. To enhance the grinding performance of cBN wheels, we have developed a new type of polycrystalline cBN abrasive grain by direct transformation from hexagonal boron nitride, which was produced from chemical vapor deposition process [3]. This new cBN abrasive grain possesses an ultrafine crystal structure composed of submicron sized primary crystal grains. Therefore, we call this new abrasive 'ultrafine-crystalline cBN (UcBN or cBN-U for short) grain'. UcBN abrasive grain is expected to be used for wide applications from high-efficiency grinding to high-quality grinding, because its wear resistance is much more effective than those of conventional monocrystalline and polycrystalline cBN abrasive grains.

This paper presents some mechanical properties and grinding characteristics of the new ultrafine-crystalline cBN abrasive grain manufactured under an optimum condition. The fracture strength of the cBN-U grain is measured using an apparatus for fracturing a single grain by a compressive force. Moreover, grinding experiments using the cBN-U grain are carried out and its grinding performance is compared with those using representative conventional polycrystalline and monocrystalline cBN grains.

2 PROPERTIES OF ULTRAFINE-CRYSTALLINE CBN ABRASIVE GRAINS

The manufacturing process of the ultrafine-polycrystalline cBN abrasive grains is shown in Fig.1. First, a hexagonal cubic boron nitride (hBN) disk made by means of the chemical vapor deposition process, whose dimension was 6 mm in diameter and 1 mm in thickness, was prepared as the

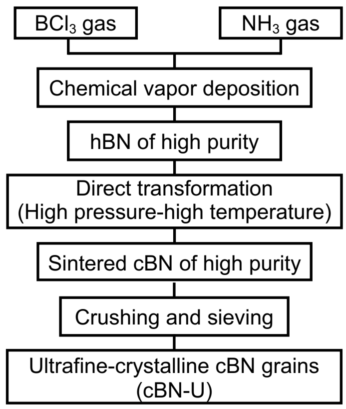
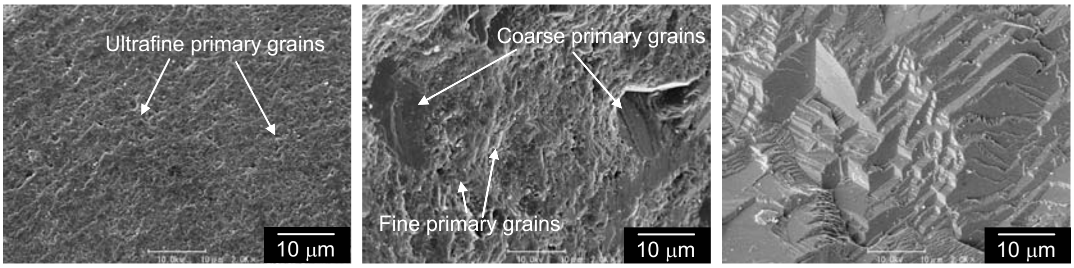


Figure 1: Manufacturing process of ultrafine-crystalline cBN (cBN-U) abrasive grains

starting material. Next, it was set in a Ta capsule, and transformed directly to polycrystalline cBN under an ultrahigh pressure 7.7 GPa and an temperature higher than 2073 K using a modified belt-type high-pressure apparatus [3][4]. Moreover, the cBN samples obtained were crushed into pieces using a roll crusher. The cBN abrasive grains with sizes of #170/230-#60/80 were obtained by classifying the crushed cBN powder.

Fig. 2 shows the typical SEM images of the fracture surfaces of the new polycrystalline cBN-U, conventional polycrystalline cBN-W5 and monocrystalline cBN-B abrasive grains. Some large primary crystal grains (Monocrystalline cBN) with a grain size of 10-20 μm are observed in the fracture surface of cBN-W5 grain. Namely, the conventional polycrystalline cBN-W5 grain has a crystal structure composed of micron sized primary-crystal grains and coarse primary-crystal grains with a grain size of 10-20 μm. In contrast, the cBN-U grain has an ultrafine crystal structure composed of only sub-micron sized primary-crystal grains.



(a) cBN-U (b) cBN-W5 (c) cBN-B

Figure 2: SEM images of fracture surfaces of ultrfine-crystalline cBN (cBN-U), conventional polycrystalline cBN (cBN-W5) and monocrystalline cBN (cBN-B) abrasive grains.

The cBN grains with a size of #60/80 were used for a compressive fracture test to measure the fracture strength [3][5]. In this test, a single grain is compressed between two sintered diamond compounds and the normal load when it is fractured is measured as a fracture load *W*. Although the applied load is compressive, brittle fracture occurs due to a tensile stress which develops perpendicular to the loading axis. This loading is similar to the loads applied to grains in grinding, and provides a useful comparison of abrasive materials relative to their bulk strength. On the basis of the Griffith's theory on brittle fracture, the fracture strength σ_t (tensile strength) can be approximated as [5]

$$\sigma_t = \frac{\overline{W}}{0.32A} \tag{1}$$

where $\overline{W}$ is the average fracture load and *A* is the average projected area of cBN grain. The loading velocity is 1.67 N/s, and number of samples is 250.

Fig.3 shows the results in weibull distribution of the fracture load *W* obtained from the compressive fracture test. The average fracture load $\overline{W}$ and the tensile strength of cBN grains obtained on the basis of these weibull plots are shown in Table 1 and Fig. 4. The tensile strength of cBN-U grain is about 1.52 times higher than that of cBN-W5 grain. In addition, the difference in weibull modulus between these cBN grains indicates that the strength of cBN-U grain is more stable than those of conventional cBN grains.

Other properties measured for three types of cBN grains are also summarized in Table 1. The hardness of the cBN sample was measured by the indentation method using the Vickers indenter with a normal load of 300 g for 15 s. The optical emission spectrometry analysis of cBN grains has proved that the contents of metallic impurities in cBN-U grain are much lower than those in cBN-W5 and cBN-B grains.

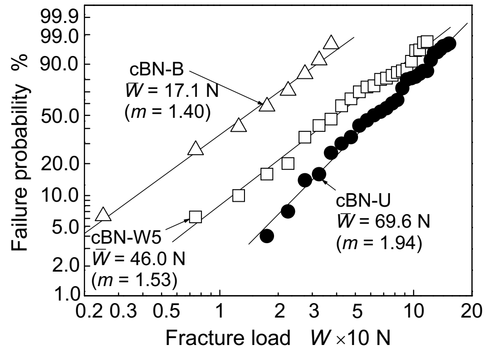


Figure 3: Weibull plot of compressive fracture load *W* ($\overline{W}$: Average fracture load, *m*: Weibull modulus).

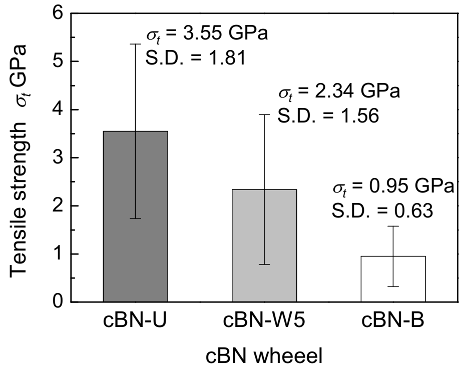


Figure 4: Comparison in tensile strength between cBN-U grain and conventional cBN grains.

Table 1: Mechanical properties of cBN grains (Grain size: #60/80).

cBN grain type (Symble)		Ultrafine-polycrystalline cBN (cBN-U)	Conventional polycrystalline cBN (cBN-W5)	Monocrystalline cBN (cBN-B)
Density	g/cm^3	3.4	3.4	3.4
cBN purity	wt%	> 99.9	> 99.9	> 99.9
cBN primary grain size	μm	< 1.0	< 20	-
Hardness HV (R.T.)	GPa	56-60	53-58	-
Average fracture load $\overline{W}$	N	69.6	46.0	17.1
Standard deviation		35.7	30.6	11.3
(Weibull modulus *m*)		(1.94)	(1.53)	(1.40)
Tensile strength σ_t	GPa	3.55	2.34	0.95

Table 2: Grinding conditions.

Grinding method	Surface plunge grinding (Up cut)
Grinding wheel	cBN80L100V (cBN-U) cBN80L100V (cBN-W5) cBN80L100V (cBN-B) Dimensions: ϕ200 × t10 mm
Peripheral wheel speed	v_s = 33 m/s
Work speed	v_w = 0.15 m/s
Wheel depth of cut	a = 10 μm
Grinding fluid	Soluble type (JIS W-2-2) 2% dilution
Workpiece	High speed steel (JIS SKH51) Hardness: 65HRC Dimensions: $100^l \times 5^t \times 30^h$ mm
Dressaing method	Rotary diamond dresser Speed ratio: 0.5 Dressing lead: 0.1 mm/rev Dressing depth of cut: 2μm × 5 times

3 GRINDING PERFORMANCE OF ULTRAFINE-CRYSTALLINE CBN WHEELS

3.1 Grinding force and grinding energy

To clarify the fundamental grinding performance of the UcBN abrasive grain, surface plunge grinding experiments were conducted on a horizontal spindle surface grinding machine. The grinding conditions are listed in Table2.

Fig.5 shows the changes in the normal grinding forces F_n' and tangential grinding force F_t' as the accumulated stock removal V_w' increases in the grinding processes with three types of cBN grains. The grinding forces in using the cBN-U grain are much lower than those in using the conventional cBN grains. In particular, after the stock removal V_w' exceeds around 5000 mm^3/mm the grinding forces for the cBN-U grain are reduced by 30-40 % compared with those in using the cBN-W5 and cBN-B grains.

Using these tangential forces, the specific grinding energy U_e was obtained from the equation $U_e = (F_t'v_s)/(av_w)$ (F_t' : tangential grinding force per unit grinding width, v_s : peripheral wheel speed, v_w : work speed, a: wheel depth of cut). The results are show in Fig. 6. The specific grinding energy in using cBN-U grain is 60-70 % of those in using cBN-W5 and cBN-B grains. The majority of the grinding energy is converted into the heat. The lower the specific grinding energy is, the less the grinding heat generated at the contact face between the wheel and the workpiece becomes. This low specific grinding energy is one of the important features in grinding with the cBN-U grain. Therefore, the cBN-U grain may be used for constructing an environmentally conscious grinding system such as a coolantless grinding.

3.2 Wheel wear characteristics

Fig.7 shows the changes of the radial wheel wear ΔR as the accumulated stock removal increases in the grinding processes with three types of cBN grains. In any case, the radial wheel wear increases rapidly at a initial stock removal range from 0 to 500 mm^3/mm (Initial wear region) and after the stock removal of 500 mm^3/mm increases gradually almost in proportion to the stock removal V_w' (Steady-state wear region). The inclination of straight line in the steady-state wear region, namely, the wheel wear volume per unit stock removal is defined as the volumetric wheel wear rate r_g. The value of r_g for cBN-U grain is much lower than those of conventional cBN grains.

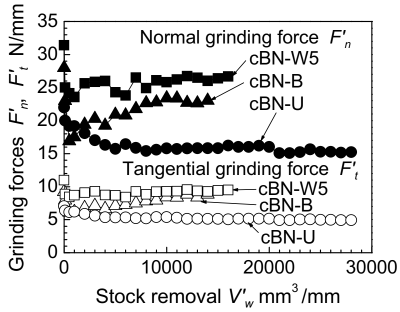


Figure 5: Changes in the normal grinding forces F_n' and tangential grinding force F_t' as accumulated stock removal V_w' increases.

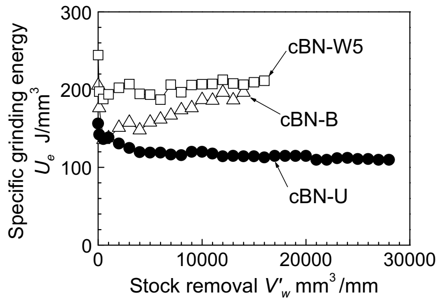


Figure 6: Changes in the specific grinding energy U_e as accumulated stock removal V_w' increases.

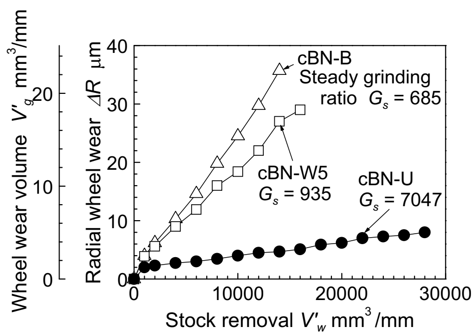


Figure 7: Change of radial wheel wear ΔR as accumulated stock removal V_w' increases.

The values of the steady grinding ratio G_s, that is the reciprocal of r_g, in the grinding processes with the cBN-U and conventional cBN grains are indicated in Fig.7. The grinding ratio G_s in using the cBN-U grain is about 7 times higher than that in using the cBN-W5 grain, and about 10 times higher than that in using the cBN-B grain. The results of three-dimensional SEM observation for the working surface of UcBN grinding wheel have indicated that the wear mechanism of the UcBN grain cutting edges is dominantly due to the micro brittle attrition wear. Such high grinding ratio is another one of the important features in grinding with cBN-U grain. Therefore, the cBN-U grain may be widely used for grinding various difficult-to-grind materials.

3.3 Characteristics of ground surface roughness

Fig.8 shows the changes in roughness Ra of the ground surfaces as the accumulated stock removal increases in the

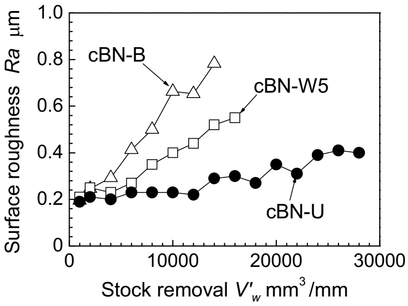


Figure 8: Change of surface roughness Ra as accumulated stock removal V_w' increases.

grinding processes with three types of cBN grains. The rate of increase in surface roughness with the stock removal in using the cBN-U grain is much lower than those in using conventional cBN grains. This low rate of increase in the roughness is brought from the high wear resistance of the cBN-U grain.

4 APPLICATION TO PROFILE GRINDING

Because the UcBN grain has some effective features as described above, it can be applied to various kinds of grinding operations, such as an ultrahigh speed grinding, high quality grinding and so on. In this paper, as a representative application, the experimental results when the cBN-U grain is applied to the creep feed profile grinding are described.

As shown in Fig.9, experiments for producing a V-shaped groove on a flat surface in one pass by creep feed grinding are carried out using the cBN-U and conventional cBN-B grains [6]. High-speed steel (SKH51/JIS) and Ni-based superalloy (Nimonic 80A) have used as the work material. When grinding with the cBN-U grain, both radial wear and profile wear are less, and hence the grinding ratio is 4 −10 times higher than that with conventional cBN-B grain, as shown in Fig.10. Grinding force in grinding with the cBN-U grain is reduced by 10-20 % compared with that in grinding with the cBN-B grain [6]. Thus, the cBN-U abrasive grain is suitable for application with a high dimensional accuracy in creep feed profile grinding, because it gives less profile wheel wear, and hence better form retention of the wheel, than conventional cBN abrasive gains.

5 CONCLUSIONS

The main results obtained in this study are summarized as follows:

(1) The fracture strength of the cBN-U grain is about 1.5 times higher than that of conventional polycrystalline cBN-W5 grain.

(2) The cBN-U grain has a very high wear resistance as compared with conventional cBN grains. Hence, the grinding ratio in grinding with cBN-U grain is 4-10 times higher than that in grinding with conventional cBN grains.

(3) The estimated tool life of the wheel in using the cBN-U grain, based on the roughness of the finished surface, is much longer than those in using conventional cBN grains.

(4) The cBN-U abrasive grain is suitable for application with a high dimensional accuracy in creep feed profile grinding, because it gives less profile wheel wear, and hence better form retention of the wheel, than conventional cBN abrasive grain.

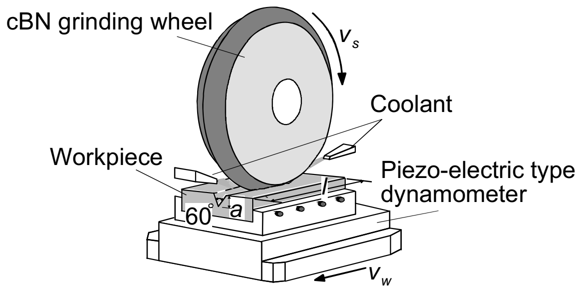


Figure 9: Experimental setup for creep feed profile grinding.

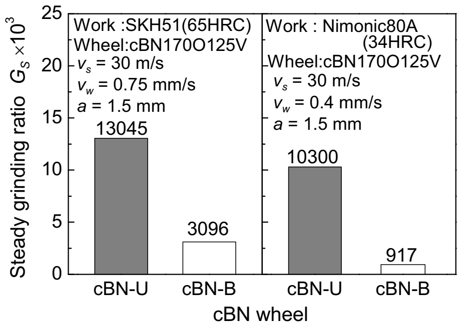


Figure 10: Comparison in grinding ratio between cBN-U and conventional cBN grains.

As mentioned above, the UcBN grains will play an increasing important role in developing an innovative grinding technology needed in the 21st century.

Acknowledgments

This research was supported in part by Grants-in-Aid for General Science Research (c) (No.19560106) from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

References

[1] Webster, J. and Tricard, M., 2004, Innovations in Abrasive Production for Precision Grinding, Annals of the CIRP, 53/2: 597-617.

[2] Ichida, Y., Sato, R., Morimoto, Y., Ohsawa, Y, and Fredj, B., N., 2006, Formation Mechanism of Finished Surface in Ultrahigh-Speed Grinding with cubic Boron Nitride (cBN) Wheels, JSME Int. J. (Series C), 49/1: 100-105.

[3] Ichida, Y. and Kishi, K., 1997, The Development of Nanocrystalline CBN for Enhanced Superalloy Grinding Performance, ASME J. Manuf. Sci. Eng, 119/1:110-117.

[4] Yamaoka, S., Akashi, M., Kanda, H., Osawa, T., Taniguchi, T., Sei, H. and Fukunaga, O., 1992, Development of belt type high pressure apparatus for material synthesis at 8 GPa, J. of High Pressure Institute of Japan, 30: 249-258.

[5] Yoshikawa, H. and Sata, T., 1960, Fracture Strength of Abrasive Grains, J. Jpn. Soc. Prec. Eng., 26/8: 47-482.

[6] Ichida, Y., Sato, R., Morimoto, Y. and Inoue, Y., 2006, Profile Grinding of Superalloys with Ultrafine-Crystalline cBN Wheels, JSME Int. J. (Series C), 49/1: 94-99.

Service Engineering

System for Planning of Resources in IPS²-Delivery

Horst Meier[1], Christian M. Krug[1]
[1] Chair of Production Systems, Ruhr-University Bochum, Bochum, Germany

Abstract
Industrial Product-Service-Systems (IPS²) represent highly specific customized products with integrated development of product and service shares. Due to the multiplicity of customized IPS² that are used parallel the responsible person has to be supported in the planning of the IPS² delivery over the life-cycle.
This support must be two-stage. At first the automated aggregation of the single resources requirements to the best possible total resources requirement is done considering the customer related boundary conditions. Thereupon the total resources requirement is analyzed in detail regarding the intensity of planning of each resource. The intensity of planning in this context specifies how intense the responsible person has to address himself to which resource.
This article describes basic methodic approaches to execute the necessary planning steps that can constitute fundamentals for a computer-aided implementation.

Keywords:
Industrial Product-Service-Systems, IPS²; Resource Planning

1 INTRODUCTION

Industrial Product-Service-Systems can be characterized by the fact that they consist of a combination of tangible product and intangible service shares, whose intangible service shares provide a value to the customer via the complete life-cycle [1, 2]. In the different function-, availability- or result-oriented use models [3] the order for the service performance is either given by the customer (function oriented) or by the provider (availability-/result-oriented). The planning of the required resources, necessary for service performance, has to be done automatically so that the huge planning scope can be dominated.

Initial Situation and Attributes of Service Processes to be planned

The initial situation for the capacity planning is the well known process plan (Figure 1).

In this plan the workflow is given. Each process shows a specific variance depending on the kind of service (e.g. breakdown <P8/9>– short variance, optimization <P7> – long variance).

This variance can be used to minimize efforts with the common methods of combining, splitting or rescheduling of lots. In case of Industrial Product-Service-Systems these possibilities for optimization are complemented by a number of influence capabilities. Among these possibilities for optimization there are the partial substitution of product and service shares, the operation with alternative resources (e.g. service technicians with different qualification level) or the using of alternative delivery processes [4]. To use this for enabling an effective capacity planning further attributes in addition to the describing matters of conventional processes have to be carried. In addition to the variance in time there are among these attributes:

- Possibilities of substitution of product and service shares
- Alternative resources
- Alternative processes

They are fixed during the development phase and may be extended during the delivery phase based on the broadened know-how.

Further on the IPS² specific restrictions listed below have to be considered:

- Place of delivery
- Use model within the feature is offered
- Available service time
- Process responsibility
- Influence on customer satisfaction (conspicuity of process result similar to FMEA)
- Service specific characteristics of the delivery process
  - Uno actu principle
  - Immateriality an thus no storability (preliminary/ end combination)
- Ramp-Up of the delivery processes

The two step parameterization of the attribute „place of delivery" here feature a particularity. During the process development at first the place of delivery is fixed concerning the decision of execution at customer's or at supplier's site.

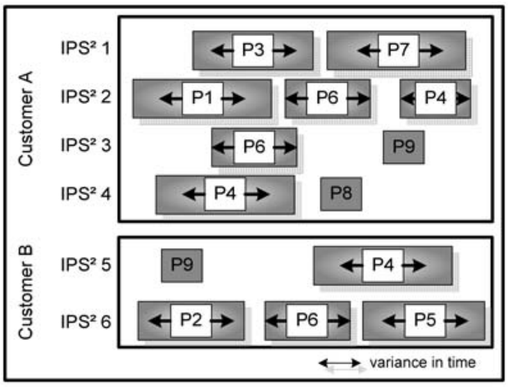


Figure 1: Workflow plan for IPS² delivery processes

In the second step when dispatching to the IPS² Execution System (IPS²-ES) the data on the supplier or network partner executing the service or otherwise on the customer are stored. The use model (fuction, availability, result oriented) characterizes due to the different turn over models the choice of resources (e.g. quality of a spare part) and the dedicated reaction time. Service times in the majority of cases are downtimes for the customer. If the remaining time for maintenance of an Industrial Product-Service-System is mostly spent for the repair of a breakdown it may become necessary to combine services or precautionary change or replace plant components due to safety reasons to ensure the warranted availability. This planning cannot be considered within the operative planning, but has to be done proactively, because at the moment of delivery the capacities are already scheduled and have to be applied according to their use as resources. Thus the remaining service time has to be stored with the process description. Furthermore the responsibility for the service execution within an Industrial Product-Service-System may pass over from supplier to customer respectively from customer to supplier [5]. Also the customer and his satisfaction take center stage of service processes so the conspicuity of the process result is a main criterion according to the process priority. The uno actu principle describes the concurrence of creation and consumption of services. Except the product share of a service process the immateriality of services leads to no storability. Corsten therefore divides the service delivery into preliminary and end combination; so the creation of performance capability (training of service personnel, production of spare parts) and the actual service delivery [6]. The ramp-up of the delivery processes represents a particular characteristic of Industrial Product-Service-Systems. It describes how many process executions have to be done by the same technician to reach the maximum efficiency.

2 STRATEGY FOR OPTIMIZATION

On process level only the mentioned methods of combining, splitting or rescheduling of lots can be used. This does not give consideration to the potential for optimization of Industrial Product-Service-Systems.

This means that the description on process level is not detailed enough to optimize the total resources requirement and to plan the future needed capacities of each resource and to optimize the needed amount by this step, because the resources are hidden behind the processes' surface. Further on the processes could be changed on resource level while delivery:

- to fulfill the aim of the process although the planned resources are missing
- due to changing customer demands or
- the dynamic of the business model.

Therefore the durability of spare parts or the qualification of the service technicians might change. This has to be considered by the capacity planning as well.

So there is need to analyze the processes up to the resource level. By doing this the weakness of the automatic planning becomes clear. For this an example will be given.

Due to the automated parallel planning of delivery processes based on the use model it may occur that within short time several processes have to be executed within the same customer organization. That can lead to combining of lots on process level to reduce transition times. The workload of the single resources as personnel or tools thereby remains unconsidered.

In the following example it is essential to exploit the potential for optimization on resource level: A maintenance process to be executed within a result oriented use model requires a qualified technician to be at the customer's organization for three days. Another delivery process, for example the extension of an Industrial Product-Service-System, requires several technicians with partly the same level of qualification. Because the workload of the technicians is always lower than hundred percent the maintenance processes can be executed by the technicians necessary for the extension by combining the lots. This results in saving the working time of one technician. This very simple example shows the potential for optimization of service delivery to only one customer.

The entire problem however features a much higher complexity. Cause every state of the resources' capacity over time is associated with customer related restrictions. These restrictions may exert influence on the occurrence of the product share and feature dependencies with each other.

Customer and process related restrictions of states in the capacity over time

From analyzing the delivery processes of Industrial Product-Service-Systems up to resource level the demands for each resource result and have to be implemented into capacities after their optimization (Figure 2).

The differing notations for the resources are based on the different type of resources. The resources „Technician" and „Measuring equipment" are not expended and thereby are only available for the process during a certain time whereas the resource "Lubricant" is expended with its application. For this reason for a resource temporary assigned to an IPS² time and duration of the demand are drawn in, whereas for expended resource only the moment of demand is noted, because this resources keep being part of the IPS² for indefinite time.

To enable the optimization of the demands the restrictions have to be noted with the resources description. The well known characteristics from the conventional planning as amount and date therefore no longer fit the specific needs of Industrial Product-Service-Systems.

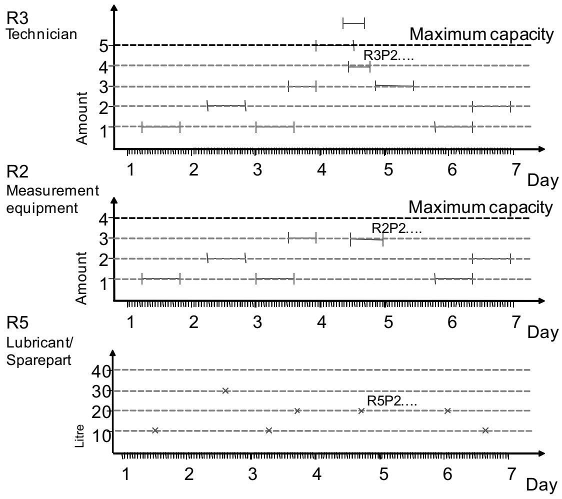


Figure 2: Capacity over time of selected resources

With the characteristics to be noted with the resources there are the already above mentioned process restrictions as place of delivery, use model within the feature is offered, available service time, process responsibility and the influence on customer satisfaction of the process result. In addition the following resource specific details have to be noted with the resources:

- Article number of the resource
- Further resources necessary for this process
- Resource responsibility
- Alternative resources
- Additional resources in case of use of alternative resources
- By-products of the resources subject to the process step
- Location of the resource
- Qualification of the respective technician

The article number of the resource is needed to identify a resource within its process context and to link the resources connected to this process. Like the process responsibility also the responsibility for the provision of a resource may switch between customer and supplier of an IPS². An example is the provision of an overhead crane by the customer within a service delivery process the supplier is in charge for. In the context of IPS² the application of alternative resources to achieve a defined process result is possible. This can be exemplified with the assignment of a lower qualified person to execute a delivery process. This requires the provision of an enhanced documentation or instructions an additional resource. With regard to sustainability it is of particular importance that the by-products emerging with the use of a resource are noted with the resource description. By the notation in the form of mathematic descriptions it is possible to determine the impact on the environment already at the time of planning [7]. The location of the required resource is on of the major variables together with the place of delivery, because by their combination the transition time between the delivery processes is determined. The locality of the involved resources and the places of delivery can partly be compensated by the allocation speed. Here the attributes location of resources and qualification of the technician are special, because they continuously change and so have to be updated.

Each point in the capacity over time estimated by the process analysis consequently features process and customer specific restrictions that have to be considered during the optimization.

Objectives

When transferring the objectives of the conventional capacity planning short throughput time, adherence to delivery dates, high capacity utilization and low stock to Industrial Product-Service-Systems it can be found out that these objectives remain valid without change. Within the capacity planning of Industrial Product-Service-Systems however the constant and high capacity utilization and the minimal stock of preliminary combinations and supplies are in the focus.

To achieve these objectives the IPS² specific possibilities for optimization using the information noted with the resources have to be exploited. Among these possibilities for optimization there are the variance in time of the delivery process (combining, splitting or rescheduling of lots), the resource variance as the use of alternative resources (technician, supporting methods), exploiting the process variance (rescheduling of the tasks' sequence), the variation of the allocation speed (Transition time, journey => current place of technician, material and IPS²), the partial long-term substitution of product and service shares as well as the service distribution, that means delivering the service by a partner of the supplier.

These tasks may partly be use both in the strategic capacity planning and the operative resource planning. The distinction between capacity planning and operative resource planning against this background is that within the capacity planning there remains enough time to alter or create capacities whereas during the operative resource planning a planning problem has to be solved only by rescheduling the existing capacities.

Figure 3 shows the recommended strategy to determine the optimized entire resource demand. The two step procedure at first reduces the transition times; therefore the variance in time of the delivery processes is used.

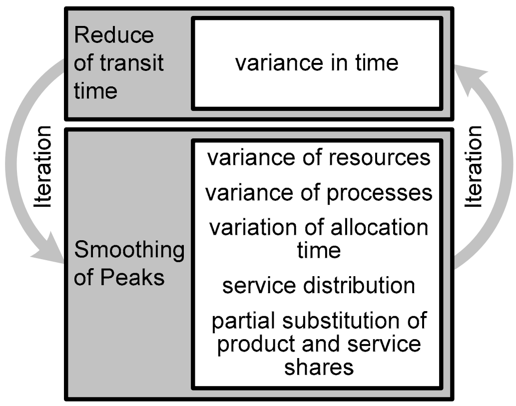


Figure 3: Optimization of the total resource requirement

The result thus represents the capacity over time optimized regarding a specific factor and therefore the corresponding capacities have to be built up.

These optimization steps have to be executed repeatedly with every alteration of the Industrial Product-Service-System mentioned before, as changing use model or substitution of product and service shares.

3 AUTOMATED IDENTIFICATION OF THE INTENSITY OF PLANNING

With the objective to ensure an efficient execution of the delivery planning and due to the multiplicity of customized Industrial Product-Service Systems the necessity of planning respectively the intensity of planning of the different resources preferably have to be automated determinable.

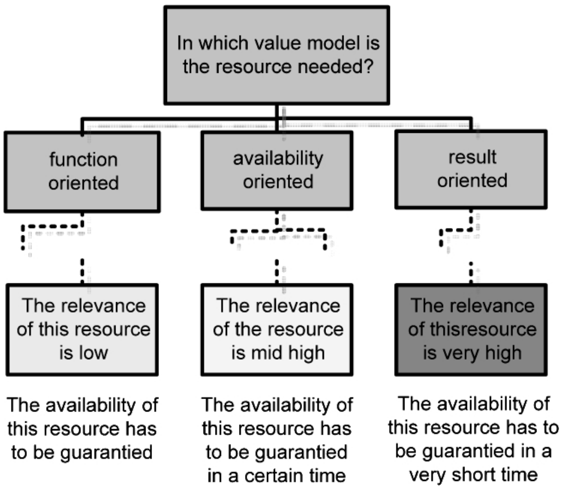


Figure 4: Part of the decision tree

The task now is not only to make a Yes/No decision regarding the resource but to judge the resource referring to its application and the chosen use model regarding the intensity of planning. Against the background of a multi dimensional problem a decision tree suits best to derive the recommended action. This method features the advantage that by the sequential chain of decisions the basic question -How far has the resource to be considered within the planning?- is clarified more and more precisely and due to the user-defined level of detail also special cases can be served with a recommended action.

Based on the question whether the resource to be used is needed for a service execution or for the operation of an Industrial Product-Service-System the decision tree goes to different branches up to the recommended action for the planner. To make clear the intensity of planning the recommended actions are coded with the colors green, yellow and red. Figure 4 shows up a part of the decision tree. A short example for a sequence of questions will be given: Does the promised availability directly depend on this resource? Yes=> Are there any alternative resources or processes? Yes=> Do they have a more efficient cost structure? ... This leads to very concrete recommended actions depending on the decisions like “No Stock in OEM-Network, Just in Time” or “This resource has to be stored in a stock within the OEM-network”. This will help the planner to have a look at the most important points in planning.

4 CONCLUSION

It was shown that for an efficient capacity planning of IPS² the delivery processes have to be resolved up to resource level to exploit the IPS² specific optimization criteria and which information therefore have to be noted with the resources. In the last part of the article the decision tree for determining the necessity of planning was presented.

5 ACKNOWLEDGEMENT

We express our sincere thanks to the Deutsche Forschungsgemeinschaft (DFG) for financing this research within the Collaborative Research Project SFB/TR29 on Product-Service-Systems – dynamic interdependency of products and services in the production area.

6 REFERENCES

[1] Tan. A.R.: McAloone, T.C.: Charateristics of strategies in Product/Service-System Development, International Design Conference – Design 2006, Dubrovnik – Croatia, May 15-18, 2006

[2] Matzen, D.; Tan A.R.; Myrup Andreasen, M.: Product/Service-Systems: Proposal for Models and Terminology, 16. Symposium “Design For X”, Neukirchen, October 13-15, 2005

[3] Tucker, A.; Tischner, U.: New Business for Old Europe – Product-Service Development, Competiveness and Sustainability, Greenleaf Publishing, Sheffield, UK, 2005

[4] Meier, H., Krug, C.M., Strategische Kapazitätsplanung hybrider Leistungsbündel, PPS Management 11, 48-51, 2006

[5] Meier, H., Kortmann, D., Leadership – From Technology to Use, 14th CIRP Conference on Life Cycle Engineering, 159-163, 2007

[6] Corsten, H., Stuhlmann, St., Kapazitätsplanung in Dienstleistungsunternehmungen, in: Bruhn, M., Meffert, H., Handbuch Dienstleistungsmanagement, Gabler, Wiesbaden, S. 177-192, ISBN 3-409-23593-0, 2001

[7] Meier, H.; Krug, Ch., M.; Planning of Resources by Using Transfer functions, 15th CIRP Life Cycle Engineering Conference, Sydney, Australia, 2007

Product Diffusion in a Market with Network Externalities: An Approach from the Viewpoint of Value-creation among Consumers

Nariaki Nishino[1], Takeshi Takenaka[1] and Kanji Ueda[1]
[1] Research into Artifacts, Center for Engineering, The University of Tokyo, Japan

Abstract
Markets with network externalities, such as those for mobile telephones and internet services, have been growing. Network externalities are defined as externalities by which a consumer's utility depends on the number of users who use the same product. In such markets, high technology does not necessarily produce value in society because value is created through interaction among consumers. This study specifically examines how value in production and service is created through network externalities. To examine this problem, we first construct a consumer agent model based on results of a general lifestyle survey, and secondly construct a multi-agent system with consumer agents. Subsequently, we analyze decision-making through interaction among producers and consumers. Finally, we discuss the problem of product diffusion from the viewpoint of value-creation.

Keywords:
Network Externalities, Product Diffusion, Value-Creation, Multi-Agent System

1 INTRODUCTION

Markets with network externalities, such as those for mobile telephones and internet services, have been expanding through development of information technology. Network externalities are defined as externalities which affect a consumer's utility depending on the number of consumers who use the same product that the consumer uses [1]. For instance, a fax machine is worth little to a person if no other people use them.

In such situations, high technology does not necessarily produce value in a society because value is created through interaction among consumers. Effects of network externalities are especially good examples of those problems. Value depends on how many products are diffused in a society if network externalities pertain. Therefore, product value cannot be defined independently merely through technological specifications.

From the viewpoint of value creation, Ueda [2] presents three basic models of value creation, as follows:

- Class I value model: *Providing value* model
- Class II value model: *Adaptive value* model
- Class III value model: *Co-creative value* model

In the Class I value model, producers and consumers are defined independently of their values. In the real world, this model is applicable to mass-produced products and routine services. In the Class II value model, the consumers' objective is defined completely. However, the product and service environments are changing and unpredictable. This model is applicable to consumer-oriented products or services such as semi-order-made goods. In the Class III value model, the producers and consumers are mutually inseparable from the viewpoint of value creation. Consequently, the producers are mutually involved with consumers to co-create the value. In the real world, knowledge databases, doctor-patient medical services, and open-source software such as Linux might correspond to this kind of value model.

As an example of a Class III value model, we construct a simple decision-making model in which network externalities are presented. The model consists of a producer and two consumers. Although we might examine numerous consumers to consider a Class III value model, we instead specifically examine simple interaction among a producer and two consumers for simplification of this explanation.

For analyzing markets with network externalities, we adopt a multi-agent simulation. However, the problem is how agents behave in the simulation: selfish, altruistic, conservative, sociable, etc. It is easy to deal with problems of value creation if the model consists of economic agents who are all perfectly rational. However, actual people in the real world are not perfectly rational and have different lifestyles. For that reason, we construct a consumer agent model based on results of a lifestyle survey, and use such agents in the simulation. This study therefore examines how products are diffused through interaction between agents, and elucidates how value is created through effects of network externalities.

2 LIFESTYLE SURVEY FOR CONSTRUCTION OF CONSUMER AGENTS

2.1 Method of lifestyle survey

The survey was conducted on eight days during July 2007 using a membership questionnaire system on a mobile-phone network. A mobile research company (Point On Inc., Tokyo, Japan) built the system. For participating in the survey, the participants earned some points that are useful for shopping. In all, 8177 (4443 female and 3744 male) people in their teens to their 70s participated in survey. Most participants were residents of the Tokyo metropolitan area. Table 1 presents the number of participants in each segment according to sex and age.

Table 1: Participants in each segment according to age and sex.

	U-20	21-30	31-40	41-50	51-60	61-70	O-70	Total
Male	352	753	1441	811	320	85	0	3744
Female	757	1634	1373	442	207	17	3	4433

We asked 41 questions about daily behavior (5 questions), experience or time spent using information technologies and internet services (4 questions), personality (5 questions), values or attitudes (11 questions), interests (2 questions about favorite magazines and TV programs), and other attributes (14 questions). Questions about personality were selected based on previous studies of "Big Five Personality Traits" [3].

2.2 Data Analysis

Data were screened and organized for descriptive statistics including the frequency distribution and cross-tabulation. Spearman's rank-order correlation coefficient was used for analysis of correlation between questions. Additionally, multivariate analyses, including factor analysis, cluster analysis, and principal component analysis, were conducted.

2.3 Analysis results

We specifically addressed questions which might be widely related to consumers' purchase-decision formation. We chose these questions based on findings in social psychology related to word of mouth communication and perception of brand value [4]. We particularly examine the underlying correlation of questions and attempt to classify the participants in groups using cluster analysis. The selected questions are as follows.

Question: I see myself as someone who…

- Question 12: …easily changes mood.
- Question 15: …often searches for information about holiday plan or leisure on the Internet.
- Question 19: …often considers other people's opinions, such as word-of-mouth, on the internet.
- Question 21: ...is considerate and kind to people around me.
- Question 24: ...is selective about fashion and products.
- Question 28: ...likes to experience new things.

Table 2 shows the respective correlations between each pair of questions using Spearman's rank-order correlation coefficient. For example, people who often seek information related to holiday plans or leisure (Question 15) might tend to consider other people's opinions, such as word-of-mouth on the internet (Question 19), and tend to enjoy experiencing new things (Question 28). Furthermore, people who easily change moods (Question 12) might tend to dislike experiencing new things (Question 28). However, such a correlation analysis is insufficient to understand the types of people. Therefore, to reveal underlying groupings, we conducted the Two-Step Cluster Analysis procedure, which is an exploratory tool designed to reveal natural groupings (or clusters) within a dataset that would otherwise not be apparent. The first step is construction of a Cluster Features (CF) Tree, which has leaf nodes. Then each successive case is added to an existing node or forms a new node based upon its similarity to existing nodes. At the second step, the leaf nodes of the CF tree are grouped using an agglomerative clustering algorithm. Each of these cluster solutions is compared using Schwarz's Bayesian Criterion (BIC) to determine which number of clusters is "best". In this study, the participants were classified into five clusters. Figure 1 shows the respective fractions of extracted clusters.

Table 2: Correlation between each pair of questions.

	Q12	Q15	Q19	Q21	Q24	Q28
Q12	1	-.00	.170**	.002	.122**	-.024*
Q15	-0.00	1	.283**	.138**	.179**	.263**
Q19	.170**	.283**	1	.090**	.168**	.130**
Q21	.00	.138**	.090**	1	.185**	.173**
Q24	.122**	.179**	.168**	.185**	1	.214**
Q28	-.024*	.263**	.130**	.173**	.214**	1

* significant at 5% level; ** significant at 1%, two-tailed test

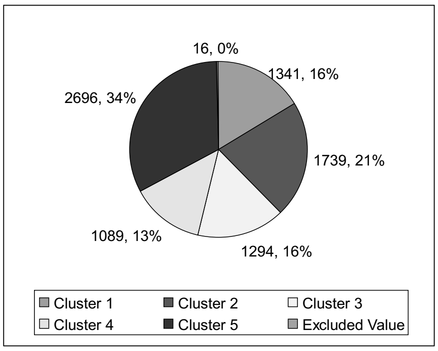


Figure 1: Shares of the clusters (Number of people, share).

Figure 2 shows the relative importance of the questions for each cluster. Each number signifies the deviation of each cluster from the average of all participants using Bonferroni's correction method. A positive number shows that the people in the cluster hold true to the question.

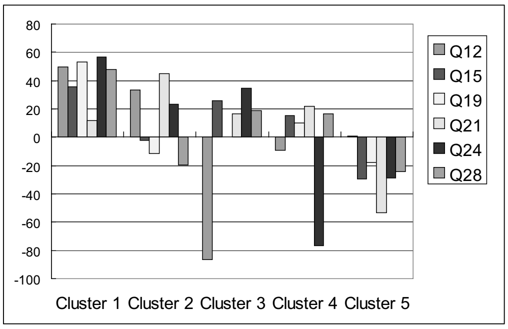


Figure 2: Degree of importance of questions for each cluster.

For example, people in Cluster 1 hold true to almost every question. On the other hand, people in Cluster 3 can be described as emotionally stable (having negative attribute to Question 12) and have normal or positive attributes to other questions. Moreover, those of Cluster 5 (the biggest Cluster) show normal emotional stability, but show negative attributes for other questions. It is difficult to understand the mechanisms of consumers' decision-making, but they might be classifiable into some types. Through these analyses, we try to construct appropriate agent models which are useful for a computational simulation of consumers' purchase behaviors.

3 MULTI-AGENT SIMULATION

3.1 Decision-making model of market with network externalities

We construct a simple market model with network externalities, in which a producer and two consumers exist. Decision-making is conducted sequentially by consumer 1, the producer, and consumer 2, as shown in Figure 3. Consumer 1 decides to purchase product A or not; then the producer makes a decision of whether to introduce product B or not. Then consumer 2 makes a decision. Consumer 2 can purchase product B only if the producer has introduced it. Also, consumer 2 makes a decision of whether to purchase product A or purchase nothing if the producer has not introduced product B.

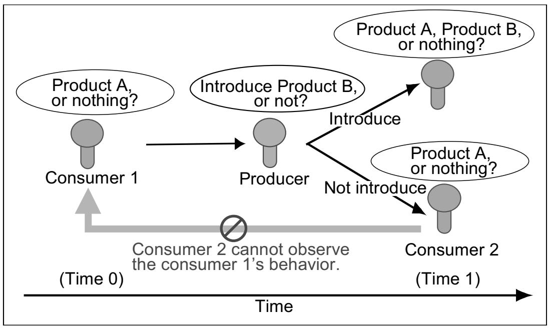


Figure 3: Decision-making model.

Consumer 1

Consumer 1 makes a decision of whether to purchase product A or nothing. Consumer 1's utility is defined as follows.

$$U_1 = \begin{cases} \sum_{t=0}^{1} \delta^t u_A(z_{A,t}) - P_A & \textit{if } \text{purchase product A} \\ 0 & \textit{if } \text{purchase nothing} \end{cases} \quad (1)$$

$$u_A(z_{A,t}) = \alpha + \beta z_{A,t}$$

Therein, P_A stands for the price of product A, δ stands for a discount factor, $z_{A,t}$ stands for the number of product A purchases at time t ($t = 0,1$), and α and β are constants. We assume that price P_A is fixed.

Consumer 2

If products A and B are on the market, consumer 2 selects one from three choices: purchase product A, purchase product B, or purchase nothing. A choice of purchasing B is removed from the choices above if only product A is on the market. Consumer 2's utility is defined as shown below.

$$U_2 = \begin{cases} \delta(u_A(z_{A,1}) - P_A) & \textit{if } \text{purchase product A} \\ \delta(u_B - P_B) & \textit{if } \text{purchase product B} \\ 0 & \textit{if } \text{purchase nothing} \end{cases} \quad (2)$$

In these equations, P_B signifies the price of product B, $z_{A,1}$ represents purchases of product A at time 1, and u_B stands for satisfaction of product B. We assume that price P_B is fixed.

Additionally, we introduce incompleteness of information: a consumer cannot observe the other consumer's behavior.

Producer

The producer decides whether to introduce product B or not. If introducing product B, product A and B are on the market; if the producer does not introduce product B, only product A is on the market. The resultant profit is represented as

$$\Pi = \sum_{t=0}^{1}(P_A - C_A)z_{A,t} + (P_B - C_B)z_{B,1} \ , \quad (3)$$

where C_A denotes the production cost of a unit of product A, and C_B denotes that of product B. Inventories are not considered and products are produced at a consumer's request. Thus, the producer incurs no loss even if no product is sold.

Network Externalities

According to equations (1) and (2), the consumer's utility depends on the number of product A, but does not depend on the number of product B. Therefore, in this model, consumers who use product A are affected by network externalities.

3.2 Decision-making rule of agents

All agents make decisions based on *Q*-learning. Agents select actions based on a *Q*-value $Q(s,a)$, where s denotes the state, and a is the action. The *Q*-value is updated as follows.

$$Q(s,a) \leftarrow Q(s,a) + \alpha[r - \gamma_r \max_{a'} Q(s',a') - Q(s,a)]$$

Therein, α is the learning rate, r is reward (profit or utility), and γ_r is the discount factor. Action a_i has the following probability:

$$P(a_i \mid s) = \frac{e^{Q(s,a_i)/T}}{\sum_a e^{Q(s,a)/T}} \ ,$$

where T stands for the computational temperature parameter. We set that $\alpha = 0.1$ and $T = 0.8$.

3.3 Two agent types implied by lifestyle analysis

Type I Agent: agent loves brand new products and depends on the network effect

Lifestyle analysis shows that Clusters 1, 3, and 4 have a positive attribute to Question 28. People in these clusters might prefer brand new products. Moreover, because attributes to Questions 15, 19, and 21 are positive, people in these clusters depend on networks and might be affected by network externalities.

Based on this inference, we construct agents with the following characteristics. The agents always prefer product B to product A: $u_B > u_A(1)$ and $u_B > u_A(2)$. In addition, the agent's utility increases if two agents use the same product: $u_A(2) > u_A(1)$. Using these constraints, we set $\alpha = 12$, $\beta = 5$ and $u_B = 45$.

Type II Agent: agent has no interest in brand new products, ignores the network effect

Lifestyle analysis shows that Clusters 2 and 5 have negative attributes to Question 28. This result represents that people in these clusters might have no interest in brand new products. In addition, these clusters show a negative attribute to Questions 15 and 19: people of these clusters might not be affected by network externalities.

Based on characteristics of these clusters, we prepare the following agents. Because the agent is uninterested in product B, the agent obtains no utility in purchasing product B: $u_B - P_B = 0$. Moreover, the agent ignores network externalities: $\beta = 0$. Consequently, we set parameters as $\alpha = 12$, $\beta = 0$ and $u_B = P_B = 27$.

3.4 Simulation results

Table 3 shows common parameters used for the simulation. Figures 4 and 5 depict the respective averaged profits (utility) of Type I agents and Type II agents. Table 4 represents the frequency of output patterns observed in 100 simulations. In the table, "A" signifies the purchase of product A, "B" stands for the purchase of product B, "N" stands for the purchase nothing, "I" stands for introduction of product B, and "NI" stands for no introduction of product B; e.g., (A, I, B) means that consumer 1 purchases product A, the producer introduces product B, and consumer 2 purchases product B.

The table shows that product A is diffused between two consumers in the case of Type I agents. In contrast, quite different characteristics are apparent for Type II agents. In this case, patterns of product diffusion are unpredictable.

Next, we compare the social surplus between two cases. Figure 6 shows the social surplus: social surplus is very low for Type II agents. Agents with complicated behavior, such as Type II agents, might reduce the social surplus.

Table 3: Parameters in simulation.

P_A	P_B	C_A	C_B	δ
18	27	10	24	0.8

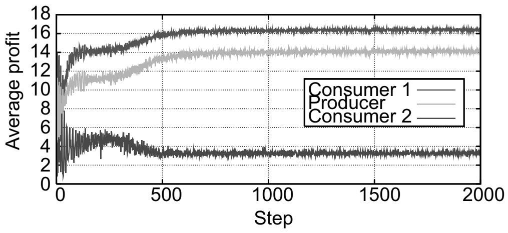


Figure 4: Average profit and utility in case of Type I agents.

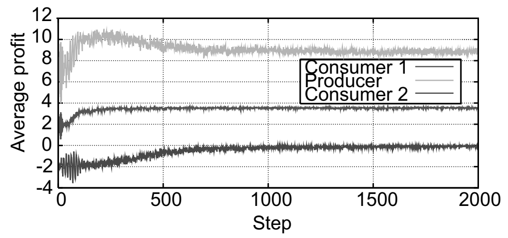


Figure 5: Average profit and utility in case of Type II agents.

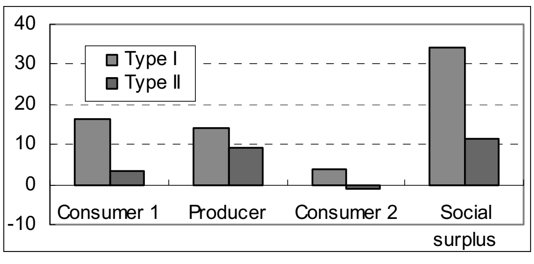


Figure 6: Social surplus in Type I and Type II simulations.

Table 4: Output patterns in 100 simulations.

Output behavior	Type I	Type II
(A, NI, A)	96	1
(A, I, A)	0	1
(A, I, B)	0	41
(A, I, N)	0	36
(A, NI, N)	4	20
Others	0	1

4 DISCUSSION: PRODUCT DIFFUSION AND VALUE CREATION

The multi-agent simulation revealed that the difference of these agent types affects the diffusion of products (Table 4).

In the Type I agent simulation, the (A, N, A) behavior occurs with high probability. Consumer 2 cannot purchase product B because the producer does not introduce it. However, in fact, consumer 2 wants to purchase product B if it is on the market because the utility of purchasing product B is greater than that of purchasing A: $U_2^A = 3.2$ and $U_2^B = 14.4$. Thus, decision-making by the producer drives products diffusion (A, N, A).

Several patterns of product diffusion are apparent in the Type II agent simulation. Because the agent has no interest in product B, $U_2^B = 0$. Consumers 2 therefore are indifferent between purchasing product A and purchasing nothing. This behavior affects product diffusion.

These results demonstrate that the social surplus is sensitive to agent behavior. Therefore, value creation in a society with network externalities depends strongly on agent lifestyle and behavior under these circumstances. Because such agents are numerous in the real world, we must examine mechanisms among producers and consumers to create value in a society.

5 CONCLUDING REMARKS

This paper presented an examination of product diffusion in a market with network externalities. Through lifestyle analysis, we constructed agents of two types for use in multi-agent simulation. Simulation results underscore that the difference of agent types affects product diffusion. Additionally, we presented discussion of value creation based on the results.

6 REFERENCES

[1] Katz, M.L., Shapiro, C., 1992, Network effects, Software provision, and Standardization, The Journal of Industrial Economics, 40/1:85-103.

[2] Ueda, K., Kito, T., Takenaka, T., 2008, Modelling of Value Creation Based on Emergent Synthesis, Annals of the CIRP, 58/1 (in press).

[3] McCrae, R., Costa, P.T., 1999, A Five-factor Theory of Personality, in Pervin, L.A. and John, O.P. (Eds.): Handbook of Personality: Theory and Research, 2nd ed., Guilford Press, New York, 139-153.

[4] Sundaram, D.S., Webster C., 1999, The Role of Brand Familiarity on the Impact of Word-of-mouth Communication on Brand Evaluations, Advances in Consumer Research, 26:664-670.

Cost Evaluation Method for Service Design Based on Activity Based Costing

Koji Kimita[1], Tatsunori Hara[2] , Yoshiki Shimomura[1] and Tamio Arai[2]

[1] Department of System Design, School of Engineering, Tokyo Metropolitan University, Japan

[2] Department of Precision Engineering, School of Engineering, The University of Tokyo, Japan

Abstract

As our economy matures, customers have begun to demand more services in addition to just industrial products. To address this problem, we require a novel engineering methodology, called Service Engineering (SE). SE aims to create value by combining services and products. SE focuses more on increasing customer satisfaction, while general service developers need to take into account economic cost in order to be successful in business. This paper proposes a method to evaluate service from the both viewpoints of customer importance and economic cost. The proposed method is verified through its application to a practical case.

Keywords:

Product/Service Systems; Design Support; Activity Based Costing

1 INTRODUCTION

Environmental problems have been quite serious over a couple of decades. To solve this problem, we should reduce the production and consumption volume of artifacts to an adequate, manageable size without making quality of our life lower than now. Consequently, we must shift to the new paradigm that aim at qualitative satisfaction rather than quantitative sufficiency, and thus the decoupling of economic growth from material and energy consumption [1]. To achieve this paradigm, products should have more values, supplied largely by knowledge and service contents, rather than just materialistic values. This dematerialization of products requires the enrichment of service contents. To this end, we need a novel engineering method to evaluate services and to design the services. This novel engineering methodology is called Service Engineering (SE), and it aims to create value by combining services and products [2]. SE differs from conventional engineering in that the design target is customer value, and the purpose is to increase customer satisfaction rather than to achieve more functional products or services.

The purpose of SE, which is currently under development, is to fulfill the requirements of customers. However, in order to be successful in business, general service designers need to take into account the economic costs. In other words, service designers should evaluate a service from the viewpoints of both customer satisfaction and economic costs. In order to serve this need, this paper proposes a method to calculate the economic cost of a service and to support service designers in finding concrete ways to reduce the service costs. First, the modeling method of SE is adopted. In SE, service contents are represented as a set of functions and entities. Then, a management accounting method known as ABC (Activity-Based Costing) [3] is applied to the SE model. ABC is a costing methodology used to trace overhead costs for cost objects such as products, processes, and departments. With respect to a service, this is an appropriate method owing to the characteristics of the service: high rate of overhead costs in the total costs. On the other hand, in SE, a method has been developed to evaluate the importance of functions from the viewpoint of customer importance [4]. This method allocates customer importance to functions by adopting two methods: QFD (Quality Function Deployment) [5] and DEMATEL (Decision-Making Trial and Evaluation Laboratory) [6]. The results of this method are used in portfolio analysis to evaluate functions from the viewpoints of economic cost and customer importance. The present method is verified through an example presented herein.

2 SERVICE ENGINEERING

2.1 Definition of service

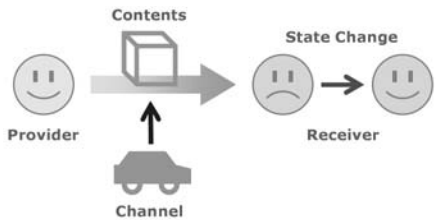


Figure 1: Definition of a service [2].

Service is defined as an activity between a service provider and a service receiver to change the state of the receiver [2]. According to the definition, a receiver is satisfied when his/her state changes to a new desired state. For the purpose of SE, the design of services must be based on the state change of a receiver. Therefore, it is necessary to find a method of expressing state changes of the receiver. States of the service receivers are represented as a set of parameters called receiver state parameters (RSPs), which represent customer value in SE [2]. All RSPs are assumed to be observable and controllable. RSPs are changed by 'service contents' and 'service channels,' as shown in Figure 1. Service contents are materials, energy, or information that directly change the receiver's state. Service channels transfer, amplify, and control service contents. The parameters expressing service contents, which influence RSPs directly, are called contents parameters (CoPs). In the same way, the parameters of service channels are called channel parameters (ChPs), which influence RSP indirectly. Thus, in SE, customer requirements are represented as RSPs, and the design of a service is based on the degree of customer satisfaction represented as the change of RSPs.

2.2 Function design of service contents

As mentioned before, the design of services must be based on customer satisfaction. Therefore, we need a modeling method that represents the relationship among customer value, service contents and service channels. In SE, a sub-model called view model is proposed to represent a functional service structure [2].

A view model represents the mutual relationships among an RSP, CoPs and ChPs. An RSP changes only according to the contents of the service received. In other words, service receivers evaluate service contents when they receive the service. Service channels are evaluated indirectly by the receiver and thus do not influence the RSP. In SE, we assume that service contents and service channels are comprised of various functions. To describe these functions, Function Names and Function Parameters (FPs) are defined. Consequently, both CoPs and ChPs belong to FPs. These functions are actualized by entities. An entity in a view model represents not only physical products but also facilities, employees, information systems and so forth. As shown in Figure 2, a view model is expressed visually using a tree structure, and thus allows service designers to obtain relationships among an RSP, functions and actual entities.

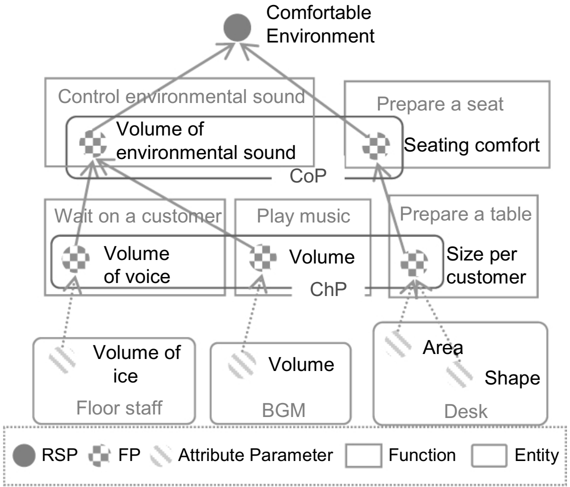


Figure 2: A simple example of view model [2].

2.3 Service activity design

A function is realized by a certain behavior of the physical structure. In this context, a service activity can be recognized as a means of actualizing a function. In SE, we propose a modeling framework to integrate the function and the service activity that was designed on the basis of the previously mentioned view models [7]. In this model, designers can describe service activities using the BPMN (Business Process Modeling Notation) model [8] according to a service blueprint [9]. The service blueprint, proposed by Shostack, is the most popular method of describing service activities in the field of service marketing. These activities are associated with functions in the view models.

2.4 Service evaluation

In SE, a method is proposed for evaluating services using QFD [5], which correspond to the view models. This method can quantify the importance of the service functions from the customer importance inputs obtained through QFD. A view model, which is represented by a graph structure, is converted into a matrix expression. A table similar to QFD is generated to appropriately reflect the Engineering Metrics of a product, i.e., a product's quality characteristics, to meet the customer's needs, or the VOC (Voice of the Customer). In addition, the DEMATEL method [6] is used to conduct a quantitative analysis by classifying the FPs into CoPs or ChPs, depending on whether or not they directly affect the RSP. Using this method, RSP importance, which is determined by the customers, is converted into FP importance. Consequently, service designers can obtain the function importance that reflects the customers' needs.

3 ACTIVITY BASED COSTING

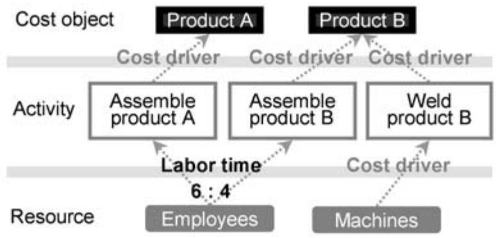


Figure 3: Relationship among resources, activities and cost objects.

ABC is a costing methodology used to trace overhead costs for cost objects such as products, processes, and departments. In ABC, the resource costs, which include the overhead cost, can be allocated to the cost objects based on the activities. Activities are the operations needed to implement tasks, and resources such as labor, electricity, and facilities are consumed to perform the activities. For example, in order to deliver a product, activities such as designing, assembly, and shipping are essential. Moreover, for these activities, the abovementioned resources are consumed.

The ABC procedure comprises two stages. In the first stage, the resource costs are associated with activities based on a cost driver. A cost driver is the criterion for cost allocation. In order to appropriately assign the resource cost to each activity, cost drivers have to be appropriately identified for each resource. For instance, the resource 'salary' may be driven by the amount of time the employee spends on an activity. In the second stage, costs are allocated to the cost objects instead of activities based on the number of activities the cost objects consume. This stage can be achieved by using cost drivers similar to the previous stage. Thus, ABC calculates the economic costs by allocating resource costs for activities. Figure 3 illustrates the relationship among cost objects, activities and resources.

4 THE PRESENT METHOD

In order to calculate the economic cost of functions, we need to identify the relationship between resources and activities, according to ABC. On the other hand, in SE, we can obtain the relationship between functions and actual entities in the services based on the view models. Here, the entities, which actualize the function, contain physical products, facilities, employees, and so on. This implies that we can identify the entities as resources in ABC. Furthermore, service activities are associated with the functions in the view models. As a result, each activity during a service delivery process is related to the entities based on the relationship between the functions and the entities. Thus, service designers can identify the relationship between entities (resources) and service activities, and calculate the economic cost of each service activity and function according to ABC.

The procedures of the present method are described in Figure 4. They begin with function design to obtain the relationship between functions and entities (resources). In this step, a view model is adopted as the method of describing the functional structure. Further, service activities are designed based on the view model. Each activity designed in this step is related to the functions in the view model. Consequently, we can obtain the relationship between service activities and entities (resources). Next, the function analysis is performed with respect to economic costs and customer importance. Finally, these results are used in portfolio analysis to evaluate functions from the viewpoints of economic cost and customer importance.

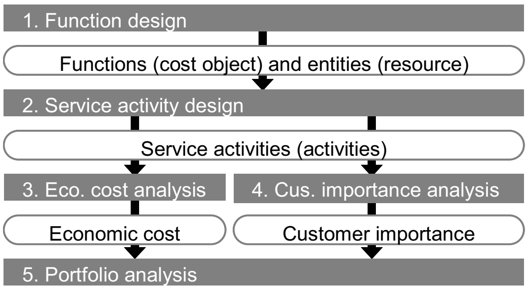


Figure 4: The proposal procedures.

5 APPLICATION

The proposed method was implemented by considering an example from the support service for the introduction of an IT system. This service involves supporting a client firm in introducing an IT system. It includes a survey of the current business process, taking decisions on the concept of IT introduction, and so on. In this application, the design-planning phase in IT introduction was considered as the scope of the application. This section explains the application by focusing on the evaluation of economic costs.

First, we implemented function design to obtain the functions of service contents and entities corresponding to resources. This was executed according to the view model. First, we determined the requirements of the client firm (corresponding to RSP) through a brainstorming session. Consequently, we enumerated three RSPs: design quality of IT introduction, planning efficiency with respect to its design, and planning quality of its design. Next, these RSPs were decomposed into functions and entities that realized the state change of the RSPs. Consequently, eighteen functions were enumerated (see Table 1). With regard to entity, we chose the employees that consist of class-A consultant, class-B consultant and system engineer. Second, the service activity design was implemented based on the view model established in the previous step. Consequently, 17 activities were enumerated (see Table 2).

Table 1: The list of the functions.

ID	Function
F1	Grasp the current situation
F2	Expose customer needs
F3	Assure the achievement of goals
F4	Design IT introduction clearly
F5	Design innovative and attractive IT system
F6	Obtain the customer's consent to the design
F7	Enhance the feasibility of the design
F8	Create the plan of IT introduction
F9	Enhance the appropriateness of IT introduction
F10	Set the price of IT introduction properly
F11	Enhance cost-benefit performance
F12	Planning the design of IT introduction quickly
F13	Lower the cost of the design planning
F14	Clarify the process to draw up the design plan
F15	Execute drawing up the design plan effectively
F16	Build organization
F17	Customer support
F18	Do precise report

Table 2: The list of the activities.

ID	Activity
A1	Survey the current situation
A2	Extract issues and problems
A3	Survey technology trends of IT
A4	Report the survey result
A5	Build organization
A6	Model the current business process
A7	Assess the current business process with the customer
A8	Report the plan on the occasion with president
A9	Model the future of the customer's business
A10	Extract issues toward the future
A11	Describe the future on the occasion with the president
A12	Make decision on the concept of IT introduction
A13	Put system modeling in execution
A14	Assess the effects of IT introduction
A15	Draw up the design plan of IT introduction
A16	Review the plan of IT introduction with customer
A17	Sum up the blueprint of IT introduction

Next, each function was analyzed from the viewpoint of economic costs by using ABC. The initial step was to allocate the resource costs to activities based on the cost driver. In this case, the resource and labor costs were allocated to service activities based on the amount of time the employees spent on each activity. As mentioned before, since service activity is a means of actualizing a function, they are associated with the functions. Therefore, the entities that are related to a certain service activity are identified from the relationship between the functions and entities, if there exist several entities in service. Next, the cost of each activity was assigned to functions. Each service activity was divided among the functions related to the service activity according to the proportion of contribution. For instance, the service activity ‘Survey the current situation’ was divided into two functions: grasp the current situation and expose customer needs. The ratio of the contribution was 0.3 and 0.7,

(Unit: 1,000,000YEN)

Activity \ Function	Activity cost	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12	F13	F14	F15	F16	F17	F18
A1	94	34	20		7			7	7	7	7				7				
A2	81		7		20		20	7	7		7				7				7
A3	35					15	5			5	5								5
A4	33			5	5		5											5	14
A5	63							19				6	6	6	6	6	6	6	
A6	134			34	34		34												34
A7	139	14	14		14	14	14	14					14				14	14	14
A8	33						17									6		6	6
A9	234				29		29		29				29	29	29			29	29
A10	166				12		36	12	59				36	12					
A11	33						7									7	7	7	7
A12	198				33				33				99						33
A13	175								44		44		44						44
A14	63				9		27		0		9							9	9
A15	175				58		58		58										
A16	33			7	7		7	7	7										
A17	148				49		49												49
Function cost (YEN)		47	41	45	276	29	307	65	244	12	71	6	228	47	49	18	27	75	250

Figure 5: Function-Activity matrix.

respectively. However, no widely authorized recommendation is provided with respect to quantifying the contribution. Thus, the allocation should be executed by a team of experts to ensure good quality of results. Finally, the economic cost of each function was calculated, as shown in Figure 5 (on the lowest line). Simultaneously, a customer importance analysis was conducted. Further, by using the results obtained from the analysis of the functions from the economic costs and customer importance, portfolio analysis was conducted, as shown in Figure 6. Depending on the area in the portfolio, the improvement strategies for each function can be generated. For instance, for an area with high economic costs and high customer importance, the improvement strategy is to realize this function by reducing its economic cost, e.g., 'Design IT introduction clearly (F4)'. For an area with high economic costs and low customer importance, the improvement strategy is to exclude this function, e.g., 'Obtain the customer's consent to the design (F6)'.

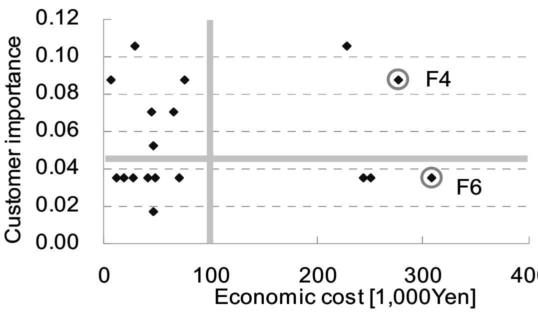


Figure 6: Portfolio analysis.

6 CONCLUSION

This paper proposed a method to evaluate service from the viewpoints of customers' demands and economic cost. Concretely, the present method enables service designers to calculate economic cost of functions of service. The application suggested that the method could support service designers with finding improvements for reducing cost.

Future works include developing well established method to quantify the contribution that is used to allocate activity costs to functions.

7 REFERENCES

[1] Tomiyama, T., 1997, A Manufacturing Paradigm Toward the 21st Century, Integrated Computer Aided Engineering, 4:159-178.

[2] Arai, T., Shimomura, Y., 2004, Proposal of Service CAD System -A Tool for Service Engineering-, Annals of the CIRP, 53/1:397-400.

[3] Cooper, R., 1988, The Rise of Activity-Based Costing-Part One: What is an Activity-Based Cost System?, Journal of Cost Management, 2/2:45-54.

[4] Sakao, T., Watanabe, K. and Shimomura, Y., 2003, A Method to Support Environmentally Conscious Service Design Using Quality Function Deployment (QFD), In Proceeding of Eco Design 2003.

[5] Akao, Y., 1990, Quality Function Deployment, Productivity Press.

[6] Warfield, J. N., 1976, Societal Systems - Planning, Policy, and Complexity, Wiley Law Publications.

[7] Hara, T., Arai, T., Shimomura, Y., Sakao, T., 2007, Service/Product Engineering: a new discipline for value production, In Proceedings of the 19th International Conference on Production Research.

[8] BPMN Information (http://www.bpmn.org).

[9] Shostack, G. L., 1982, How to Design a Service, European Journal of Marketing, 16/1:49-63.

Definition of Design Operation for Service

Koji Harada[1], Tamio Arai[1], Tatsunori Hara[1] and Yoshiki Shimomura[2]

[1] Department of Precision Engineering, School of Engineering, The University of Tokyo, Japan

[2] Department of System Design, Tokyo Metropolitan University, Japan

Abstract

Manufacturing industry has directed strong attention to services. As well as products are defined as objects consisting of functions and attributes in a conceptual design process, services are also recognized as those objects to design. Service Engineering, proposed by authors, has introduced a computational model to services, which can be transferred from a provider to a receiver. In this paper, we analyze the cases of Product-Service System by representing their models and clarify the different structures between existing product sales and PSS. As a result, design operations for describing changes of services are organized to 4 patterns.

Keywords:

Service Engineering; Design Operation; Product-Service System

1 INTRODUCTION

Manufacturing industries have directed strong attention to services with their shift from being simple sellers of products to being service providers [1]. Products and services are merging together toward a new industrial structure providing high value for customers. We have been doing research on Service Engineering [2] to evaluate and design services as a series of activities to provide value. In Service Engineering, service includes manufacturing, distribution and sales of physical products. The purpose of the study is to establish design methodology of both products and services.

As well as products are defined as objects consisting of functions and attributes in a conceptual design process, services can be also recognized as such objects to design. We have introduced a computational model to services, which are transferred from a provider to a receiver [2]. A computer-aided design system called Service Explorer [2] [3] [4] has been in development since 2002.

A provider frequently changes his/her services, for example in improvement design. Although the provider needs to execute tailored operations to the service model in this case, design operations have not been studied systematically in early literature. Mathematical definition of them is required for implementation of design operations on Service Explorer.

This paper aims at organizing design operations for describing changes of services based on case studies. We analyze cases of Product-Service System [5] and clarify the operations for describing the changes in each case. After that, the clarified operations are organized according to function and attribute model of service. The results enable us to define these operations as mathematical operations of functions and attributes.

This paper consists of 5 sections. Section 2 illustrates the modeling method of services and mentions change of services in the model. Case studies are done and the design operations are organized in Section 3. Section 4 discusses and Section 5 concludes the paper.

2 FUNCTION AND ATTRIBUTE MODEL OF SERVICE

2.1 The definition of service

A service is defined as an activity between a service provider and a service receiver to change the state of the receiver [2]. A provider causes a receiver to change from a state to a new state that the receiver desired, where both contents and a channel are means to realize the service. A service receiver is satisfied with just contents, and a service channel is used to transfer, amplify and control the service contents [2].

This definition is broader than typical definitions in the traditional management and marketing fields [6] [7]. According to our definition, most business activities are services, including selling physical products. Services to be targeted in this study correspond to PSS (designed to change the state of the receiver).

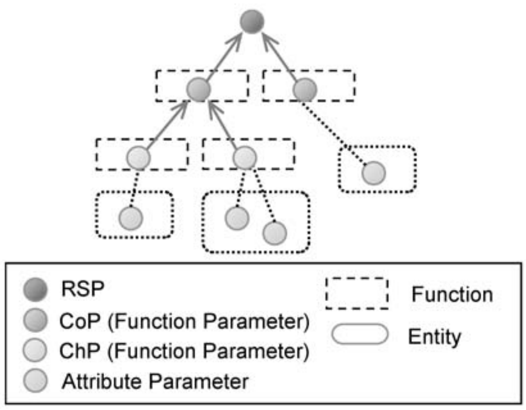


Figure 1: View Model

2.2 Service Modeling

Receiver's states are parameterized as a set of Receiver State Parameters (RSPs) [2]. RSPs represent customer value and they are indices of state changes of the receiver.

Figure 1 shows the service contents, their functions, and the corresponding "view model" method [2] [8]. After identifying the customer value as RSPs, functions and entities for each RSP are described as a view model. Functions are deployed as several sub-functions, and these functions have Function Parameters (FPs). Parameters representing the state of service contents are called Contents Parameters (CoPs), and those of service channels are called Channel Parameters (ChPs). Both CoPs and ChPs belong to FPs.

A view model works as a bridge between the customer value and actual entities using a tree structure. When detailing upper functions and constructing a functional view model structure, hardware, humanware and software are associated with the lowest-level functions as actual entities that enable these functions. These entities also have their own attribute parameters (APs). In a view model, the mutual relationships among an RSP, CoPs, ChPs and APs are expressed visually.

2.3 Change of services

A provider frequently changes his/her services according to change of promotional strategy, management strategy and other reasons. These changes of services are represented as modification of functions and entities in the view models. In some cases, a receiver may have new RSPs because the receiver sets different criteria to evaluate the changed service after the change of service delivery system. To organize these operations for representing changes of services and to clarify the patterns of operation, we analyze service cases in next section.

3 CASE STUDIES

3.1 Product-Service System

Product–service system (PSS) is a specific type of proposition that a business offers to its clients, consisting of a mix of tangible products and intangible services designed and combined so that they jointly are capable of fulfilling specific customer needs [5]. PSS is familiar with representation by functions and attributes. We analyze cases of PSS by comparing "typical product sales" with "the corresponding PSS". After that, we organize design operations for describing modification when shifting from the former type to the latter type.

3.2 Results of case studies

In this paper, we collected and analyzed 20 kinds of PSS cases in the referring literature (e.g. [9]). View models of typical product sales and PSS are compared as stated above. As a result, the operations are classified into 4 patterns as follow from the view point of modification to RSP, function and entity.

Pattern.1: **Addition of new function and entity**

Pattern.1 is the operation which adds new function and entity in view model.

In *chemical management service*, chemical company undertakes training for using chemical, waste management and other additional services as well as sales of chemical. In chemical sales, function of enabling to tighten a seal has been provided for the RSP concerning "use safely". In this PSS case, its view model for the same RSP is described as illustrated in Figure 2. The red square shows the changed structure after the operation: new function of undertaking training is added; new entities "manual" and "trainer" are also

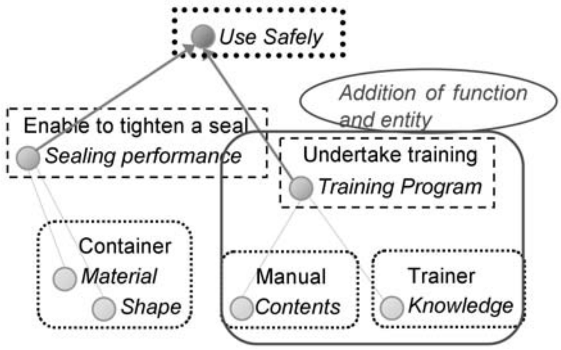


Figure 2: Chemical management service

added and related to the function.

In above case, a new function changing RSP directly is added, that is, CoP is added. Furthermore there is a PSS case in which a provider enhances the services' value by adding sub-functions, that is, ChP is added.

Remote control service of medical equipment introduces X-ray diagnosis system which diagnoses patients by remote-controlled support. This system enables patients to be examined by a distance expert operator. In typical medical service, function of diagnosing by using advanced equipments has been provided for the RSP concerning "Reliability of diagnosis". In this PSS case, its view model is described as illustrated in Figure 3. The red square shows the changed structure after the operation: new function of introducing remote-controlled system is added as sub-function of above function; new entity "remote-controlled system" is also added and related to the function. The quality of entity "operator" is

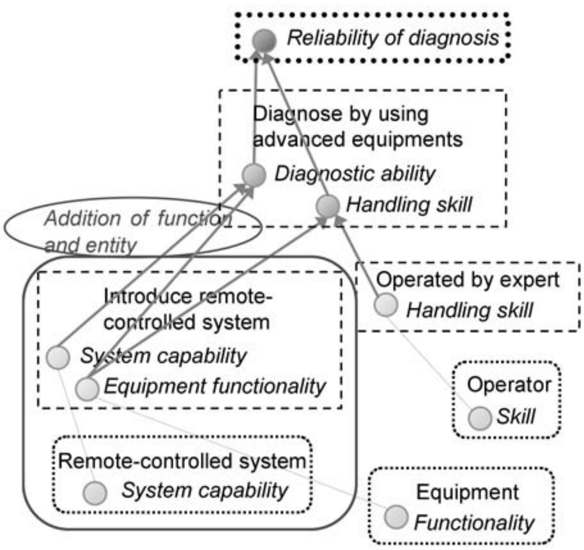


Figure 3: Remote control service of medical equipment

enhanced by adding the new function in the PSS case.

Pattern.2: **Replacement of function and entity**

Pattern.2 is the operation which replaces present function and entity with new function and entity in view model.

In *food reservation sales service*, alternate trader assumes sales of fruit and vegetables direct from the farm in place of retail store. They are delivered after reservation and their price are fixed. In retail sales, functions of holding a lot of stock and lengthening store hours have been provided for the RSP concerning "certainly available". In this PSS case, its view model for the same RSP is described as illustrated in Figure 4. The red square shows the changed structure after the operation: new functions of accepting a reservation and selling on the internet are replaced with present two functions. In this case, CoPs are replaced, but ChPs may be replaced in other cases.

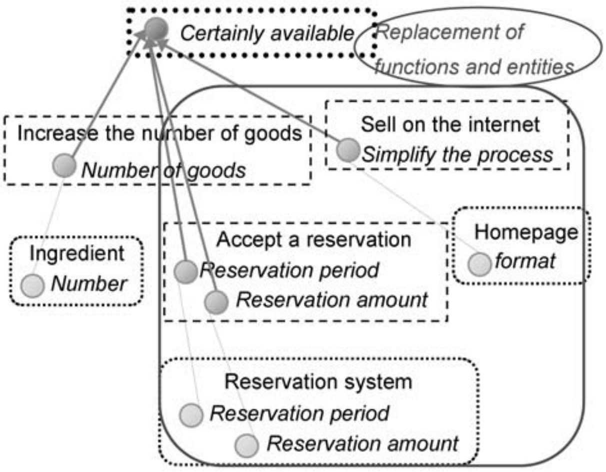


Figure 4: Food reservation sales service

Pattern.3: **Addition of new RSP**

Pattern 3 is the operation which adds new RSP and new function and entity for the RSP as new view model.

Carpet leasing service is lease service including the maintenance and cleaning of leased carpet by the carpet maker. In carpet sales, customers have RSP concerning "maintenance labor", but any functions haven't been provided for the RSP. They maintain and clean their carpet themselves. In this PSS case, its view model for the same RSP is described as illustrated in Figure 5. The red square shows the changed structure after the operation: new function of

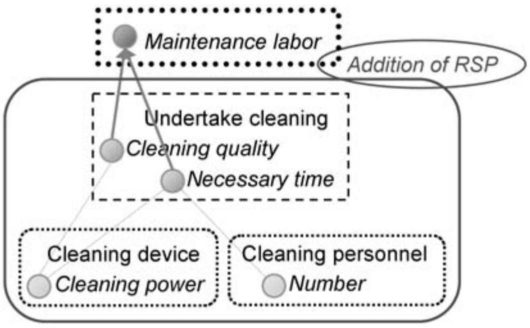


Figure 5: Carpet leasing service

undertaking cleaning is added; new entities "Cleaning device" and "Cleaning Personnel" are also added and related to the function. The previously missed RSP is satisfied in this case.

Pattern.4: **Replacement of RSP**

Pattern 4 is the operation which replaces present RSP with new RSP and adds new function and entity for the new RSP as new view model.

Car sharing is typical PSS case, which is a membership car rental system. An organization contracts car purchase and management, and a user pay a fee corresponding to the mileage. In car sales, functions of facilitating maintenance and enriching after-sales service have been provided for the RSP concerning "Risk of buying". In car sharing, its view model is described as illustrated in Figure 6. The red square shows the changed structure after the operation: new RSP concerning "risk of using" is replaced with present RSP; functions of increasing rentable/returnable place and keeping car interior clean. Customer's criteria (i.e. RSPs) for received services change because the service delivery system changes.

4 DISCUSSION

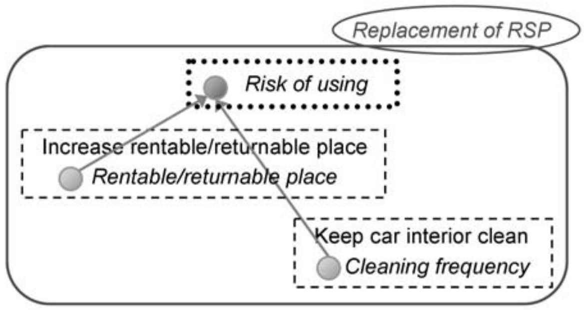


Figure 6: Car sharing

20 kinds of PSS cases analyzed in this paper have 8 cases applying to pattern 1: addition of new function and entity, 1 case to pattern 2: replacement of function and entity, 7 cases to pattern 3: addition of new RSP, 4 cases to pattern 4: replacement of RSP. As the above results, pattern 1 and pattern 3 are mainly observed. This is because PSS provides additional service element to typical product sales in many cases.

It's just conceivable that there is deletion as a design operation for modifying models besides addition and replacement. No case corresponding to the deletion is found in this research, but we need to leave open a possibility that there is *deletion of functions and entities, deletion of RSP* as design operations.

In practical flow of the services among many stakeholders, contents of the services are processed in the same manner as manufactured products. Dual roles of a receiver and a provider on one agent are unified in the service model. In addition, contents of the service in different receivers are identified according to the realization structure of the service. These operations of process also need to be defined in service design.

5 CONCLUSION

In this paper, design operations of function and entity due to changes of services are organized by the case studies of Product-Service System. As a result, design operations are classified into 4 patterns: *addition of new function and entity*; *replacement of function and entity*; *addition of new RSP*; and *replacement of RSP*. When a service designer changes his/her services, these changes can be expressed using a combination of these operations.

Future work will include as follows: defining mathematical formulation of these operations; implementing them to the computer-aided design system called “Service Explorer”.

6 REFERENCE

[1] Tomiyama T., 1997, A Manufacturing Paradigm Toward the 21st Century, Integrated Computer Aided Engineering, 4, 159-178.

[2] Arai T., Shimomura Y., 2004, Proposal of Service CAD System -A Tool for Service/Product Engineering-, Annals of the CIRP, 53(1), 397-400.

[3] Arai T., Shimomura Y., 2005, Service CAD system -Evaluation and Quantification-, Annals of the CIRP, 54(1), 463-466.

[4] Hara T., Arai T., Shimomura Y., and Sakao T., 2007, Service/product Engineering: a new discipline for value production, In proceedings of the 19th International Conference on Production Research (ICPR-19), Valparaiso, Chile, CD-ROM.

[5] Tukker A., Tischner U., 2006, New Business for Old Europe: Greenleaf Publishing.

[6] Sasser E., Olsen R.P., Wyckoff D., 1978, Management of service operations: Allyn and Bacon, Boston.

[7] Lovelock H.C., Wright L., 1999, Principle of Service Marketing and Management, Prentice-Hall.

[8] Sakao T., Shimomura Y., 2007, Service Engineering: A Novel Engineering Discipline for Producers to Increase Value Combining Service and Product, Journal of Cleaner Production, Special issue on sustainable production and consumption - making the connection, 15(6), 590-604.

[9] Goedkoop M., van Halen C., te Riele H., and Rommens P., 1999, Product Service Systems, Ecological and Economic Basics, Pre consultants: The Netherlands.

Industrial Product-Service-Systems - Typology of Service Supply Chain for IPS² Providing

Horst Meier[1], Oliver Völker[1]

[1] Ruhr-University Bochum, Chair of Production Systems, Bochum, Germany

Abstract

Today's organizational forms either focus on the production of products or of the delivering of services. That leads to the creation of supply chains by outsourcing parts from the original manufacturer to a supplier. Due to the increasing global competition companies are forced to stand out by using Industrial Product-Service-Systems (IPS²). The possibility to manage the increasing complexity by outsourcing processes to suppliers and thereby generating supply chains cannot be simply adapted to the delivery of an IPS². Thus, the configuration of supply chains in the delivery phase of IPS² requires the development of new methods and tools. With this approach a service supply chain typology for IPS² providing and the process workflow to generate IPS² networks are pointed out.

Keywords:

Industrial Product-Service-Systems; IPS²; Organization; Supply Chain; IPS² Network

1 INTRODUCTION

Today a great problem is the rapidly increasing competition in the industrial markets over the last years. Drivers like the globalization and the growing shortness of resources lead to increasing manufacturing costs and higher prizes for the offered products. To keep the competiveness the companies try to outsource their parts into low cost countries. But emerging markets like China or India are also able to offer the parts for lower prizes. This reveals the need for thinking for another solution to link the customer to the supplier. In Germany, the plant construction and machine sector has tried to develop better and more functional products by using its technical advances in order to stay competitive. But the developed products can also be provided by emerging markets with e.g. imitation. Only new ways of business relationships are able to get a higher value for the customer and better sales for the supplier. Many companies are transforming from mere product or service provider to a provider of solutions for customer problems. The offering of industrial product-service-systems (IPS²) with dynamic interdependency of products and services in the production area are transforming the traditional definition of organization into a new form of relationship between the customer and the providing company [1]. IPS² solutions are designed flexibly in regard to possible changes of individual customer needs and requirements during the use-phase [2].

New opportunities are driven out of the IPS² providing. The interdependent bundle of products and services of the IPS² generates a greater value for the customer. Not like the normal product offering with his defined functionality the IPS² includes a greater variety of functions and therefore a higher value. It is developed as an integrated solution and thus the customer should not be able to separate parts of the IPS² to get them from another supplier. And third the customization of the IPS² makes is difficult for the customer to compare it with another offered solution [3]. Providing such an individual solution is a great challenge of current organizations. New ways for measuring the prize limit of an IPS² have to be established and the information and telecommunication technology have to be used to generate a life-cycle management to make information available [4] [5]. This approach shows the difficulties to establish an optimal organization. Furthermore it shows a way to outsource the delivery of IPS² parts to suppliers efficiently.

2 TYPOLOGY OF SERVICE SUPPLY CHAIN FOR IPS² PROVIDING

2.1 Conceptual organization design

Supply chains describe the organization to move materials or products from the supplier to the customer. Therefore the supply chain can include the transformation from raw materials over components to products. In the typical supply chain the information and material flow is contrariwise directed. This traditional supply chain structure fits perfectly in the manufacturing of products but for delivering of services or IPS² some additional characteristics are necessary. Some aspects of services require another view of supply chains in the context of IPS² providing. This is for example the characterization of services itself. A service cannot be stored because of his immaterial character. Furthermore services are offered and consumed simultaneously to their production. Another characteristic of prime importance is the necessity of integrating an external factor, e.g. persons, goods, rights of the customer. In addition to these service characteristics the IPS² characteristics have to be considered as well. This could be for example the possibility to substitute products for services or contrariwise [6]. An IPS² with an automated maintenance component can be substituted in the case of failure with a manual observation and data transmission.

The supply chains of IPS² describe the extraorganizational structure to manage the delivery during its life-cycle. The supplier of the IPS² has to manage many different and complex processes during a long period of time. The characteristic of the IPS² to provide a problem solution and not a given functionality in the form of a product leads to the demand that more than one company has to join the network to fulfill all necessities. Interviews with possible customers have shown that the outsourcing of IPS² components is not of particular importance provided that one contact person is constantly available for the customer [7]. Thus, the supply chain in the meaning of an IPS² organizational structure can implement a flexible using of suppliers for the processes in the delivery phase.

This approach shows the typology of the service supply chain under the context of the different targets of delivery flows. Figure 1 shows the relation and the main elements of the IPS² supply chain. It consists of the customer of the IPS², the IPS² provider, the IPS² module supplier, the component supplier and the service supplier. In the following the different elements will be detailed.

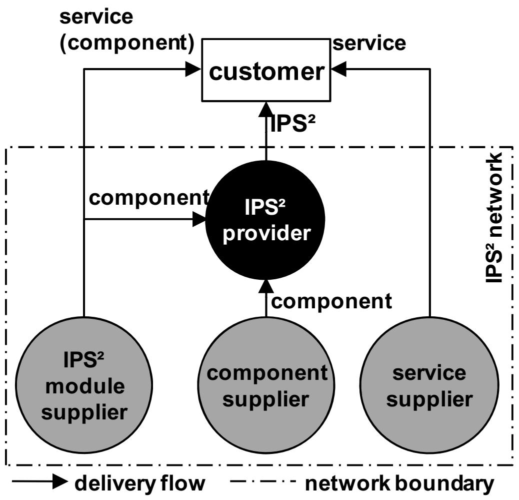


Figure 1: Typology IPS² Network Organization

Customer

The customer is the receiver of the IPS². He buys the IPS² from the IPS² provider and his requirements have to be recognized in the IPS² development and organization. Dependent on the optimal solution for the IPS² the customer sells for the delivered function or the availability or the output. The resources of the customer can be embedded in the delivery processes.

IPS² Provider

The IPS² provider has the business relationship with the customer and every communication is between the customer and the provider. The delivery flow from the IPS² provider goes directly to the customer and includes the main parts of the IPS². The IPS² provider takes all risk of the IPS². Therefore all necessary coordination is demanded by the IPS² provider. Parts of the IPS² can be outsourced to IPS² module, component or service supplier.

IPS² Module Supplier

The IPS² module supplier is responsible for a specific part of the IPS². Therefore, he delivers the processes by himself or optional with offering parts of it to sub-suppliers over the life-cycle. The components or the services can be either delivered to the IPS² provider or the customer. The IPS² module supplier shares a part of the risk and he can manage all necessary processes without coordination with the IPS² provider.

Component Supplier

The component supplier delivers a tangible part, a component of IPS², to the IPS² provider. This supplier does not have the prefix IPS² and is therefore not responsible for delivering the component over the life-cycle.

Service Supplier

This supplier delivers service related parts of IPS². This could be a service share for which the service supplier has technological expertise. Thus, the characteristics of services and IPS² make it necessary that the service flows directly to the customer.

2.2 Demands of the IPS² organization and the supply chain for IPS² providing

The longer lasting responsibility for their offered solutions and the willingness to react on changing demands of the customer leads to high demands for the IPS² organization and the belonging supply chain. The most important demand for an IPS² providing organization and this includes the extraorganization is to manage the dynamic of the IPS² itself. A new organization has to adapt the processes in the delivery phase to provide the required resources in the desired time and location with a flexible and changeable structure. The need for a flexible and changeable structure does not exist in traditional organization. The product development is mostly limited to the existing manufacturing technology. Efficient usage of the existing capacity is more important than the best fulfillment of the desired solution of the customer. Therefore, a flexible assignment of the IPS² parts on network partners has to be developed. In this regard each network partner is evaluated and chosen due to its performance fulfillment. As a result the assignment of different parts can be changed dynamically. Growing scopes of delivery processes and rising amount of provided IPS² are influencing the intra- and extraorganization as well. In the following an approach to manage the dynamic of the IPS² with an autonomous IPS² organization concept is presented.

3 AUTONOMOUS IPS² ORGANIZATION

3.1 IPS² network organization concept

The typology can be used to build up a new network organization concept for the delivery of IPS². The initial demand to build up an IPS² organization is the responsibility of the provider for the IPS² over the life-cycle to achieve a high customer satisfaction. Thus the permanent adaptions to the customer demands and the making of necessary repairs or optimizations have to be managed. The tasks for the IPS² organization are dependent on the IPS² and are mainly fixed with the development. The advantage of the IPS² is the solution cluster that gives a variety of acting to deliver the same result. Thus, every task of the IPS² can be managed with a supply chain in the form of the IPS² provider the supplier and the connection between them. The IPS² provider is the central point who manages to deliver all necessary jobs. He answers the question for the best fitting partner for the job deliver. The necessary connections for this will be created autonomously with the software agent concept as shown in figure 2. Scholz-Reiter defines the term autonomously as 'the independence of a system in making decisions by itself without external instruction and performing actions by itself without external forces' [8]. Two different kinds of software agents, job-agent and offer-agent, are implemented to build up new network connections. The job-agents are assigning possible jobs to the network partner and the offer-agents are checking the feasibility and further the calculating of costs. A supplier who does not have the required resources can offer critical parts to the members of the network also with an own job-agent. In this way an optimal allocation of the job can be found. The ideal solution is to create automatically a supply chain structure only with offering the job by the IPS² provider. The premise of it is that all agents have to get access to every data needed. This could be the production planning and control data (PPC) or the enterprise resource planning data (ERP). Thereby job- and offer-agents are working autonomously without external control. To implement and program the agents their process workflows have to be designed.

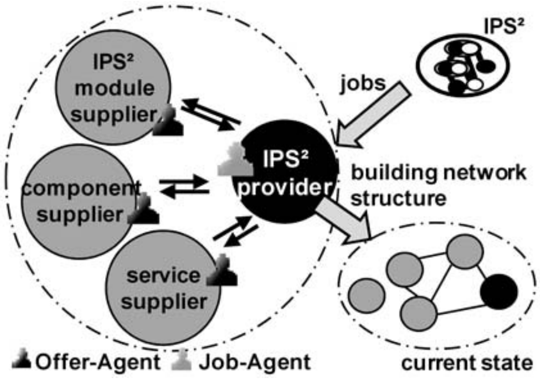
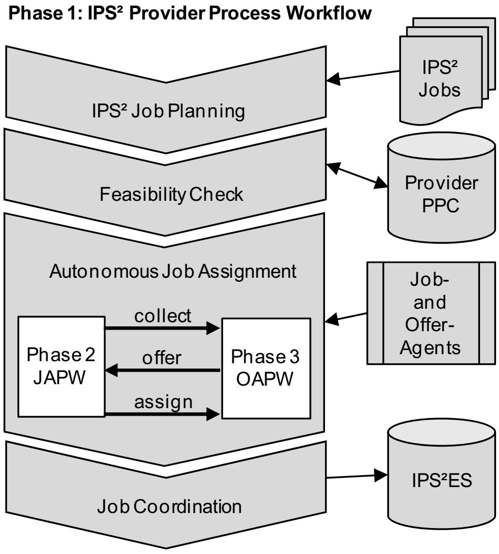


Figure 2: Concept of Autonomous Network Organization

3.2 Workflow management

The workflow management for the IPS² network organization can be subdivided into three different phases. Phase 1 explains the process workflow of the IPS² provider. Phase 2 describes the assigning of jobs to IPS² network partner and it runs parallel with phase 3 which checks the feasibility of the supplier for the job delivery. The relations of three phases are illustrated in figure 3. Phase 1 is surrounding the other phases. Phase 2 and phase 3 are running during the process step 'autonomous job assignment' of phase 1.

Phase 1 starts with the need of planning and controlling of IPS² jobs. The provider will separate these jobs by determining a responsible partner to each specific part of IPS². Reasons for outsourcing these parts could be the inexperience with the technology or a lack of resources or contractual agreements between the IPS² provider and the IPS² module supplier. The parts without his responsibility do not have to be considered in the following steps. Next the provider will check his technical and organizational ability to deliver the job. The necessary data comes out of the PPC of the IPS² provider. Possible criteria for unfeasibility can be e.g. quality demands or cost limits or resource reservation from other processes. The not feasible parts of the IPS² have to be assigned to the network partners. Step 3 explains the assignment of jobs with autonomous software agents. In this step the phase 2 and phase 3 are embedded. The job- and offer-agents are part of a program pool and can be initiated when necessary. A description of the phases is following. The fourth step is necessary to ensure the coordination of the IPS². All information about the jobs and which partner is responsible for their delivery is collected in the IPS² execution system. A supervisory and control regarding specific problems of the IPS² Execution System is to be made amenable to the IPS² provider.



Legend: IPS² provider process workflow (IPS²PPW), Production Planning and Control (PPC), job-agent process workflow (JAPW), offer-agent process workflow (OAPW), IPS² Execution System (IPS²ES)

Figure 3: IPS² Provider Process Workflow

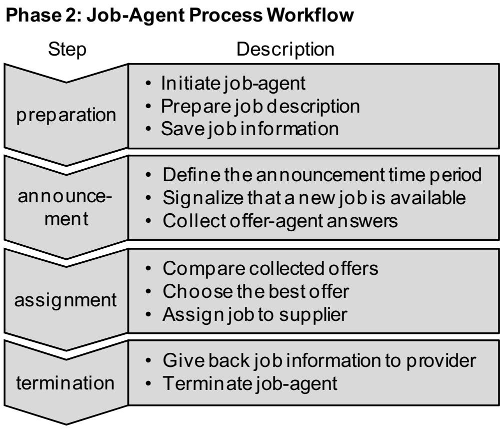


Figure 4: Job-Agent Process Workflow

A detailed description of the job-agent process workflow is illustrated in figure 4. The phase 2 includes the main process steps: preparation, announcement, assignment and termination. With beginning of the former step 'autonomous job assignment' in the IPS²PPW every job initiates its own job-agent. The life time of a job-agent ends with the final process step termination. Working with heterogeneous suppliers makes it necessary to prepare the job description for an automated using of the agents. Thus a similar understanding of the delivery for every member in the network is enabled. In the description of the deliverable part are the following information comprised:

- For which IPS² is the job necessary and how is it identified?
- A detailed description of the deliverable processes and the belonging resources.
- What is the internal transfer prize?
- What is the target destination for the delivery?
- What are the quality demands?
- At which point of time has the job been delivered?

The announcement step starts after the preparation step. There the offered job will rest for a limited period of time until he will be assigned. During this time the offer-agents have to work. Information can be exchanged and the leading prize offer is visible for the competitive offer-agents. During the step assignment the job-agent is comparing the offered prizes and is choosing the best offer based on cost, time, quality and risk criteria. The chosen supplier and the IPS² provider will be informed of the assignment. The work of the job-agent is done and it will be terminated. Therefore a conflict with other jobs will be prohibited.

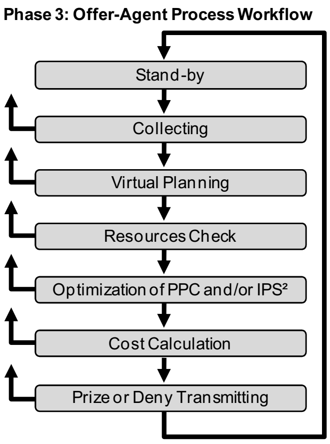


Figure 5: Offer-Agent Process Workflow

Figure 5 explains the third phase, the offer-agent process workflow. The offer-agents can be software programs that are continuously working for every network partner. It is also possible to use human resources in a first step. If no job is offered the offer-agents remain in stand-by status. The work begins when a new job is detected in the IPS² network. Then the offer-agent collects the new information of the job. To check the feasibility and make a cost calculation possible the steps virtual planning, resource check and an optional optimization are necessary. The virtual planning integrates the new job in the PPC of the supplier. Therefore problems like missing or reserved resources can be detected within the step resource check. The autonomous aspect of the members in the IPS² network makes it possible to solve the resource problems in two ways. The first way is to shift non-IPS² processes in the planning and scheduling table. The second way is using the IPS² specific resource planning methods, that allows e.g. the substitution of product and service or the use of time variance or check for alternative resources from the network. In the case of a positive planning of the job the offer-agent can calculate the costs of the delivery. After that it transmits the prize for the delivery with his supplier or if the supplier is incapable it transmits a deny signal to the job-agent. In the last step the offer-agent changes his status from working to stand-by. One important point for the calculation is the competition of different network partners for the same job. The offer-agents have access to the offered prizes and can react with underbidding with their prize offer. If a lower prize is offered the information of the job-agent will be updated and the offer-agent will awake out of his stand-by modus.

4 CONCLUSIONS

It has been shown that the delivery of IPS² is a challenge for traditional organizations and that a new approach for the organization structure is necessary. Therefore, the roles of the network members require a new understanding. The new typology described different types in the network and can be used to manage the delivery processes. The autonomous concept has been emerged to be the best fitting approach. A real-time reactive IPS² network can be established under consideration of all possible circumstances, if the shown agent concept and the process workflows can work autonomously.

5 ACKNOWLEDGMENTS

We express our sincere thanks to the Deutsche Forschungsgemeinschaft (DFG) for financing this research within the Collaborative Research Project SFB/TR29 on Product-Service-Systems – dynamic interdependency of products and services in the production area.

6 REFERENCES

[1] Meier, H., Uhlmann, E., Kortmann, D., 2005, Hybride Leistungsbündel – Nutzneorientiertes Produkt-verständnis durch interferierende Sach- und Diesntleistungen, wt Werkstattstechnik online, Jahrgang 95 H.7/8, 528-532.

[2] Meier, H., Kortmann, D., 2007, Leadership – From Technology to Use, 14th CIRP Conference on Life Cycle Engineering, 159-163.

[3] Johansson, J.E., Krishnamrurthy, C. Schlissberg, H.E., 2003, Solving the solutions problem, The McKinsey Quaterly, Number 3, 117-125.

[4] Rese, M., Otte, H. O., Strotmann, W. C., 2007, Industrial product service systems as a problem solution to meet future challenges in B-to-B markets -a value based approach, In: Proceedings of the International Conference on Business Market Management at the TU Delft.

[5] Abramovici, M., Schulte, S., 2006, Lifecycle Management für hybride Leistungsbündel (HLB), wt Werkstattstechnik online, Jahrgang 96 H. 7/8, 467-471.

[6] Meier, H., Krug, C.M., 2006, Strategische Kapazitätsplanung hybrider Leistungsbündel, PPS Management 11, 48-51.

[7] Stremersch, S., Wuyts, S., Frambach, R.T., 2001, The Purchasing of Full-Service Contracts: An Exploratory Study wihtin the Industrial Maintenance Market, Industrial Marketing Management 30, 1-12.

A Method of Supporting Conflict-Solving for Service Design

Yoshiki Akiyama[1], Yoshiki Shimomura[1] and Tamio Arai[2]
[1] Department of System Design, Tokyo Metropolitan University, Japan
[2] Department of Precision Engineering, The University of Tokyo, Japan

Abstract
Recently, the importance of "service" has been emphasized in various industries. However, few researches have focused on service design in spite of its great importance. This paper proposes a methodology of supporting service design, by solving conflicts in service design solution. The proposed methods enable service designers to find out the existing conflicts in service design solution, and obtain the basic strategies to solve them on computers. The verification of the proposed methods is carried out, applying it to the existing service cases.

Keywords:
Service Engineering; Design Support; CAD

1 INTRODUCTION

Service is recognized as the important element in many industries, and it's imperative in order to make industries more competitive [1]. In spite of its importance, there are few researches on service design or its evaluation (e.g., [2]). Service Engineering (SE) [3] provides a methodology from the view points of engineering and makes it possible for service designers to come up with the design solution without ad hoc trial and error. Based on the above understanding, the authors have been developing an integrated design-support tool of product and service, called Service CAD system [4]. However, in general, design solutions obtained with the current Service CAD frequently have various kinds of conflict, and such conflicts make it difficult to practice services. Namely, the quality and efficiency of design largely depend on how rapidly designers can find out and solve such conflicts in design solutions.

In this paper, the authors propose a methodology of supporting design, which enables designers to find out the existing conflicts in design solutions and obtain the basic strategies to solve them on computers. Two different approaches for detecting conflicts are proposed; one is the approach with lexical expressions of functions, and the other is the approach with the ranges of design parameters. The methods for solving the detected conflict with TRIZ [5] methodology is also suggested. The use of TRIZ methodology helps the service designers to come up with the solution easily. The verification of the proposed method is carried out, applying it to the existing two service cases.

2 EXISTING RESEARCHES ON CONFLICT-SOLVING

TRIZ is one of the well-known methodologies, which provides the methods to detect and solve conflicts in product design field. TRIZ consists of various methods and the knowledge base grounded on the past cases of invention, which enable designers to solve problems effectively. For example, Technical Contradiction Table [5] suggests the principles to solve conflicts between elements in design solution. The principles are created on the basis of the analysis of millions of patents.

There are some interesting researches on applications of TRIZ to service design. Two of the researches are reviewed and their effectiveness is analyzed in this chapter. One is the research, where effective usages of TRIZ tools in order to tackle the conflicts in service design are discussed [6]. In this research, some TRIZ tools (e.g., Problem Formulator [7], 40 inventive principles [5], Separation Principles [5]) consists in entire conflict-solving process. The conflict-solving process is very useful in that a simple flowchart of conflict-solving process with TRIZ tools is shown. In this research, however, the way of service-modeling is ambiguous and largely dependent on each designer.

The other is the research in which the modification of 40 inventive principles of TRIZ for service design is suggested [8]. In this research, the new principles are proposed by matching up the existing principles. For example, the principle 'Combination-Separation' is described as the following statement; 'Combine parts of an object or the phases of a process to form a uniform object or process. Separate a uniform object or a uniform process to form independent parts or phases'. However, in this research also, the lack of defined service-modeling method makes it difficult to achieve conflict-solving. In addition, the effectiveness of the new principles is left to be verified.

Therefore, we believe that the operation of conflict-solving have to be done based on the defined service-modeling. In the following sections, the definition and modeling of service in SE is presented. The model of service is a target of the operations of conflict-solving proposed in this paper.

3 SERVICE MODELS IN SERVICE ENGINEERING

3.1 Components of service

In Service Engineering, a service is defined as an activity by a service provider to change the state of a service receiver

The 41st CIRP Conference on Manufacturing Systems, 2008

[9]. A service is delivered by means of service contents and service channels. While service contents directly change the receiver's state, service channels transfers, amplifies, and controls service contents and indirectly influences the state change of a service receiver. Service contents and channels make up the realization structure of a service. Therefore, designers need to reinforce both service contents and channels in order to design competitive service.

In Service Engineering, it is assumed that a service and the state change of a service receiver can be expressed as the combination of parameters that represents the service and the relationships among the parameters [4]. A receiver's state is represented by a set of Receiver State Parameters (RSPs). Contents and channels are expressed by Contents Parameters (CoPs) and Channel Parameters (ChPs) respectively.

3.2 Sub-models of a service: view model

In Service Engineering, various sub-models of service are proposed. Among them, view model is treated as a target of conflict detection in this paper.

A view model describes a functional structure to realize a change in an RSP and expresses a part of the realization structure of a service through the relationship between the RSP and the functional structure, described in the form of the functional relations among the RSP, CoP, and ChP [4].

The functions of channels and contents are expressed by function names (FNs) as lexical expressions and function parameters (FPs) as target parameters of functions. Each function is related to another. The FPs that are directly related to RSPs are recognized as CoPs, and those indirectly influencing RSPs are ChPs. The lowest functions, which have been deployed enough, are related to Entities. An entity is what exists in the real world. An entity includes not only a product but also a person and organization. An entity has one or more attribute parameters (APs). Figure.1 is an example of view model which describes a part of realization structure of a coffee shop service.

4 DEFINITION OF CONFLICT

In order to develop an argument on conflict, 'conflict' needs to be clearly-defined. In this paper, a conflict is defined as a state; 'When a designer improves a part of service design solution, another part of it is deteriorated'. In the following two chapters, two different methods to detect and solve conflicts are explained.

5 AN APPROACH OF CONFLICT-SOLVING USING LEXICAL EXPRESSIONS OF FUNCTIONS

5.1 Conflict detection with lexical expressions of functions

This is one of the conflict detection methods. In this method, conflicts in view models are detected, by using the lexical expressions of functions. In general, a lexical expression of a function consists of "object" and "predicate". Therefore, conflict detection is executed, analyzing these elements from the lexical point of view. For example, two functions 'Increase staff' and 'Decrease staff' have the anonymous predicates each other, whereas these functions have the same objects. It is usually impossible to increase and decrease the staff at the same time, so the conflict between these two functions is detected. In another case, two functions 'Increase staff' and 'Decrease employees' have the antonymous predicates, whereas these functions have the synonymous objects. The conflict between these functions is also detected.

The biggest advantage of using lexical information of functions is that conflicts can be detected in earlier phase, where the rough concept of the service is determined. In this phase of design, it is rare for designers to describe the parametric information of the service and product. This method makes it possible for designers to find the conflicts before the phases, where more detailed specification or providing process is determined. Earlier detection saves troubles and additional cost in the later phases.

5.2 Building up lexical relation database

In order to detect the conflicts with the above method, the frequently used vocabularies, which are in synonymous and antonymous relations, need to be accumulated and ready to be searched. So a database (Lexical Relation Database) including such information needs to be built up on the computer. However, it is almost impossible to accumulate all these vocabularies in it, for a trillion vocabularies which are in such relations exist. Therefore, the database should be

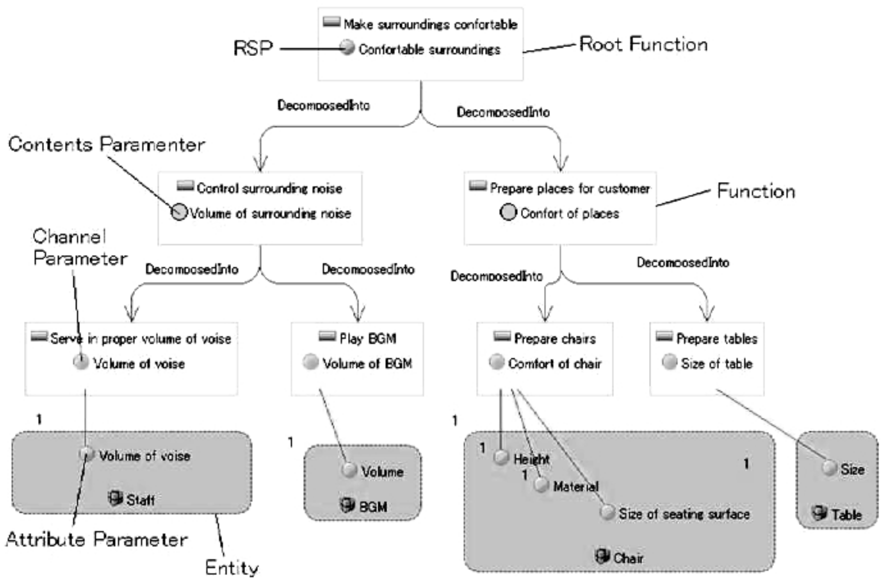


Figure. 1: View model of a coffee shop service.

expandable, adding new vocabularies by designers or a system administrator optionally.

The above-mentioned requirements for Lexical Relation Database are fulfilled by building it up with the schemes of RDF [10] and OWL [11], which are used extensively in the research field of Semantic Web. A property [10] makes it possible to describe the vocabularies and relation between them. In this detection method, two properties are additionally defined by the authors; 'IsSynonymOf' and 'IsAntonymOf' for the synonymous and antonymous relation respectively. 'IsSynonymOf' is defined to be a transitive property [11]. Thanks to this definition, for example, when a pair ('staff' and 'employee') is an instance of IsSynonymOf, and another pair ('employee' and 'worker') is an instance of IsSynonymOf, then the pair ('staff' and 'worker') is also inferred to be an instance of IsSynonymOf from the existing information in the database. Figure.2 is a screen shot of Lexical Relation Database.


```
:
<Predicate rdf:ID="decrease">
  <IsSynonymOf>
    <Predicate rdf:ID="diminish">
      <IsSynonymOf rdf:resource="#decrease"/>
      <IsSynonymOf>
        <Predicate rdf:ID="lessen">
          <IsAntonymOf>
            <Predicate rdf:ID="multiply">
              <IsSynonymOf>
                <Predicate rdf:ID="enlarge">
                  <IsAntonymOf rdf:resource="#decrease"/>
                  <IsSynonymOf rdf:resource="#increase"/>
                </Predicate>
              </IsSynonymOf>
              <IsAntonymOf rdf:resource="#decrease"/>
              <IsAntonymOf rdf:resource="#lessen"/>
            </Predicate>
          </IsAntonymOf>
          <IsSynonymOf rdf:resource="#diminish"/>
        </Predicate>
      </IsSynonymOf>
      <IsSynonymOf>
        <Predicate rdf:ID="minify">
          <IsSynonymOf rdf:resource="#diminish"/>
        </Predicate>
      </IsSynonymOf>
    </Predicate>
  </IsSynonymOf>
  <IsAntonymOf rdf:resource="#enlarge"/>
  <IsAntonymOf rdf:resource="#increase"/>
</Predicate>
:
```


Figure. 2: A screen shot of Lexical Relation Database.

5.3 Process of conflict detection

Figure.3 is the flowchart of conflict detection using the above-mentioned Lexical Relation Database. Detail of conflict detection process is explained below.

The lexical expressions of the functions in the design solution (or a part of it) are compared by pairwise comparison. For each pair of the functions, the following process is executed. The numbers correspond to those in Figure. 3.

1. Determine whether the two functions have the same object.

It is determined whether or not the two functions have the same object. For instance, two functions 'Increase staff' and 'Decrease staff' have the object 'staff' in common.

2. Determine whether the two functions have the object in a synonymous relation.

When the two functions don't have the same object in common, it is determined whether or not these functions have the object in a synonymous relation by accessing the Lexical Relation Database.

3. Determine whether the two functions have the predicates in an antonymous relation.

When the two functions have the same or synonymous objects in common, it is determined whether or not these functions have the antonymous predicates by accessing the Lexical Relation Database.

4. Alert a designer to conflict.

When a conflict is detected in the above processes, the conflict is alerted to the designer.

5.4 Conflict-solving with TRIZ tool

In order to solve a conflict detected with the above method, Separation Principles in TRIZ is useful. Separation Principles are the ways a designer solve a self-conflict in TRIZ. When two different states are demanded for a parameter, the states can be separated in the view point of time or place with these principles. Separation in time, for example, provides a solution that a parameter is in state A in a time and state B in another time.

The two functions with a conflict detected in the above-mentioned method, have the same (or synonymous) object and antonymous predicate. One function requires the object to be in state A, while the other function requires it to be in state B. A predicate behaves as an operant which requires an object to be in a state. Therefore, the demands required by these two functions cannot be fulfilled at the same time usually. The use of Separation Principles, however, makes it possible that an object is in state A in a time and in B in another time.

5.5 Verification experiment of the proposed method

In this section, a verification experiment for the methods proposed in this chapter is conducted. The methods are applied to a greatly-simplified realization structure of a restaurant service described by the authors. The conflict detection method is executed for the realization structure. In the realization structure, a function 'Increase time for customer-care' is required to fulfil the upper function 'Give finely-textured service', while another function 'Decrease time

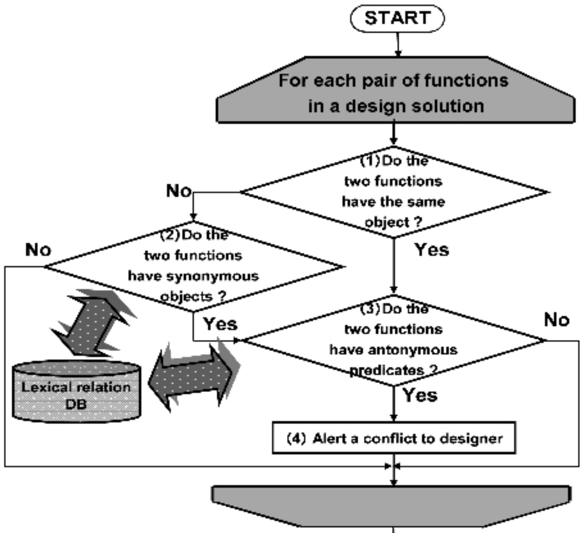


Figure.3: Flowchart of conflict detection using Lexical Expressions of Functions.

for customer-service' is required to fulfil another upper function 'Respect the customers' private time'. The process of the conflict detection between functions is as follows.

First of all, it is checked whether 'Increase time for customer-care' and 'Decrease time for customer-service' have the same object. But, the two functions don't have the same object. So it is checked whether or not two functions have the objects in a synonymous relation. If the Lexical Relation Database contains a synonymous relation between "customer-care" and "customer-service", it is determined that these functions have the synonymous objects. Second, it is checked whether or not two functions have the predicates in an antonymous relation. If the Lexical Relation Database contains an antonymous relation between 'increase' and 'decrease', it is determined that these two functions have the antonymous predicates. As a result of the above process, a conflict between 'Increase time for customer-care' and 'Decrease time for customer-service' is successfully detected.

Next in the solving step, the detected conflict is solved with Separation Principles in TRIZ. In this case, a separation by time seems helpful to solve it. For example, a time can be divided into 'the time before taking order' and 'the time after taking order'. It can be a solution that 'Increase time for customer-care before taking order' and 'Decrease time for customer-service after taking order'.

6 AN APPROACH OF CONFLICT-SOLVING USING DESIGN PARAMETERS' RANGES

6.1 Conflict detection with the ranges of design parameters

This is another conflict detection method proposed in this paper. In this method, the conflicts between functions in view models are detected by using the ranges of design parameters, while the lexical expressions are utilized in the method proposed in the previous chapter. Take a part of realization structure that a parameter influences multiple upper parameters of the functions for example, as shown in Figure.4. The information of the ranges of parameter C is used in order to detect a conflict between function X and function Y. If there isn't shared value between these two ranges, the conflict is alerted to the designer.

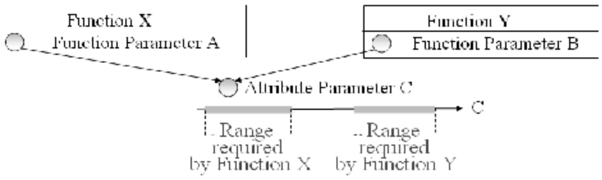


Figure.4: Inconsistency between parameter ranges.

6.2 Range calculation with Set-based Theory

The range required to develop the upper functions need to be calculated for each parameter which influences multiple functions. Thus, Set Based Theory [12], which is a theory that formulates parametric operations with ranges, is utilized. In Set Based Theory, the variation of each variable in a system is expressed as a set. And, the set is expressed as an interval. The influence of each parameter's variation on other parameters is calculated with the rule of interval operation. Set Based Theory consists of four operations, but among them, RangeOperation and DomainOperation are utilized in this method. The former is an operation that the range of an unknown dependent value is calculated from the ranges of multiple independent valuables, while the latter is an operation that the range of an unknown independent valuable is calculated from a dependent valuable and multiple independent valuables. Set Based Theory is introduced in some researches on product design methodology (e.g., [13]).

6.3 Process of conflict detection with Set Based Theory

Figure.5 is a flowchart of conflict detection using Set Based Theory. Detail of the conflict detection process is described in this section with an example of a part of the realization structure (Figure.6), where a parameter influence on multiple upper parameters of the functions. For each of such structures, the conflict detection method is executed through the below process. The numbers correspond to those in Figure. 5.

1. Append the desired ranges to parameters of functions.

The desired ranges to parameters of functions, parameter A

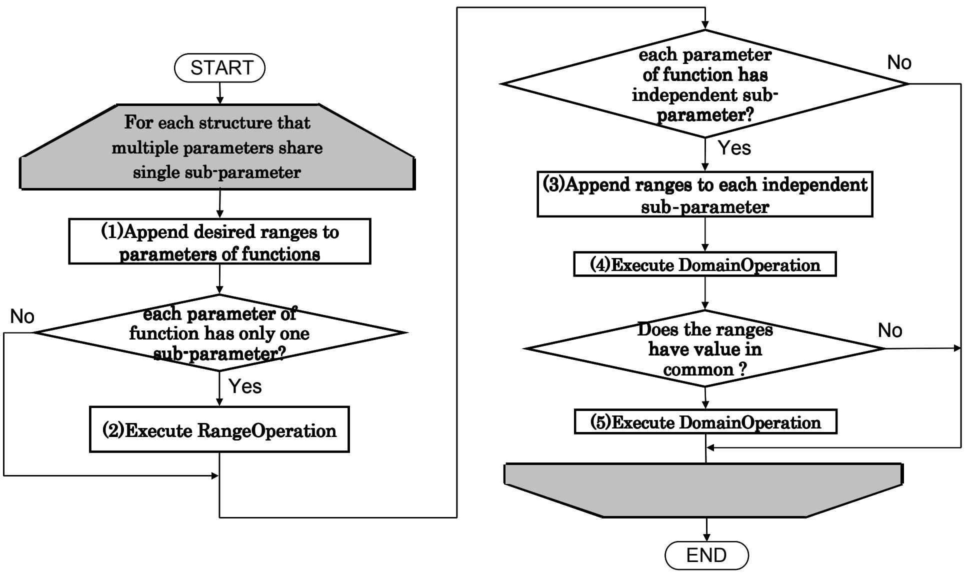


Figure.5: Flowchart of conflict detection using ranges of design parameters.

and B, are appended.

2. Execute RangeOperation.

The range of parameter D, which influences multiple upper functions, is calculated with Set Based Theory. In this case, the value of parameter B is determined by only parameter D. Therefore, the range of parameter D, which is required to put parameter B within the desired range, is calculated by using RangeOperation.

3. Append ranges to each independent sub-parameter.

In this case, RangeOperation cannot be utilized to calculate the ranges of parameter D, for the value of parameter A is determined by more than one parameter, C and D. Thus, the range of independent parameter C is determined preliminarily.

4. Execute DomainOperation.

DomainOperation is executed in order to calculate the range of parameter D required to put parameter A within the desired range.

5. Alert a designer to the conflict.

In process 1 to 4, two different ranges for parameter D are calculated; one is the range required to put parameter A within the desired range and the other is the range required to put parameter B within the desired range. If these two ranges don't have value in common, the self-conflict of parameter D (in other words, the mutual conflict between function X and Y) is alerted to designer.

6.4 Conflict-solving with TRIZ tool

In order to solve a conflict detected with the above method, there are two possible solutions. One is almost the same method as used in the previous chapter. In this method, a detected conflict is solved, by separating in terms of time or place. Separation in terms of time, for instance, enables for value of a parameter to be in a time, and in another time. The other solution is the method with Technical Contradiction Table in TRIZ methodology. Detail of this method is explained below with Figure.6.

1. Determine the parameters influenced by a parameter.

The use of Set Based Theory has found that parameter D is determined to be conflicting; it is required to be in mutually exclusive ranges simultaneously. Parameter A and parameter B, which are influenced by parameter D, are utilized in this conflict solving method.

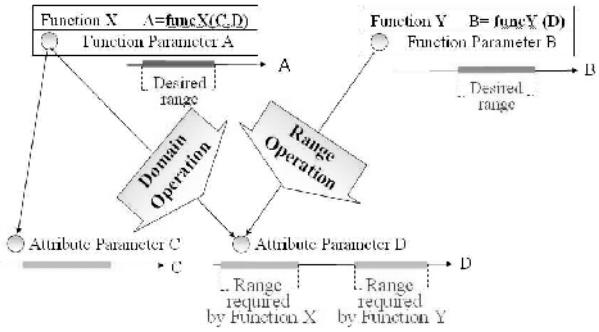


Figure.6: Detection of inconsistency between parameter ranges.

2. Apply each parameter to one of the 39 attributes.

In Technical Contradiction Table, a list of 39 attributes is defined – such as size, weight, ease of manufacture. Technical Contradiction Table is a 39 by 39 grid matching each attribute to another. In each cell, there is a list of inventive principles. Therefore, in order to use this table, each conflicting parameter needs to be applied to one of the 39 attributes.

3. Use Technical Contradiction Table.

Possible inventive principles are identified, by matching one parameter to another. If there is a useful principle in a cell, it is utilized to solve the conflict by a designer.

6.5 Verification experiment of the proposed method

In this section, a verification experiment for the method proposed in this chapter is conducted. The method is applied to a realization structure of house-cleaning service described by the authors. As a part of it, functions of a vacuum used in the service is described as Figure.7. In Figure.7, there are two functions; 'Increase Suction' and 'Lower Power Consumption'. It is assumed that a parameter of 'Increase Suction' is formulated as formula (1), and a parameter of 'Lower Power Consumption' is formulated as formula (2).

$$F = \frac{0.974 D_t^6 N^2 \phi^2}{\beta} \qquad (1)$$

$$E = \frac{\pi N T}{30\alpha} \qquad (2)$$

The conflict detection method is executed for the realization structure. The process of the conflict detection is as follows. The desired ranges to parameters of functions, 'Power Consumption' and 'Efficiency of Suction', are described as F = [0, ∞] and E= [0,300] respectively. Next, the ranges of

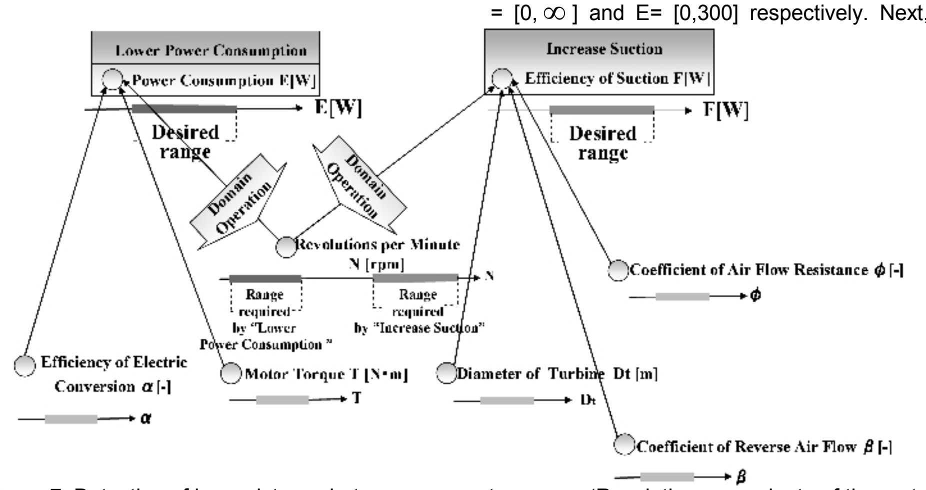


Figure.7: Detection of inconsistency between parameter ranges 'Revolutions per minute of the motor'.

independent parameters are described; α =[0.35,0.40] , T=[0.75,0.80] , ϕ =[0.05,0,0861] , β =[1.31×10^{-4} , 1.33×10^{-4}]. The range required by ‘Increase Suction’ is calculated as N = [1980,∞] with DomainOperation. Also, the range required by ‘Lower Power Consumption’ is calculated as N = [0, 1528] with DomainOperation. These two ranges have no value in common. Therefore, it is determined that there is a conflict between them.

The detected conflict can be solved with Technical Contradiction Table in TRIZ in this case. Two parameters, ‘Power Consumption’ and ‘Efficiency of Suction’ are applied to ‘Force’ and ‘Waste of energy’ respectively. As a result, the principle ‘Spheroidality’ is found as one of the way to solve the conflict. ‘Spheroidality’ gives designers an advice to use not-flat shape, for example, rollers or rotation. This principle gives the authors an inspiration; a cylindrical filter makes it possible to create more surface area than a flat-shaped filter. It is expected that a cylindrical filter creates more efficiency of suction without deteriorating suction power. In fact, this idea is used in some vacuum like Figure.8. This vacuum is not designed with TRIZ methodology. However, this example clearly shows that the solution got with this method is useful.

Figure.8: An application of the Spheroidality principle – a vacuum made by TWINBIRD Corporation [14] (Left: The vacuum cleaner, Right: a cylindrical filter)

7 DISCUSSION

7.1 Effectiveness of the proposed methods

Two different methods to detect and solve conflicts are proposed. Thanks to the proposals with different approaches, conflict-solving in broader scope of the design process is achieved. Thus, the approach with lexical expressions of functions enables to detect and solve the conflicts in the earlier design process, where the abstract concept of the service and product is discussed and described. On the other hand, the approach with the ranges of design parameter enables to detect and solve the conflicts in the later process, where more detailed specification of service activity and product; the concrete number of staff, or the valuables of parameters in product.

7.2 Possible improvements of the proposed method

As for the approach of conflict-solving using lexical expressions, proposed method can detect the conflicts from a realization structure, where functions are expressed by only predicate and object. However, functions can be expressed in various ways. A function with a subject, modifier, and direction can be described. Therefore, the detection method needs to be modified in order to treat such functions.

In the previous chapter, the realization of a vacuum is used as an example through verification experiment. And the effectiveness of the detection method with the ranges of design parameter is verified. However, it is difficult to apply this method to a realization structure, where highly-qualitative parameters are described; ‘skill of employee’ or ‘diversity of brand’, for example. Thus, the detection method needs to be modified in order to detect conflicts from a realization structure with highly-qualitative parameters.

As for the method to solve the conflict detected with the ranges of design parameters, Technical Contradiction Table in TRIZ is utilized. However, this tool was originally developed on the assumption that it is used for product design. Therefore, the modification of Technical Contradiction Table for service design makes the method more useful.

8 CONCLUSIONS

In this paper, a series of methods to detect and solve the conflict is proposed. The effectiveness of proposed methods has been proven through the verification experiments.

9 ACKNOWLEDGMENTS

This research was partially supported by the Ministry of Education, Science, Sports, and Culture through a Grant-in-Aid for Scientific Research (B), 18360079, 2006.

10 REFERENCES

[1] Tomiyama, T., 1997, “A Manufacturing Paradigm toward the 21st Century,” Integrated Computer Aided Engineering, 4, pp. 159-178.

[2] Shostack, G.L., 1981, “How to Design a Service,” in Donnelly, J.H. and W.R. George, eds., Marketing of Services, American Marketing Association.

[3] Shimomura, Y., and Tomiyama, T., 2002, “Service Modeling for Service Engineering,” Proceedings of the 5th International Conference on Design of Information Infrastructure Systems for Manufacturing 2002 -DIISM2002-, pp. 309-316.

[4] Shimomura, Y., Sakao, T., Sundin, and E., Lindahl, M., 2007, “A Design Process Model and a Computer Tool For Service Design”, Proceedings of the ASME 2007 International Design Engineering Technical Conferences & Computers and Information in Engineering Conference IDETC/CIE 2007, CD-ROM

[5] Mann, D., 2002, “Hands On: Systematic Innovation,” Creax

[6] Zhang, J., Tan, K., Chai, K., 2003, “Systematic innovation in service design through TRIZ,” Proceedings of the EurOMA-POMS 2003 Annual Conference, Vol.1, pp. 1013-1022.

[7] Zlotin, B., Zusman, A., 2001, “Directed Evolution: Philosophy, Theory and Practice,” Ideation International Inc, Southfield, MI.

[8] Ruchti, Bruno., Livotov, P., 2001, “TRIZ-based Innovation Principles and a Process for Problem Solving in Bussiness and Management,” Proceedings of the European TRIZ Association

[9] Shimomura, Y., Watanabe, K., Arai, T., Sakao, T., and Tomiyama, T., 2003, “A Proposal for Service Modeling,” Proceedings of the 3rd International Symposium on Environmentally Conscious Design and Inverse Manufacturing (EcoDesign 2003), pp. 75-80.

[10] Beckett, D., 2004, “RDF/XML Syntax Specification (Revised),” (http://www.w3.org/TR/rdf-syntax-grammar/).

[11] Patel-Schneider, P., Hayes, P., Horrocks, I., “OWL Web Ontology Language Semantics and Abstract Syntax,” (http://www.w3.org/TR/owl-semantics).

[12] Finch, W. W., Ward, A. C., 1995, “Generalized set-propagation operations over relations of more than three variables,” Artificial Intelligence for Engineering Design, Analysis and Manufacturing 1995, pp. 231-242.

[13] Umemori, Y., Kondoh, S., Umeda, Y., Shimomura, Y., Yoshioka, M., 2001, “Design For Upgradable Products Considering Future Uncertainty,” In Proceedings of Eco Design 2001, IEEE, pp. 87-92

[14] YC-5015BL, TWINBIRD CORPORATION (http://www.twinbird.jp/)

Learning networks: a method for Integrated Product and Service Engineering – experience from the IPSE project

Mattias Lindahl[1], Gunilla Ölundh Sandström[2], Erik Sundin[1], Anna Öhrwall Rönnbäck[1] & Johan Östlin[1]

[1] Department of Management and Engineering, Linköping University, Sweden

[2] Department of Machine Design, School of Industrial Technology and Management, Royal Institute of Technology, Sweden

Abstract

The aim with the Integrated Product and Service Engineering (IPSE) project is to develop a methodology for companies that want to make the journey of moving from selling products to also sell Integrated Product and Service Offerings. In order to achieve that major changes are needed in the companies. In this paper the learning network approach is described as well as the content of the workshop series that the companies participated in. The findings show that a learning network approach is beneficial methodology for achieving changes in the companies, since the participants learn from each other and from the researchers.

Keywords:

Product Service Systems (PSS), Integrated Product Service Offerings (IPSO)

1 INTRODUCTION

In recent years there has been a shift from a product focus into a growing focus on the early stages of product development and also service and business development [1, 2]. There is an understanding that in order to reach more far fetching benefits, business concepts needs to be affected and how companies develop customer offerings becomes central [3, 4]. This means that companies need to address how they develop their business models, their customer offerings and the products and services that fulfil customer offerings. To create value for the customer is more in focus [5]. To offer product and service offerings is not a new concept per se though it is a new concept for several manufacturing companies. Research show that the transition from being focused on selling products to becoming a more service oriented company is a process filled with possibilities and difficulties such as; organisational, financial and handling a new relationship to the customer [6].

The transition to product-service offerings places new and more demanding requirements on product and service development and production, along with new requirements for companies in the way they relate to and build up relationships with customers. Previous authors' research shows that existing product-service offerings are developed by the companies' marketing departments and based on existing products optimized for traditional sale [6, 7]. With product-service offerings, the skill to combine different types of products and services into a desired function becomes more crucial. In order to be able to deliver, companies need to continually develop their value chains and the competence of their personnel [3, 8]. It is important to organize the company and develop its logistics to be able to deliver a solution and create opportunities for take back of the products used in the offerings.

At the same time, modern product development involves increasingly more teamwork as well as incorporating more and more people. This development is caused, for example, by an increased technical complexity in products and an increased time pressure, as discussed in Wheelwright and Clark [9]. Wheelwright and Clark [9] and Cooper *et al.* [10] have emphasized an increased need and importance for multifunctional teams and similar types of interdisciplinary collaboration in order to minimize missed communication, provide a broader knowledge base and increase the cross-fertilization of ideas.

2 RESEARCH ON PSS AND IPSE

The research in Integrated Product and Service Engineering builds on research performed in different research disciplines such as; environmental research, product development, remanufacturing and business models for industrialized services. The research performed in the IPSE (Integrated Product and Service Engineering) research group is partly inspired from the PSS-concept as presented by e.g. Goedkoop *et al.* [11].

Furthermore, earlier research performed by the IPSE research group was partly driven by an interest in eco-design and the area was studied with an environmental filter for remanufacturing and product and service development, this is also one part of the research at present. The business models and the business strategies used in the companies are important to understand also when studying how the product development process is affected [12].

The current research is also influenced by the work in the Product/Service Engineering, where focus is on the receiver of the service and tools are developed that support the companies in describing and developing offerings to the customer. The word Engineering within the IPSE-project is a term for stating that a methodology that supports companies in developing Integrated Product and Service Offerings is developed.

3 THE IPSE PROJECT

Much research in this area, the PSS, has been more focused on theoretical issues and on making conceptual proposals for how PSS can and should work. Our research has instead been based on a close co-operation with companies that already have or are in the process of starting to sell integrated product service offerings (IPSO). The focus, e.g. in the *Integrated Product and Service Engineering (IPSE)* research project, is on developing a methodology for efficient development and production of integrated product and service offerings. A methodology developed in close co-operation with potential users and evaluated in real cases. The 2.5 year long project is founded by the Swedish Governmental Agency for Innovation Systems (VINNOVA).

The project is based on an indentified need for further research on how to develop integrated product and service offerings. The research group has also found that much research, and therefore also models and guidelines, were adjusted for large companies, though there is large potential for SMEs to gain from adjusting their business models. Furthermore, this and previous research have also shown that for small companies, product and service development is very tightly connected to business development. This means that the methodology also needs to address those issues.

The research in this project focus both on what to develop as in understanding the business models applied and how and what implications it leads to for the companies that move into selling integrated products and services, e.g. how to adjust different processes such as the product development process.

The aim of the IPSE-project is to develop a methodology that supports SMEs to develop integrated product and service offerings, i.e. a methodology for companies that wants to differentiate their product portfolio with different kinds of integrated product and service offerings suitable to their business.

The development of the methodology is made in co-operation with some twenty Swedish, small and medium sized manufacturing companies and two banks. The majority of the companies are members of two different networks and each network are related to a bank.

4 THE OBJECTIVE

In order to be able to develop an IPSE methodology that is useful for SMEs, it is essential and preferable to understand their current situation. This incorporates an understanding on how the potential users of the method are working today and what needs they have.

The objective with this paper is to analyse how SMEs can be supported, and what the requirements of a methodology for developing Integrated Products and Services are.

5 RESEARCH METHOD

The IPSE project started with a state-of-the-art analysis of the participating companies. This was made in order to get a deeper understanding about their current business models and product development and to investigate potential needs for methodological support. Qualitative research interviews [13] was used as the primary data collection method. The face-to-face interviews were recorded and performed with product and service developers as well as CEO. Customers were contacted via telephone interviews.

Based on the results from the state-of-the art analysis, the selected research approach has been to develop the methodology together with the participating companies in a series of workshops (further described in section 7). The research idea has been to create the networks as learning networks, meaning that the participating companies learn from and stimulate each other. At the workshops the researchers has participated by discussing different themes with the companies and guiding them in how they can develop integrated products and service offerings. The majority of the workshops have been recorded and all written material during the workshops such as whiteboards has been documented.

In order to be able to directly implement some of our findings from the series of workshops, we have divided all participating companies into three major groups. This has enabled for us to implement improvements to the methodology and then test the improvements with another group.

6 THE IMPORTANCE OF THE USABILITY OF METHODS

The general attitude, in both industry and academia alike, is that methods are important for improving product development performance [14]. However, the number of methods are broad, and are often met with the mixed attitudes of, for example, enthusiasm, curiosity and scepticism [15]. According to Ritzén [16], the usage of methods only becomes a regular activity if they support the users with their own work. Considering the above in combination with the low level of industry utilization of methods, a developer of methods should consider why methods have such limited use in industry. One possible explanation could be that the method does not fulfil the users' – e.g. managers' – requirements[1]. If so, the application of those requirements could be useful for further development of methods [14], e.g. the an IPSE methodology.

Lindahl [14] lists several context related aspects that influence the use of methods. For example organizational arrangements, social factors, physical settings and education levels. According to him, a method must more or less attract and fulfil requirements raised by different actors in order to be "actively used". He further states that, depending on the context, different actors' requirements are more or less important. The conclusion he makes is that it is more or less impossible to discuss requirements for a method unless considering the context in which the method will be utilized. Further, actors involved must gain something, for example a more time-efficient product development, from using the method unless it is likely that the utilization will stop or perhaps never even start.

According to Lindahl's [14] findings it exists four major requirements, of which three are interlinked. A method must exhibit the following:

- *be easy to adopt and implement* – whether a method fulfils the three following requirements is of lesser importance if it is due to a problem with adoption and

[1] Requirement is context defined as "a specific description of an attribute".

implementation and becomes seen as having a low degree of usability, and therefore is not utilized by the designers in their daily work. This requirement is the key for a method to become actively used.

- facilitate the user to fulfil specified requirements on the presumptive product and at the same time
- reduce the risk that important elements in the product development phase are forgotten.

Both of these two latter requirements relate to a method's degree of appropriateness. The second and the third requirements are related to the fourth requirement, which is considered by the author to be the most important, that the use of the method:

- *must reduce the total calendar time (from start to end) to solve the task.* If the method helps the user to fulfil specified requirements, it will also most likely help them to reduce the calendar time as well as the number of working hours needed to accomplish the product development. This is also something that enables the user to introduce changes in early phases of the development project when changes still are easy to make. Likewise, if the method reduces the risk that important moments in the product development are forgotten, it will most likely have a positive effect and reduce the calendar time and number of working hours needed.

Much would be gained if these four requirements were used as a first overall validation of the usefulness of methods.

The knowledge described above on how methods needed to be extended within the IPSE project, since it was a methodology that affects several actors in the companies and is also a methodology that supports a change process in the company when it comes to changing, business models and the product and service development, as well as the design of the products used.

7 LEARNING NETWORKS AS PART OF THE IPSE METHODOLOGY

7.1 Findings from the state-of-the art study of participating companies

From the state-of-the –art study it became clear that only a few companies had offers that were in line with a PSS offer. However, it was also clear that they had a potential to develop PSS offers and some of the companies were already performing activities in line with what is described in the PSS literature.

One important aspect that was noticed was that several of the interviewees lacked understanding the IPSE concept and the business logic. They also lacked insight into how they could gain from it, e.g. how they could do better business using a different business model and having a new perspective on their knowledge and products and how they could create more value for their customers.

The methodology was initially intended to primarily support the development process for the development of different offerings, and the products and services included in the offerings. The initial idea therefore mainly focused on the engineering aspects of the development of integrated product service offerings. During the analyses of the state-of-the-art it, however, became clear that the companies involved in the project, mainly SME´s, first needed a learning cycle for understanding the change of the business and understanding a new business logic.

The participating companies have their core competence in developing and marketing physical products. A shift into developing and delivering offerings is a major change of their business models and mindset. The considerations about this shift are often neglected when methodologies for PSS for manufacturing companies are developed.

In the participating companies few structured processes, such as stage-gate processes, were used. In small companies structured and formal processes are less used than in large companies that have a more complex organisation and more people to organise. In a small company the communication and organisation is easier. An example is that from the participating companies, the managing director is the one participating in the network and the workshops. This lack of experience of using methods also needs to be considered when developing the methodology.

7.2 A Learning Network approach

From previous research projects insights has been gained of the benefits with Learning Networks. Ritzén *et al.* [17] have described how a Learning Network approach can be used for achieving organisational change and for modelling development work. Their model is partly based on the model of "the cycle of experimental learning" by Kolb [18]. That model promotes that practitioners need to disconnect form an action experiencing loop and need support to learn from their experiences by discussing their practice in a more conceptual way and then act according to the new insights, figure 1.

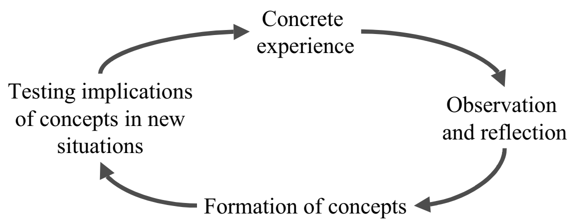


Figure 1. The cycle of experiential learning (from Kolb [18])

The basis of the learning network is that the companies bring in their own experiences and share their knowledge with each other and the researchers present their reflections and can give input from their knowledge. The researchers and the practitioners then both benefit from the network.

The goals for companies to participate in the Learning Networks are according to Ritzén *et al.* [17] to:

- gain theoretical insights and new knowledge
- share experiences
- have time and space for reflection
- gain motivation to overcome organizational barriers for change
- get support in process management

The workshop series described below are developed and organised as a learning network, see figure 2.

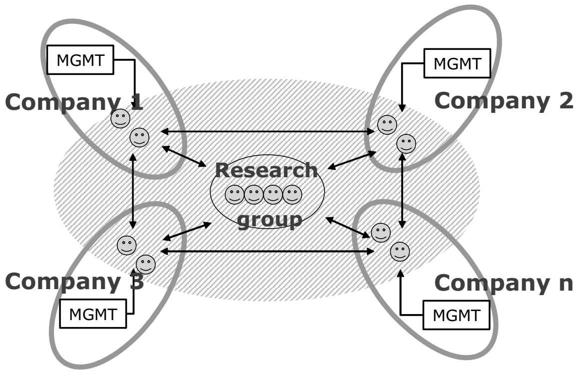


Figure 2. An illustration of a Learning Network [17].

7.3 The series of workshops

A workshop series was developed with focus both on understanding a new business model *and* how that affects product development, including the design of the products. At the same time the series of workshops it self forms a method. The first step of the workshop series was to inform and educate the participating companies of the new business logic and what the benefits for each of the companies could be. That understanding was crucial for the companies to understand and after that considerations for the physical products in the offering could be addressed. This is in line with Lindahl's [14] comment that the use of a method must be more or less useful for all users, i.e. the general manager need to be convinced that it is worth to let his company use this methodology. In SME companies, that in general have smaller economical resources, this is even more important.

When an understanding for the business models were developed the following workshops could focus more on specific aspects and on the design of the products used in the offerings from a lifecycle perspective.

7.4 Themes and a workshop series

A series of 5 workshops was developed. The development has been an interactive process and the content of the five steps have been developed and have been partly changed during the development in order to incorporate the input and conclusions from the workshops with the companies.

The five workshops topics are:

1. **Workshop 1: State-of-the-art – The current business model –** A state-of-the-art analysis is accomplished. Theoretical models as e.g. Oliva & Kallenberg [19] and .Brady *et al.* [20] are used to describe the companies current position regarding combined hardware-, software- and service offerings in order to be a base for the further business development. The environmental considerations taken today by the companies are also analysed from a lifecycle perspective.
2. **Workshop 2: Sale and customer requirements –** In this workshop customer' requirements are analysed and how those are satisfied, both in order to meet existing but also potential customers needs. This can be expressed in forms of value for the customer or user, and with advantage in the form of value chains. The aim is to identify current and potential customer values for existing as well as potential customers. The workshop ends with an analysis of the company's needs for development. Keywords in this part are SWOT, service mapping, triggers, business logic, customer communication and value arguments).
3. **Workshop 3: Planning a project for a new business logic and new a new offering –** This workshop is about planning and illustrate a new business logic and offering in terms of (1) customer requirements, and (2) project method. Customer requirements could be about how customer- and market requirements are collected. Project method is about how the company is planning its development project and if it exists any project methodology or IT supports that is suitable for Integrated Product Service Engineering.
4. **Workshop 4: Business offer development –** The focus is on organization and management of development projects and product development methods. (A special focus is on how to efficient and effectively plan the whole product's / offers' lifecycle in early phases of the development phase). This incorporate issues like trademarks and signal values coupled to the new type of offering, i.e. the integrated product service offering, e.g. by industry design and design of marketing communication material. Marketing communication comprise that *all staff* represents the company, and even partners and customers.

   In this step is also an analysis made of how the design of the product used can be improved in order to be adjusted to fit into a product and service offering. Aspects like easy to maintenance, take-back of products and environmental issues are addressed.
5. **Workshop 5: Follow-up and refine –** In this, the focus is on how to reach out with an IPSO business offer. E.g. how to in the right way influence channels and decision-takers within the target organisation, e.g. by communicating life cycle cost advantages so that the customer can understand (and can calculate and evaluate) the advantages. Furthermore, this workshop also covers how to make follow-up of calculations in order to support the providing company to in the best way use its resources.

8 ANALYSIS OF THE IPSE METHODOLOGY FORMAT

After the workshop 3 in the 1st company network a discussion took place regarding the companies' requirements on a methodology that would actually support them in their journey of changing their business model and also offering integrated products and services. The findings from that discussion, our continuous dialog with the companies and our experiences from the workshop series are described here.

One important aspect found was that the companies really appreciated the learning network approach. On of the aim with the IPSE project was from the beginning to develop an IT-tool support for the companies to use by themselves. However it has turned out to be a good strategy to put those plans aside and instead support the companies in a methodology that is of a learning network nature. The issue is complex and it is important to have a more initiated person that, from an outside perspective can, guide the company in a consultancy kind of role. This is especially important in the beginning when the concept is new for the participants.

One of the major benefits of a learning network is, according to the participants, that they get time for reflection and can learn from each other. Another crucial factor for the success of the methodology as a learning network was that the participants by themselves have recognised that *the process is the important thing*. The participants and we as researchers have experienced how their mindsets have changed since the project started and they have started to think in new ways. It has been clearly stated by the participating companies that they have experienced a transformation journey and to accomplish that a learning network is needed. That could not have been accomplished by giving each company a method to use separately. Through the workshop series they also have time to reflect between the meetings and the change process is allowed to take time.

Apart from the findings of the benefits of a learning network the findings of the methodology can be grouped into three categories; (1) the formation and organisation of the learning network, (2) the content and the themes of the workshops, (3) the methodology and documentation that the participants can keep and take with them to be used by the participants in their companies.

8.1 The formation and organisation of the learning networks

It is important to make a formation of companies in the network that really have the potential to learn from each other and also to have suitable actors from the companies. In the IPSE – networks are the managing directors participating and they are therefore on the same level in the companies and have authorities to perform changes in their companies. To have actors that have authorities to perform changes are important, but it was also found that the participants would have liked to be able to be more than one participator from each company. Ritzén *at al. [17]* also puts that forward. The companies have sometimes managed that but it has been difficult to achieve. In one of the networks the companies are part of the same group of companies and they then develop both their own company as well as the group of companies. To have more than one participator from each company also increases the possibility that someone from the company always will be able to participate in the network meetings.

It was also found that the participating companies could have had customers also participating.

The meetings should be held at “neutral ground”, meaning that it should preferably be held so that the participators have to leave their normal working environment. That makes it easier for them to focus and from being less disturbed. Several of the meetings have in the IPSE project been held at “natural ground”.

To have a moderator for the workshop series were seen as crucial. The moderator(s) are the ones organising the meetings and provide theoretical and methodological knowledge and support the companies.

8.2 The content and the themes of the workshops

When designing the workshops and themes for a learning network it is, according to the participants, important to remember that to change the business is something that takes time and shall take time. The pace and intensity of the gatherings needs to be adjusted to the changing pace of the companies.

One way is to have to put emphasis on understanding and changing the mindset of the participants in the beginning and then later on, when the participants have learned, focus on the changes that the participants can make in their company.

The themes and meetings must always connect to the previous workshop so that the deliverables and results from the previous workshop are useful and further elaborated in the next meeting. Then the participant can see the progress they are making.

The methodology need to support the understanding of the whole overall picture, i.e. how this new type of offerings relates and differs from traditional, mainly, product based offerings. The aspects that needs to be changed in the companies and the aspects that the companies should focus on needs to be pin pointed in the methodology. Such aspects can for example be; *business model, customer relationship, production aspects, environmental aspects, customer value, customer requirements and product development*.

8.3 The documentation and methodology to be used in the companies by participants

The methodology should support the work of breaking down complex issues into smaller less complex issues in order to make them easier to understand and solve.

To participate in the workshop series is found to be the most important factor but for some of the participants it would also have been useful to have guidelines formulated that they can use on their own. The participants ask for simple tools such as guidelines, checklists or questions that they need to address in the company. The method must support the work in identifying the most important aspects that need to be considered in their work.

The four major requirements stated by Lindahl [14] Is in line with this and has been a base for the further development of the IPSE methodology.

9 CONCLUSION

To develop a methodology that supports SMEs in making a shift into selling integrated products and services turned out to be something else than the researchers had initially expected. In the project it has been learned that to move into integrated products and services is a long journey for the companies and a major shift of mindset and business models. To change mindset for persons and a company is learning cycle and takes time. The methodology needs to be adjusted for that rather than primarily supporting designers and service developers in how to develop the products and services. So, before focusing on engineering aspects of integrated product service offerings, the management issues must be solved and clear, e.g. the management must understand the use and business logic of this type of offers and all people involved in the development of integrated product service offerings must understand the business logic and language.

To let the companies participate in learning network organized as a workshop series with themes has been found to be a suitable methodology. The companies bring their experiences and share knowledge with each other. The researcher’s role is to educate the participants in different subjects and to moderate the workshops. One important aspect of the workshops is that they really have been workshops and the participants have during the meetings had time to reflect how the input from the researchers relates to their company and how it could be applicable. The methodology should be supported with checklists and guidelines that the participants can use on their own in their company.

The methodology still needs improvements such as securing that several participants from each company are participating and to finalize the guidelines and checklists for the companies to be used between and after the workshops in their own company.

10 ACKNOWLEDGMENT

The authors of this paper would like to thank all participating companies for their support in the development of the method. This research was partially supported by the Swedish Governmental Agency for Innovation Systems (VINNOVA).

11 REFERENCES

[1] Ministry of Economy Trade and Economy (METI), Outline of "Policies of Support for Environmental Management and Environmental Business". 2003, Ministry of Economy, Trade and Economy (METI): Tokyo, Japan. p. 8.

[2] Charter, M. and U. Tischner, eds. Sustainable Solutions - Developing products and services for the future. 2001, Greenleaf Publishing Limited: Sheffield, UK.

[3] Söderström, J., Från produkt till tjänst - Utveckling av affärs- och miljöstrategier i produktorienterade företag, in Handelshöskolan. 2004, Stockholm University: Stockholm.

[4] Tomiyama, T. Service Engineering to Intensify Service Contents in Product Life Cycles. in 2nd International Symposium on Environmentally Conscious Design and Inverse Manufacturing (EcoDesign 2001). 2001: IEEE Computer Society.

[5] Shimomura, Y., K. Watanabe, T. Arai, T. Sakao, and T. Tomiyama. A Proposal for Service Modelling. in EcoDesign 2003. 2003. Tokyo, Japan.

[6] Ölundh, G., Environmental and Developmental Perspectives of Functional Sales, in Division of Integrated Product Development, Department of Machine Design. 2003, Royal Institute of Technology: Stockholm, Sweden. p. 108.

[7] Sundin, E., M. Lindahl, T. Sakao, Y. Shimomura, and M. Björkman. New Engineering Design for Functional Sales Business. in International Conference on Engineering Design, ICED 05. 2005. Melbourne, Australia.

[8] Brännström, O. and B.-O. Elfström. Integrated Product & Service Offerings - Their Rationale and Creation. in International Product Development Conference. 2002. Sophia-Antipolis (near Nice), France: European institute for Advanced Studies in Management.

[9] Wheelwright, S.C. and K.B. Clark, Revolutionizing Product development, Quantum Leeps in Speed, Efficiency and Quality. Vol. 35. 1992, New York, US: The Free Press. 29.

[10] Cooper, R.G., S.J. Edgett, and E.J. Kleinschmidt, Portfolio Management for New Products. 1998, Reading, Massachusetts: Perseus Books.

[11] Goedkoop, M.J., C.J.G. Van Halen, H.R.M. te Riele, and P.J.M. Rommens, Product Service Systems, Ecological and Economic Basics. 1999, VROM: Hague, the Netherlands.

[12] Lagerholm, B. and A. Öhrwall Rönnbäck, IKT för funktionsförsäljning i nya värdeskapande nätverk: Nya möjligheter med informations- och kommunikations-teknik (IKT). En studie av affärsstrategier för ökad lön-samhet och livskraft i svensk verkstadsindustri-slutrapport. 2003: Slutrapport 2003-12-31, till VINNOVA projekt med diarienr 2002-01502 och 2002-01492.

[13] Kvale, S., The qualitative research interview - a phenomenological and a hermeneutical mode of understanding. Journal of Phenomenological Psychology, 1983. 14: p. 171-196.

[14] Lindahl, M., Engineering Designers' Requirements on Design for Environment Methods and Tools, in Industrial Engineering and Management. 2005, KTH: Stockholm, Sweden.

[15] Beskow, C., J. Johansson, and M. Norell. Implementation of QfD: Identifying Success Factors. in Proceedings of the International Conference on Engineering and Technology Management, IEMC'98. 1998. San Juan, Puerto Rico, USA.

[16] Ritzén, S., Integrating environmental aspects into product development - Proactive measures, in Integrated Product Development Division, Dept. of Machine Design. 2000, Royal Institute of Technology: Stockholm, Sweden.

[17] Ritzén, S., A. Zika-Viktorsson, J. Forslin, and M. Engwall, Learning to change – a network approach to engineering management development, in In proceedings of the International Conference on Engineering Design, ICED 05 2005: Melbourne, Australia.

[18] Kolb, D.A., Experiental Learning – Experience as the Source of learning and Development. 1984, Englewood Cliffs, NJ, USA.: Prentice Hall PTR.

[19] Olivia, R. and R. Kallenberg, Managing the transition from products to services. International Journal of Service Industry Management, 2003. Vol. 14(No. 2): p. pp. 160-172.

[20] Brady, A., A. Davies, and D.M. Gann, Creating value by delivering integrated solutions. International Journal of Project Management, 2005. Vol. 23(No.5): p. pp. 360–365.

Proposal of Idea Generation Support Methodology for Eco-Business

Yasushi Umeda[1], Takuya Nishioka[1], Shinichi Fukushige[1], Shinsuke Kondoh[2], and Shozo Takata[3]

[1] Department of Mechanical Engineering, Graduate School of Engineering, Osaka University, Japan

[2] National Institute of Advanced Industrial Science and Technology, Japan

[3] Department of Industrial & Management Systems Engineering, Faculty of Science and Engineering, Waseda University, Japan

Abstract

In order to solve the environmental issues, transition from the conventional business to the environmentally conscious business (eco-business) is eagerly required. Although many projects of the eco-business have been planned, it is difficult to satisfy environmental consciousness and success in business at the same time. This paper proposes a methodology for supporting idea generation of eco-business. First, we collect 130 cases of the eco-business in Japan, investigate and classify them to develop a guideline for their success. Second, based on this guideline, we propose a prototype system, named Eco-business Generator, for supporting idea generation of eco-business.

Keywords:

Eco-business; Product-Service System; Idea Generation

1 INTRODUCTION

Toward the sustainable society, environmentally conscious design and manufacturing have been eagerly studied. However, environmentally conscious products are not always successful in business. Rather, achieving environmental consciousness and success in business at the same time is very difficult. Solving this problem will promote the shift from the conventional mass production to the sustainable production. To do so, business, in addition to product life cycle, should be appropriately designed as a part of life cycle design. In this perspective, Hauschild *et al*. [1] clarified the concept of life cycle design, Masui *et al*. [2] and Kobayashi [3] proposed design methodologies for supporting life cycle design, Janz *et al*. [4] proposed an optimization method of life cycle cost, and Sutherland *et al.* [5] analyzed recycling markets from the economic viewpoint. However, the issue of business design in the context of life cycle design has not been well studied yet.

The objective of this study is to propose a design methodology for environmentally conscious business (or eco-business). Especially, this paper proposes an idea generation support methodology of eco-business, which encourages a designer to generate various business ideas that shift a conventional business to an eco-business. Here, we define eco-business as business that reduces directly or indirectly environmental loads of the society and, at the same time, makes profit. The approach taken here consists of three steps; namely, investigating about 130 cases of eco-business in Japan, extracting rules for planning eco-business, and developing a system 'Eco-business Generator' based on the extracted rules. In other words, we intend to develop a methodology like TRIZ [6] in the domain of eco-business design.

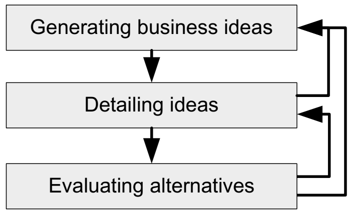


Figure 1: Business design process.

Figure 1 depicts a simplified design process of eco-business we assume. First, various business ideas are generated and some hopeful candidates are detailed by, for example, describing life cycle scenario, and then evaluated from the environmental and economic viewpoints. This process can be viewed as a life cycle design process focusing on business perspective. In general, a method for finding out creative ideas consists of divergence process, *i.e.*, generating ideas, and convergence process, *i.e.,* evaluating and selecting ideas [7]. For example, 'brainstorming' is a method for supporting the divergence process. The idea generation in this paper also focuses on the support of the divergence process in the context of eco-business. In this sense, the method aims at encouraging a designer to generate as many ideas as possible, which may include creative or innovative ideas, before evaluating feasibility, environmental effect, or profitability of each idea. While the convergence process, which corresponds to the second and third steps in Figure 1, is out of the focus of this paper, we are proposing life cycle scenario description support system [8] and life cycle simulation system [9] as tools for supporting the second and third steps, respectively. Note that, while the discussion in this paper has a lot of commonality with the conventional design of business, this paper focuses on the design of *eco*-business and, therefore, does not discuss, for example, how to increase added value in business without considering the environmental impact. And basically, this paper tries to find out a method to turn a traditional environmentally unconscious business into eco-business with increasing or keeping its profitability.

Table 1: Summary of practical eco-business.

Category	No. of cases	Examples
Reuse	32	car parts, digital camera
Recycling	25	PC, mobile phone
Maintenance	16	PFI, ESCO
Rental/Lease	13	air conditioner, carpet
Authentication	11	green purchasing, eco-labeling
Information mediation	10	reusable car parts, transportation capacity
Information sharing	7	SCM, RFID
Evaluation	3	eco-efficiency, LCA
Others	16	IPP, green brand

2 STRUCTURE OF ECO-BUSINESS

As defined in Section 1, eco-business makes profit by providing eco-products or eco-services for the customers. Eco-business includes both business-to-business and business-to-consumer business. In order to identify success factors for eco-business, we collected about 130 cases of the eco-business in Japan [10]. Table 1 summarizes the result of the investigation. Note that the number of cases in each category does not reflect the number of eco-business in the real world. Since there are too many cases of material recycling, we mainly focus on reuse, servicizing, and rental and lease systems. One of our findings of this survey is that 'information mediation' (*e.g.,* mediating wastes, secondhand goods, and transportation capacity) is unexpectedly a hopeful type, since it does not need large initial investment but has remarkable environmental effects.

First, we extracted common basic structure of the eco-business from the cases (see Figure 2). As shown in this figure, an eco-business provides products or services for customers and receives money from the customers. Here, customer's value can be classified into four kinds:

- Cost reduction; *e.g.*, cheaper products and waste treatment with cheaper cost
- Improvement of service quality; *e.g.*, products with higher environmental performance
- Avoidance of risks; *e.g.*, proper end-of-life treatment service
- Improvement of image (mainly of companies and their products); *e.g.*, EMS (environment management system) support service

In terms of customer's value, there is no difference between the eco-business and the conventional business. The most distinctive feature of the eco-business is the prerequisite that such business should directly or indirectly reduce the environmental loads from the viewpoint of the society. This is the most difficult point in planning a new eco-business. In other words, relating reduction of environmental loads in the society to customer's value with economical feasibility is the critical path in finding out new eco-business ideas and this paper aims at supporting this point. In Figure 2, this point means to find out economically feasible thick arrows relating reduction of environmental loads of the society to customer's value.

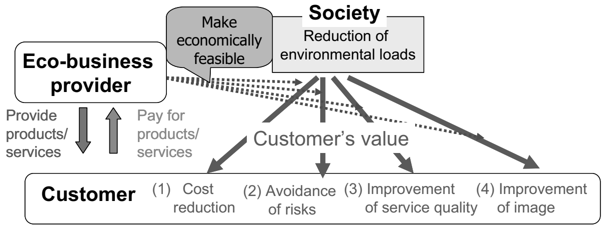


Figure 2: Basic structure of eco-business.

In order to extract general rules for finding out these thick arrows, we find out key stakeholders and key success factors from the collected cases.

2.1 Key stakeholders

We found out the following five key stakeholders in the eco-business.

Eco-business provider (EBP): Eco-business provider is the main stakeholder who operates the eco-business. In many cases, the eco-business provider offers their products to its customers, but, in some cases, other stakeholders than eco-business provides provide products (*e.g.*, consultancy firms, mediators etc.). In the former case, the eco-business provider is identical to the product provider but, in the latter case, the eco-business provider is different from the product provider.

Product provider: Product provider offers products or services to its customers. As discussed above, the product provider and the eco-business provider are not always identical. For example, Energy Service Company (ESCO), which provides solutions to save electricity for the customer (industrial units, hospitals etc.), is not the same company as the provider of electricity.

Customer: Customer is the customer of the eco-business. Customers include both of consumers and firms. A customer consumes and pays for provided products/services to satisfy its needs or to produce its own products/services for its own customer (*viz.*, customer's customer). At the same time, the customer discharges wastes and emissions to the society.

Customer's customer: When the customer of the eco-business is a company that offers its products/services to its own customers, we should take this customer's customer into account as a stakeholder of the eco-business.

Society: In this model, all environmental loads go to the society in the end. All stakeholders other than the society are responsible for these environmental loads resulting from their activities. Therefore, we consider the society as an important stakeholder of the eco-business.

2.2 Key success factors for eco-business

In order to make their eco-business feasible, the eco-business provider should provide customer's value and reduce operational costs of these activities. The analysis of the cases of eco-business revealed various key success factors. We summarized them as eight rules for providing value and eight rules for reducing costs of eco-business providers.

Rules for providing value

1) Use one more time

Reusable goods and energy are sometimes thrown away because of, for instance, too small amount of them is collected. If they are used one more time by largely colleting them, an eco-business provider can make profit for reusing these goods and energy. This may, of course, reduce waste generation and resource consumption. Examples of this rule are reuse of components and secondhand market of products.

2) Servicizing

Servicizing [11][12] refers to selling services or functionality rather than products. While the product is still owned by an eco-business provider, customers pay for use or maintenance. E-learning and videoconference substituting for transportation are examples of servicizing. Especially, servicizing gives the provider a chance of the life cycle management, such as management of used products.

3) Timesharing

Capacity of products such as personal automobiles and industrial equipment is sometimes under-utilized. By encouraging customers to abandon individual ownership, more intensive utilization of products can be realized. This may reduce resource consumption and waste generation by decreasing production and disposal of under-utilized products. At the same time, this can reduce customer's procurement costs and costs for disposal of products. Leasing and rental schemes are examples of this rule.

4) Management of hidden loss

Identification and proper management of hidden environmental loss of customer's activities often reduces customer's costs and environmental risks. An example is ESCO business, where ESCO comprehensively manages use of electricity of its customer to save electricity.

5) Finding out new use of wastes

Wastes and waste energy can be sometimes transformed into valuable resources. In this case, an eco-business provider makes profit for selling them. The zero emission concept [13], which aims at reutilizing wastes from a site as resources for another site by organizing industrial clusters, is an example of this rule. This rule may reduce the amount of wastes and, indirectly, the amount of resource consumption.

6) Application of cleaner methods to satisfy customer needs

Providing more environmentally conscious products or services to customers directly improves environmental performance of products/services and also increases the customer's image from the society. Examples of this rule include introduction of photovoltaic power generation. Resource consumption, waste generation, and environmental emissions may be reduced by increasing the environmental performance by this rule.

7) Compliance with the environmental legislation and acquisition of green authentication

Prompt compliance with environmental legislation and acquisition of green authentication (*e.g.*, ISO14000 series and eco-labeling) improve customer's image. Supporting and consulting these kinds of customer's activities are an established area of the eco-business. This rule motivates a company to increase her environmental performance.

8) Undertaking environmental loads

In our examples, some eco-business providers (*e.g.*, waste disposers and recyclers) contract for proper treatment of environmental loads (*e.g.*, end of life products) from their customers. This rule encourages proper treatment of wastes that reduces environmental emissions.

2.3 Rules for reducing costs

Many innovative projects of the eco-business fail because of lack of profitability. Therefore, it is also important to reduce the provider's costs for increasing its profitability. For this purpose, we derived the following eight rules. Some of them are almost the same as the value provision rules in Section 2.2. This means that some rules are effective for both of providing value and reducing costs.

a) Management of life cycles

Proper management and control of product life cycles (especially, after they are sold) can reduce both of environmental loads and costs. Closed-loop manufacturing of one-time use camera is an example of this rule.

b) Expansion of the business scale

This is a typical rule in the mass production, but still effective in the eco-business. Monopoly of the secondhand market of personal computers is an example of this rule. Especially, this rule is important in the secondhand type of the eco-business, because larger amount results in wider variety of products in this business. Collection of an enough amount of recyclable materials, which increases profitability, is another example of this rule.

c) Reutilization of wastes

When an eco-business provider undertakes waste treatment or product disposal from its customer, reutilizing wastes can reduce the provider's treatment costs.

d) Utilization of knowledge and information

As in the conventional business, utilization of knowledge and information can increase efficiency of energy and material usage, labor, and facilities. In addition to operational cost reduction of the provider, this rule confirms consultancy business that makes profit by helping customers' cost reduction.

e) Linkage and cooperation among various industries

Related to the rule 'Expansion of the business scale,' cooperation among various industrial sectors sometimes contributes to reduction of the operational costs. This pattern is often found when a firm enters into a new eco-business domain.

f) Combining various values

Providing multiple products/services bundled into a package can reduce the operational costs. Product service system, which is a marketable set of products and services jointly fulfilling customer's needs, is an example of this rule.

g) Technological innovation

Technological innovation such as development of more environmentally conscious devices can reduce both of the costs and environmental loads. Of course, this is a dominant rule in the conventional business.

h) Outsourcing

This is also a typical cost reduction rule in the mass production, but still valid in the eco-business. Because many cases of the eco-business cover multiple life cycle stages, the provider cannot execute the entire task alone. Therefore, making right outsourcing decision is quite important for cost reduction. While the rule 'Linkage and cooperation among various industries' assumes a flat alliance of stakeholders, this rule assumes the vertical relation between a client and an undertaker. Undertakers of the outsourcing can reduce their costs by applying the rule 'expansion of the business scale.'

As a result, we can summarize the rules in Sections 2.2 and 2.3 into three principles; namely, life cycle thinking, making lean throughout a product life cycle, and adding environmental value. Although these principles are obvious, the advantage of this methodology is to indicate how to realize these principles to a designer in the form of these rules and the database of the collected cases, as shown in the next section.

3 IEDA GENERATION SUPPORT SYSTEM

3.1 Basic approach

We developed a prototype system named 'Eco-business Generator.' for supporting a designer to generate eco-business ideas, based on the collected cases and the extracted rules. For supporting the idea generation step of eco-business design in Figure 1, we here take an approach to encourage a designer in generating as many ideas as possible by applying the rules to an existing business so as to change it into an eco-business.

For this purpose, we identified the functions of the system and the approaches for implementing these functions as follows:

1. *Providing a workspace for idea generation*
   We represent a business in the form of a life cycle flow model proposed in Section 3.3. The system provides a designer to describe and modify business ideas flexibly as the life cycle flow model.
2. *Clarifying stakeholders, including the society, and relations among them*
   We explicitly represents stakeholders and their relations in the above model. As discussed in Section 2.1, since many stakeholders are involved in an eco-business and they consist of a supply chain (or life cycle chain), we consider that flow model is appropriate for representing an eco-business idea. And especially, explicit representation of 'Society' is important in considering environmental consciousness.
3. *Encouraging a designer to apply the extracted rules for generating ideas*
   For this purpose, we represent the rules in Section 2.2 and 2.3 as the operations modifying the life cycle flow model. As a result, a designer can generate diverse variants of a business idea by applying different rules.
4. *Provide a case base of the eco-business*
   We prepare a database of eco-business cases (eco-business case base) including those described in Section 2. Since the cases are indexed by the rules, a designer can understand the meaning of the rules by browsing the case base and find out new ideas from the cases by analogy.
5. *Choosing hopeful ideas from a lot of generated ideas*
   This methodology assumes that a designer generates a lot of business ideas. In order to support him/her to choose hopeful ideas from them, we develop 'idea map,' which represents all ideas, and employ 'Pugh's method' [14] for comparing several ideas.
6. Clarifying a procedure to generate eco-business ideas
   In order to guide a designer, we propose the procedure of idea generation as shown in Section 3.4.

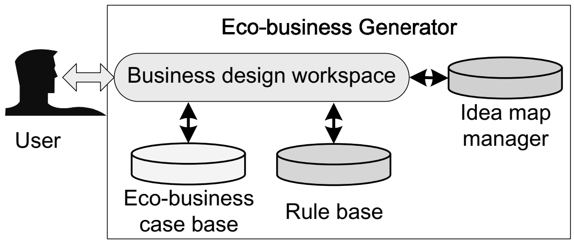


Figure 3: Architecture of eco-business generator.

3.2 System architecture

Figure 3 depicts the system architecture of Eco-business Generator. It consists of Business design workspace, Idea map manager, Eco-business case base, and Rule base. Business design workspace provides a workspace for describing a business idea as a life cycle flow model, for modifying the model, and for comparing several ideas. Eco-business case base stores practical cases of eco-business and each case consists of its name, description, keywords, eco-business provider, customer, increased value, reduced environmental loads, applied rules, and references.

The rule base stores the rules for providing value and for reducing costs and each rule consists of its name, conditions to be applied, and resulting modification to the model. For example, the value provision rule 1 (use one more time) has its condition 'this rule is applicable to objects disposed by a customer' and its modification 'the disposal flow is changed to a collection flow by the eco-business provider and monetary flow appears from the provider to the customer.' Although the explanation of the condition and the modification are represented in the form of text here, they are implemented as patterns of modifications of nodes and links in the system.

And Idea map manager manages ideas generated by the user in the form of tree (see Figure 5), in which a node is a life cycle flow model (*i.e.*, a business idea) and a link denotes that lower nodes are generated by modifying the upper nodes.

3.3 Representation of eco-business

We represent a business idea as a life cycle flow. We here employ IDEF [15] as a basic representational scheme. Figure 4 shows an example of this representation denoting a rental system of mop, a cleaning tool, which is popular in Japan. In this figure, rectangles and ovals denote processes and objects flowing through processes, respectively. In this example, 'Rental' and 'Mop' are a process and an object, respectively. Rows are stakeholders (*e.g.*, Producer, EBP (Eco-Business Provider) in this example). A stakeholder 'Society,' which receives all kinds of wastes and emissions, should be included in the model. Each flow that crosses a border between stakeholders may be accompanied with monetary transfer, represented as arrows in this figure. The monetary transfer may be in the same direction with the object flow (*e.g.*, a recycler collects wastes with some fee paid by a waste generator) and in the opposite direction (*e.g.*, a consumer buys a product). In this example, value provision

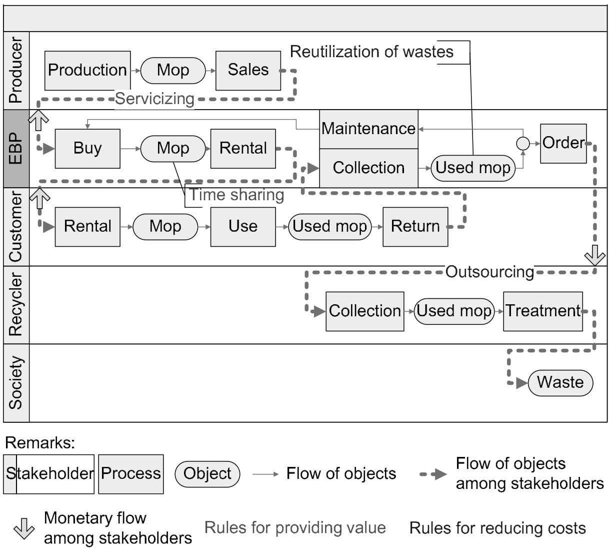


Figure 4: Example of business representation.

rules 'servicizing' and 'timesharing' (red letters) and cost reduction rules 'reutilization of wastes' and 'outsourcing' (blue letters) are applied.

3.4 Procedure of idea generation

In this system, eco-business idea generation proceeds as follows:

1. *Description of a conventional business*
   First, a user describes his/her conventional business in the form of the life cycle flow model as the start point.
2. *Objective setting*
   Next, the user determines objectives of eco-business to be designed. For supporting this step, Eco-business Generator is equipped with a tool for SWOT (Strengths, Weaknesses, Opportunities and Threats) analysis [16] and a tool for describing IFR (Ideal Final Result), which is a technical term in TRIZ representing ideal situation of the target. This means that a practical eco-business resides between the conventional business and IFR.
   The typical objectives include elimination of serious environmental loads. In other words, such environmental loads should be described in the 'Society' box, the user can easily find out the target in our representation.
3. *Application of eco-business rules*
   In this step, the user selects rules to be applied. The procedure of this step consists of finding out an applicable rule for providing value, modifying the life cycle flow model (Step 4), finding out an applicable rule for reducing costs, modifying the model again, and repeating the steps 3 and 4 until the user judges the objectives can be achieved. For supporting this step, the system indicates applicable rules and places where the rules can be applied by searching through the model. Moreover, the system shows the meaning of each rule by denoting cases that use the rule in the case base. When the user applies a rule, the model is automatically modified according to the modification pattern of the rule (see Section 3.2).
4. *Modification of the business*
   Since the modification by the rule in Step 3 is partial and not complete, the user should modify the model so as to include the selected rules. First, the user should choose an eco-business provider (EBP), which is the main stakeholder of this business. In some cases, an existing stakeholder, *e.g.*, the manufacturer, becomes the EBP, the user may add a new stakeholder as EBP in other cases.
5. *Selection and comparison*
   By repeating Steps 3 and 4, the user generates many business ideas and stores them into the idea map. In the map, leaves denote more complete ideas than other nodes and the tree structure qualitatively indicates categories of ideas. Therefore, the user first selects hopeful ideas from leaf nodes and then compares them. Since the ideas are purely qualitative, we employ 'Pugh's method' [14] that can compare concepts in the qualitative level. As a result, the user determines some hopeful eco-business ideas with which he/she proceeds to the 'detailing ideas' stage of the eco-business design (see Figure 1).

4 CASE STUDY

We generated various eco-business ideas by using Eco-business Generator. Examples include a bicycle rental business at railway stations [10], battery rental systems, recycling systems of office documents, and recycling system of flat panel display (FPD). Let us take the recycling systems of FPD as an example. Eco-business of FPD is important because, while its market share increases rapidly, recycling of FPD is very difficult although it contains a variety of scarce materials. In this case, we set the business objective to be reduction of consumption and disposal of scarce metals with satisfying consumers' needs of 'prefer larger screen.' For this objective, we generated several types of eco-business ideas as shown in Figure 5, including A) providing information (*e.g.*, disposal time and measure) to consumers, B) the manufacture collecting disposed products, C) the shop collecting disposed products, D) a new stakeholder entering the FPD life cycle business, and the rest (*viz.*, E), F) and G)) are the business in which consumers participate for circulating the products. After the selection and comparison process, we selected the most hopeful eco-business idea that combines two business ideas. On one hand, the shop, which is the EBP, lends showcase to consumers who want to sell their old FPD. On the other hand, the shop also buys old FPDs for trade-in, if the consumers want, and refurbish them for the second hand market. Both ideas may extend the product lifetime.

5 DISCUSSIONS

As described in the previous section, Eco-business Generator successfully supports us for generating various eco-business

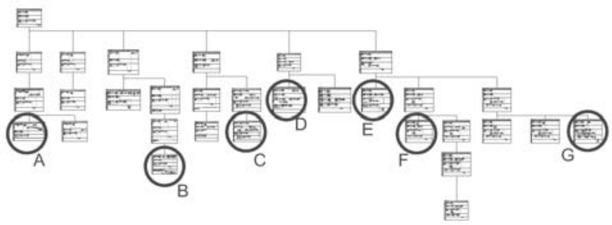


Figure 5: Idea map of the case study.

ideas quite easily. Especially, an advantage of this system is to support a user to find out many business ideas comprehensively, which may include innovative ides. Moreover, the case base helped a lot to understand the meaning of each rule. In this sense, the objective of this study, support of business idea generation in Figure 1, is successfully achieved. Furthermore, by reviewing this system, experts commented that this methodology is useful especially for training and the case base is useful for learning ideas of the eco-business and for sharing knowledge among experts in different domains. These comments imply that Eco-business Generator will work as TRIZ in the domain of eco-business design.

On the other hand, we found several issues to be solved:

- Many generated ideas are not innovative, since the proposed methodology is based on existing cases although they are abstracted as the rules. This requires the system to be more flexible for combining rules very easily. For reducing this problem, we are extending the methodology so as to encourage the user to generate ideas by employing brainstorming and KJ method.
- Success of business in the real world deeply depends on business scale and customers' selection and it is very difficult to model these factors even if we employ the life cycle simulation. It will be helpful for estimating these factors to add information about business scale, sales amount, profit, and potential market size to each case in the case base.

Note that the proposed methodology focuses on the divergence process of the eco-business design. This means that quantitative evaluation of each idea and clarification of quantitative prerequisites of each idea are out of its scope. While we are studying these issues (*e.g.*, [8][9]), it is our important future issue to support the whole design process of the eco-business by extending the proposed methodology. Nevertheless, as far as the support of the divergence process, we conclude that we succeeded in proposing a feasible methodology.

6 CONCLUSIONS

This paper analyzed 130 cases of the eco-business in Japan. The analysis of the cases identified four customer's values related to reduction of environmental loads of the society and success factors as two sets of rules; namely, eight rules for providing value and eight rules for reducing costs of the eco-business provider.

We also proposed a prototype system for supporting idea generation in planning of the eco-business by utilizing the rules and the case base of eco-business. The case study indicated that the system successfully supports us in the divergence process for generating various eco-business ideas quite easily. Future works include the support of generation of more innovative ideas, the support of the convergence process and quantitative evaluation of ideas, and application of the methodology to more practical case studies.

This study was executed by the research committee for life cycle engineering in the Japan Society for Precision Engineering. We also thank Naoto Yoshida (Osaka University) for his contributions on implementing the system.

REFERENCES

[1] Hauschild, M. Z., Jeswiet, J. Alting, L., 2004, Design for Environment? Do We Get the Focus Right?, Annals of the CIRP, 53/1: 1-4.

[2] Masui, K., Sakao, T., Aizawa, S. Inaba, A., 2002, Quality Function Deployment for Environment (QFDE) to Support Design for Environment (DFE), Proc. of DETC'02, ASME, (CD-ROM).

[3] Kobayashi, H., 2005, Strategic evolution of eco-products: a product life cycle planning methodology, Research in Engineering Design, 16/1-2:1-16.

[4] Janz, D., Sihn, W., 2005, Product Redesign Using Value-Oriented Life Cycle Costing, Annals of the CIRP, 54/1: 9-12.

[5] Sutherland, J. W., Gunter, K. L., 2002, A Model for Improving Economic Performance of a Demanufacturing System for Reduced Product End-of-Life Environmental Impact, Annals of the CIRP, 51/1:. 45-48.

[6] Terninko, J., Zusman, A., Zlotin, B., 1998, Systematic Innovation: An Introduction to TRIZ (Theory of Inventive Problem Solving), Saint Lucie Press.

[7] Schwenk, C. R., 1984, Cognitive Simplification Processes in Strategic Decision-Making, Strategic Management Journal, 5/2: 111-128.

[8] Suesada, R., Itamochi, Y., Kondoh, S., Fukushige, S., Umeda, Y., 2007, Development of Description Support System for Life Cycle Scenario, Proc. of 14th CIRP Int. Conf. on Life Cycle Engineering, pp. 29-34.

[9] Umeda, Y., Nonomura, A., and Tomiyama, T., 2000, Study on life-cycle design for the post mass production paradigm, AIEDAM, 14/2:149-161.

[10] Kondoh, S., Kurakawa, K., Kato, S., Umeda, Y., Takata, S., 2006, Analysis of Key Success Factors for Eco-Business through Case Studies in Japan, Proc. of 13th CIRP Int. Conf. on Life Cycle Engineering, pp. 275-280.

[11] Lindahl, M. and Olundh, G., 2001, The Meaning of Functional Sales, Proc. of the 8th CIRP Int. Seminar on Life Cycle Engineering, pp. 211-220.

[12] Goedkoop, M. J., et al., 1999, Product service systems, ecological and economic basis, PricewaterCoopers.

[13] Graedel, T. E., and Allenby, B. R., 1995, Industrial Ecology, Prentice Hall.

[14] Pugh, S. 1996, Concept Selection – A Method that Works, Creating Innovative Products using Total Design, Addison-Wesley, pp.167-176.

[15] Clarkson, J., and Eckert, C., 2005, Design process improvement: a review of current practices, Springer.

[16] Jenster, P. V., Hussey, D., 2001, Company Analysis: Determining Strategic Capability, John Wiley & Sons.

Multi-Agent Market Modeling Based on Analysis of Consumer Lifestyles

Tomomi Kito[1], Kousuke Fujita[2], Takeshi Takenaka[1] and Kanji Ueda[1]

[1] Research into Artifacts, Center for Engineering, The University of Tokyo, Japan

[2] Department of Precision Engineering, School of Engineering, The University of Tokyo, Japan

Abstract

This paper describes the multi-agent modeling of a service market which consists of a service provider and various types of consumers. The difficulty in determining, and thus reacting to, the needs of markets, lies in the fact that consumers have diverse value concepts, which can differ through interaction with others. This study conducts a questionnaire on consumer lifestyles, and constructs models for the several types of consumers, based on the survey results. It also clarifies characteristics of a service market. The multi-agent simulations of this service market are executed to verify the validity of the proposed model.

Keywords:

Multi-Agent; Service Market; Value Creation

1 INTRODUCTION

Recent demands on manufacturing and service industries have shifted from low-cost mass-production to needs satisfaction, through diversification of consumer preferences and lifestyles. The so called “Long tail phenomenon” [1] is one of those phenomena showing such diversification of consumer preferences. This phenomenon generates niche markets, which have been thoroughly researched in order to find new business opportunities [2]. Awareness of diversified value concepts of consumers is essential to provide products or services which fit the market.

The difficulty a producer has of understanding and matching with the needs of markets lies in the fact that consumers change their value concepts through interaction with others. One such instance of this is expressed as “network externality”. Network externality is defined as a technological externality dependant on the number of users of a certain product or a service, by which consumers’ utilities are improved [3]. To deal with such dynamics, a producer must create products or services adaptively. This characteristic is more distinctive in service markets rather than product markets.

This paper describes the multi-agent modeling of service markets. Section 2 presents the model of a multi-agent market, which comprises of one producer and various types of consumers. In this study, consumers’ value concepts are determined based on quantitative analysis of the results of a lifestyle’s questionnaire, in order to bring the model close to a real market. This section also constructs models of several types of consumer based on the survey outcome.

Section 3 shows the results of the multi-agent market simulation, in which a producer is modeled as a learning agent. By comparing with the case in which a producer doesn’t have a learning mechanism, it indicates the effectiveness of an adaptive provision of services. Section 4 discusses a future strategy for studies of services, by classifying value creation based on the concept of emergent synthesis.

2 MODELING OF MULTI-AGENT MARKETS

2.1 Outline of the model

The service market model consists of one producer and N consumers. The producer produces services and each consumer decides, based on its own value concept and demand, whether to use these. The detailed models of a service, consumers, and a producer are described below.

Service

Let each service, S_m, comprise three functions:

$S_m = S_m\,(f_1\ f_2\ f_3)$,

where f_1, f_2, f_3 = 0, 1, 2, 3. Here, f_i (i = 1, 2, 3) is the level of each function. The price of service S_m is given as follows:

$$P_m = S_m \Gamma = \begin{pmatrix} f_1 & f_2 & f_3 \end{pmatrix} \begin{pmatrix} \gamma_1 \\ \gamma_2 \\ \gamma_3 \end{pmatrix}, \tag{1}$$

where $\Gamma(\gamma_1\ \gamma_2\ \gamma_3)$ denotes the unit prices of f_1, f_2 and f_3.

Producer

The producer produces services with unit cost $T(\tau_1,\ \tau_2,\ \tau_3)$. The profit the producer gains when (s)he provides S_m is defined as

$$\Pi = \sum_m (P_m - S_m \mathrm{T}) \times N_m\ . \tag{2}$$

Here, N_m is the number of cosumers who are using S_m. The producer makes a decision as to which service to produce at every step, with the intention of creating a service whereby the total profit is maximized. The producer makes these decisions at every step based on the Q-learning algorithm.

Consumer

Each consumer, C_n (n = 1, 2,…, N), has a demand level, D_n (d_{n1} d_{n2} d_{n3}), and her or his own reservation value V_n(v_{n1} v_{n2} v_{n3}), which expresses how much (s)he can pay for one level of each function. RP_n is the reservation price, i.e. the price C_n is willing to pay for the service, given by:

$$RP_n = D_n V_n = \begin{pmatrix} d_{n1} & d_{n2} & d_{n3} \end{pmatrix} \begin{pmatrix} v_{n1} \\ v_{n2} \\ v_{n3} \end{pmatrix}. \quad (3)$$

C_n makes the decision of whether to use S_m or not. When it uses S_m, it gains the utility $U_n = RP_n - P_m$.

In addition, network externality is introduced only into the second function:

$$V_n(v_{n1} \quad v_{n2} \quad v_{n3}) = V_n(v_{n1} \quad a + b\sum_m N_m \quad v_{n3}) \,, \quad (4)$$

2.2 Consumer value concept modeling based on the lifestyle survey

In order to bring the model closer to a real market, this section determines the demand levels, *D,* and the reservation values, *V,* of consumers based on actual data.

This study conducted a questionnaire on consumer lifestyles in the central Tokyo area of Japan and collected data from approximately 7,000 people. The questionnaire looked at various topics, including life attitudes, leisure activities, personality and the use of information technologies. To highlight the structure of this survey let us consider producers who provide information and communication services. The values associated with this service may be categorised into the following three types: (i) the value linked to the functionality of the product; (ii) the value connected with the communications themselves; and (iii) any added value such as brand loyalty. Taking the example of cellular telephones, the types above can be related as follows: (i) the time a consumer spends on functions other than talking or text messaging; (ii) the time spent actually talking or send text messages; and (iii) the depth of passion for particular stylistics. Based upon the answers to these, d_{n1}, d_{n2} and d_{n3} are respectively determined in three levels (0, 1, 2). A market model composed of 787 consumers was constructed based on this survey data. The detail of consumers' demands is shown in Figure 1.

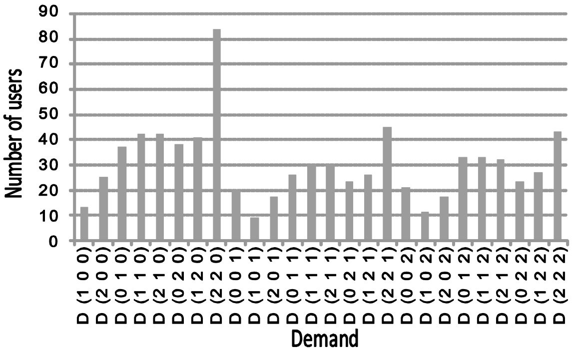


Figure 1: Consumers' demand

3 MUTI-AGENT MARKET SIMULATION

3.1 Simulation set-up

The next section demonstrates three simulations: Ex-1, 2 and 3; with the settings as shown in Table 1. The other parameters are fixed as follows: $\Gamma = \Gamma$(2 1.5 2), T = T(1.5 0.5 1), v_{n1} = 3, v_{n3} = 2, a = 2, b = 0.005.

Table 1: Simulation settings

	Ex-1	Ex-2	Ex-3
The producer is a learning agent	No	No	Yes
Network Externality is considered	No	Yes	Yes

3.2 Simulation results

Figure 2 shows producer surplus and the number of users when the producer provides each service for Ex-1. The top three services were S_m = S_m(2 2 2), S_m(2 2 1), S_m(2 3 1). When S_m = S_m(2 3 3) or S_m(3 3 3), the producer profit was high for the number of users. These services can be regarded as those which target niche markets.

Table 2 shows the results of Ex-2; specifically: consumers' utility, producer's profit and the number of users, when each of the services is provided to the market. The numbers shown in parentheses are the differences in the values when compared to Ex-1. The results indicate that network externality can accelerate diffusion of popular services. In particular, the services with higher functions than the most generally-accepted service, S_m(2 2 2), become accepted by users whose original demand levels were not high; positive network externality enhances their motivation.

Table 3 shows the results of Ex-3, where learning convergence is added to the simulation. The producer decided to provide one of: S_m(2 2 3), S_m(2 3 2), S_m(3 2 2), or S_m(2 2 2). As shown in the table, the producer reaped bigger profits in comparison to those of Ex-2. This is because the producer changed the service according to increases in consumer demand. Consequently, the services with higher function levels than S_m(2 2 2) can become widely used in the market. Here we see that a win-win like relationship between a producer and consumers can sometimes be achieved through dynamic interaction amongst them. The mechanisms underlying a "de facto standard" or "brand value" are probably also related to factors such as network externalities.

Table 2: Diffusion of services and market profits for Ex-2

	Consumers' utility	Producer's profit	Number of users
S (2 2 2)	2103 (+1520)	1864 (+648)	466 (+162)
S (2 3 2)	960 (+696)	1550 (+815)	310 (+163)
S (2 2 3)	960 (+696)	1395 (+815)	310 (+163)
S (3 2 2)	673 (+483)	1157 (+495)	257 (+110)

Table 3: Diffusion of services and market profits for Ex-3

	Consumers' utility	Producer's profit	Number of users
S (2 2 3)	1075	1578	351
S (2 3 2)	1075	1755	351
S (3 2 2)	1267	1580	351
S (2 2 2)	1665	1864	466

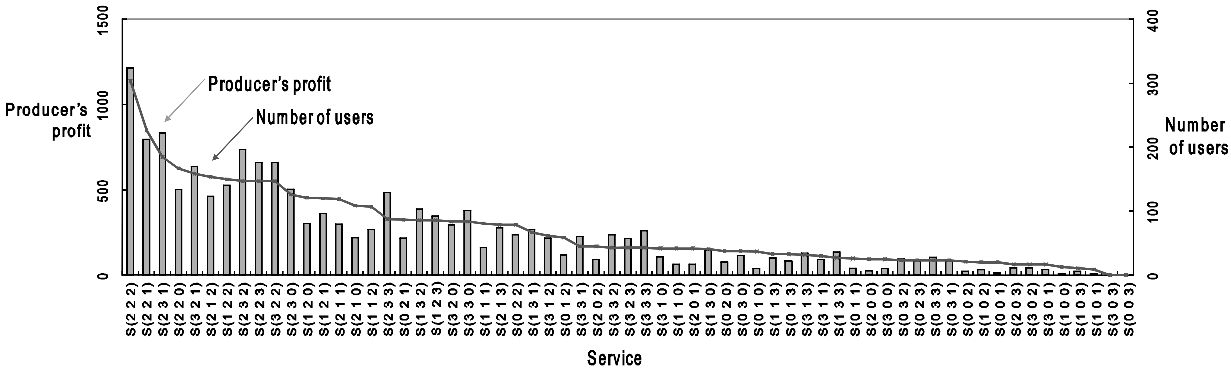


Figure 2: Producer's profit and number of users according to each service for Ex-1

4 VALUE CREATION IN SERVICE MARKETS

This section presents a review and discussion of the model of value creation [4] based on the concept of emergent synthesis [5], and rethinks the proposed model.

4.1 Concept of emergent synthesis

Design, in general, has synthetic aspects; the authors named such design processes "emergent synthesis" [5]. Within this concept, synthesis is classifiable into three classes, with respect to the incompleteness of information about the environment and/or specifications, as described below.

- Class I – Problem with a complete description: if information related to the environment and specifications are given completely, then the problem is described entirely. However, it is often difficult to find an optimal solution.
- Class II – Problem with an incomplete environment description: the specification is complete, but information related to the environment is incomplete. The problem is thus not described completely. Therefore, it is difficult to cope with dynamic properties of the unknown environment.
- Class III – Problem with an incomplete specification: both the environment description and the specification are incomplete. Problem-solving must therefore start with an ambiguous purpose and human interaction becomes important.

Emergent synthetic approaches can deal efficiently with these three problems. Such approaches include: evolutionary computation, self-organization, reinforcement learning, multi-agent systems, and game theory. Traditional methods such as analytic or deterministic ones might not be adequate to solve Class II and Class III problems.

4.2 Classification of value creation

The respective values of products and services are characterized as emergent synthesis problems, by which value becomes apparent through interactions among decision-making agents. Figure 3 presents three value models classified from the viewpoint of emergent synthesis: Providing Value Model, Adaptive Value Model, and Co-creative Value Model. As this figure shows, producers, consumers, and products/services can be treated as agents.

Figures 4 – 6 show detailed value models based on the concept of emergent synthesis.

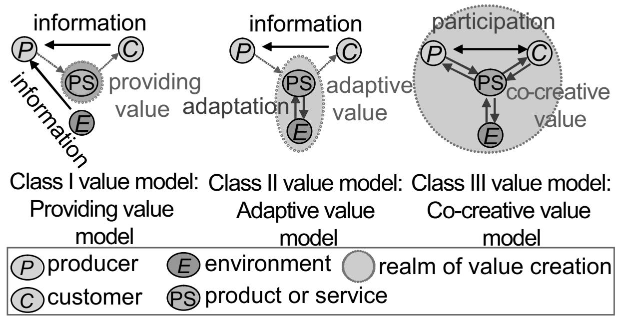


Figure 3: Classification of value models.

In the Class I value model, product and service producers, as well as consumers, are defined independently of their values. The objectives and environment are clear. The model can be described completely using a closed system. However, in most cases, too many feasible solutions exist, which engenders combinatorial explosion and creates so-called NP-hard problems. Therefore, it is necessary to develop efficient and robust search methods in order to identify optimal solutions. In the real world, this model can be applied to mass-produced products or routine services.

In the Class II value model, the objective of the consumer is defined completely. However, the product and service environments are changing and unpredictable. Therefore, the value model is an open system. In our models, the environmental changes can occur in two types: changes attributed to consumers (e.g. diversity of preferences or social influence) and changes attributed to producers (e.g. alterations in technologies and resources). Approaches based on learning and adaptation, such as reinforcement learning or adaptive behaviour based methods, are feasible to resolve this class of problems. The model is applicable to consumer-oriented products or services, such as semi-order-made goods and recommendation systems of books based on collaborative filtering. Adaptive strategies are necessary to respond to the diversity of consumers' preferences.

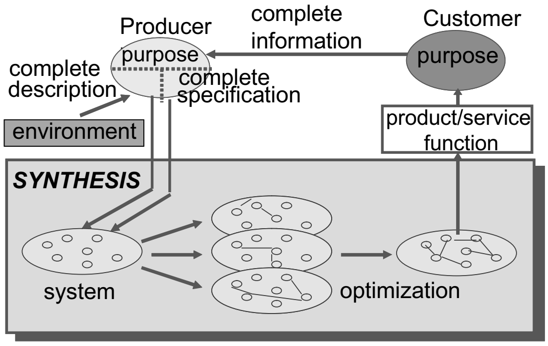


Figure 4: Class I – Providing value model.

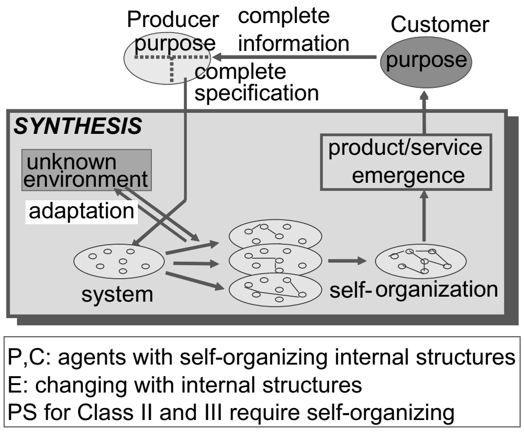


Figure 5: Class II – Adaptive value model.

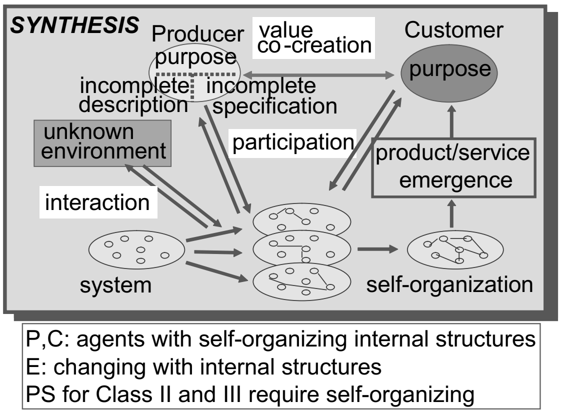


Figure 6: Class III – Co-creative value model.

In the Class III value model, along with the lack of environmental information in advance, the consumer objectives are ambiguous. The producers and consumers are mutually inseparable from the viewpoint of value creation. Consequently, the producers are involved mutually with consumers to co-create the value. In the real world, open source software such as Linux, knowledge databases, and doctor-patient medical services might correspond to this kind of value model. In such cases, it is usually difficult to control the value that emerges through the interaction between producers and consumers. A de facto standard can also be treated as a Class III problem. In such a case, network externalities can play important roles.

4.3 Rethinking the proposed model

Although the simulation models were simple and limited, the results can qualitatively explain real-world business activities. Static and predictable service environments, such as Ex-1, are classifiable as Class I. Here, a mass production strategy might secure the service producer's profit; whilst simultaneously undermining the consumers' profit. Moreover, when a problem is classifiable as Class II, for example Ex-3, a producer's adaptive strategy to service environment changes might enhance both the producer's and the consumers' profit. If Ex-3 is expanded into a Class III problem, it is also notable that value can be co-created through processes of interaction between the producer and consumers.

5 CONCLUSION

This paper presents the multi-agent modelling of service markets. The validity of the model was enhanced by determining consumers' value concepts based on the data from a lifestyles questionnaire. Recent concerns in manufacturing and services have focused not only upon minimizing cost and thus maximizing producer profit, but also maximizing the value of the whole market. For the latter to achieved, it is important to pay attention to the Class III value creation problem. Co-creation is a promising concept to enhance value and to create new value in a society.

6 REFERENCES

[1] Leibenstein. H, 1950, Bandwagon, Snob and Veblen Effects in the Theory of Consumers' Demand, Quarterly Journal of Economics, 64:183-207.

[2] Hultink EJ, Griffin A, Hart S, Robben HSJ, 1997, Industrial new product launch strategies and product development performance, Journal of Product Innovation Management, 14, 243-257.

[3] Katz, M.L., Shapiro, C., 1992, Network effects, Software provision, and Standardization, The Journal of Industrial Economics, 40/1:85-103.

[4] Ueda, K., Takenaka, T., 2007, Classification of Service Model and Value Creation, Proc. of the 9th IEEE International Conference on E-Commerce Technology, Tokyo, Japan, 497–498.

[5] Ueda, K., 2000, Emergent Synthesis, Artificial Intelligence in Engineering, 15/4:319–320.

Strategic Servicification –
A Quality based approach beyond Service-Engineering

Robert Schmitt[1], Sarah Hatfield[1]

[1] Department of Production Metrology and Quality, Fraunhofer Institute for Production Technology, Germany

Abstract

In order to provide hybrid Service-Product-Solutions, especially small and medium sized companies face challenges regarding the adaption of business processes and structures to new tasks in customer interaction and solutions improvement. The basic philosophy of the Aachen Quality Management Model is to harmonise corporate skills with strategic objectives in order to achieve a highest possible overlap with customer requirements. A systematic approach is presented which demonstrates the use of methods to identify customer needs, to compare them with the companies skills in order to fulfil these and to use change management and customer information to improve the strategic orientation of the so-called hybrid producer.

Keywords: Service, Model, Process

1 INTRODUCTION

Service-Engineering has brought on immense insights for the structured process of generating ideas to the implementation of services. So-called hybrid producers rely on methods for Service-Product-Solutions (SPS). In order to become a hybrid producer, traditional producing companies have to consider new challenges such as higher customer interaction, a higher degree of immateriality and the integration of products and services itself [1]. Relying on engineering mechanisms adapted to services, these tasks and their strategic implications have so far been underestimated by the approach of Service-Engineering. Results are small and medium sized firms struggling to render their services beneficial due to insufficient strategic reflection and decision-making as well as lacking change management to make sure all necessary processes are implemented. Servicification needs to consider strategic customer evaluation, a re-structuring of the organization and therefore leadership aspects in order to make Service-Product-Solutions profitable. An approach based on the Aachen Quality Management Model is presented to ensure strategic change based on customer, manager and operational perspectives.

2 THE AACHEN QUALITY MANAGEMENT MODEL

2.1 Motivation of the model

The Aachen Quality Management Model (AQMM) has a relative short history as it was developed in 2007 at the Laboratory for Machine Tools and Production Engineering in Aachen, Germany. This model comprises strategic aspects concerning corporate orientation and are further described in chapter 2.2. The main aim, however, was to create a model which clearly emphasises the importance of customers within engineering processes and to demonstrate how vital information gained from the customer is re-used in the right places of the engineering process to improve the quality of products and processes. Because of this orientation, the model is especially useful for the creation of customer solutions. In the model, the customer dominates the activities in the quality forward chain, i.e. all activities concerning SPS-development, as the customers' requirements serve as triggers. On the output-side of the model the customer is also the main impulse for the quality backward chain (see fig. 1), meaning the feedback of satisfaction data for the improvement of the core processes.

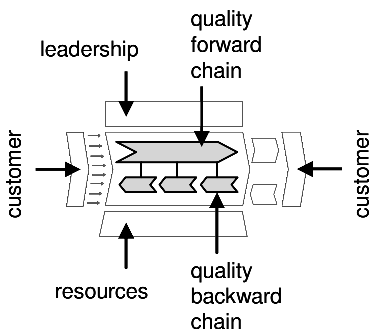


Figure 1 : Main components of the Aachen Quality Management Model

"Quality is the overlap degree of articulated and implicit customer requirements on the one hand and the characteristics of product criteria (including price and time of delivery) on the other hand" [2]. This definition implies the need for identifying customer needs that customers can not even articulate themselves, in order to achieve so-called 'delighting characteristics' [3]. The enablers of high product quality, though, lie within corporate processes and structures. In order to consider these factors, the philosophy of the Aachen Quality Management Model is presented.

2.2 Philosophy of the model

Looking at a company, there are two aspects that lead to customer satisfaction: the company's strategic objectives (orientation) and the corporate skills (ability). Only if these two factors are fulfilled, customer requirements can be met and therefore customer satisfaction achieved [4].

As the achievement of product and service quality relies on three aspects, corresponding perspectives have to be considered when producing companies want to increase customer satisfaction by the delivery of individualised services. These perspectives are the customer, the management and the operations perspective (see fig. 2) and will be defined according to the challenges of hybrid producers.

2.3 The three quality perspectives for SPS

Focussing on the customer, the performance of SPS is vital. The main task here is to transfer customer requirements into delighting ideas for hybrid products. The challenge is to anticipate which SPS characteristics actually meet implicit customer requirements. Here methods are needed to anticipate these needs without having to ask customers explicitly [5].

The management perspective encompasses the whole organisational system. This is where visions and strategies for change processes towards becoming hybrid in the sense of combining products and services from early phases on are developed. The challenge here is to find out which core skills are mature enough to meet the customers requirements and where abilities are yet to be cultivated in order to achieve competitive advantages.

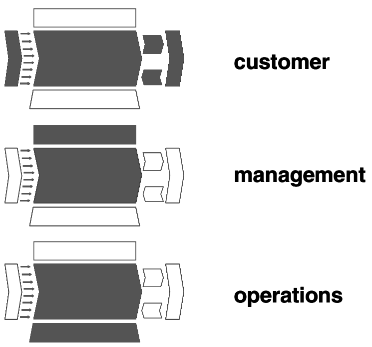


Figure 2: The three perspectives embedded in the Aachen Quality Management Model

Corporate processes are focussed by the operational perspective of the model. Here, resources need to be assessed and provided to ensure the continuous improvement of all relevant activities towards higher product and service quality. Feedback from complaints and other customer-related information needs to be institutionalised and communication between all relevant departments ensured, so that all customer and improvement related information is transparent at all times.

3 METHODS FOR STRATEGIC SERVICIFICATION

Regarding the three perspectives for SPS quality, it is acknowledged that methods exist for the anticipation of implicit customer needs, such as the 'Anticipative Customer Requirement Analysis', a systematic approach including several methods [6]. Also, the Gap-Model developed by Parasuraman and Zeithaml is state-of-the-art, although it does not necessarily consider implicit customer needs but has a strong quality focus [7].

Also, customer complaint processes are fairly well established, although companies, especially small and medium sized ones, still have a lot of potential for optimisation. The challenge for SPS is to find out which departments have to communicate with each other in order to ensure that all feedback is adequately received and in such a form that it can be used for SPS improvement.

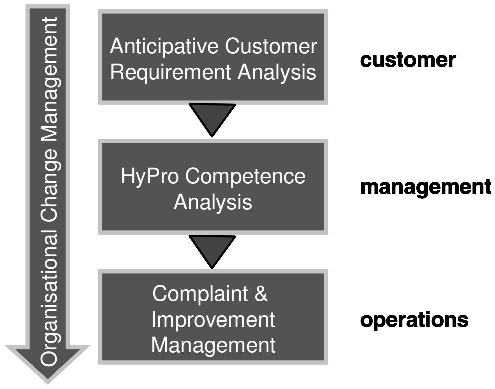


Figure 3 : Overview of methods for Strategic Servicification

The perspective, which has not been considered to such an extent yet, is the management perspective, where the core task lies in identifying corporate skills enabling services to solve individual customers' problems [8]. Corporate skills are here defined by organisational structures, efficiency of processes, qualified customer oriented staff and a visionary leadership. Therefore this aspect will be focussed especially in the following sub-chapter.

Nonetheless, the Anticipative Customer Requirement Analysis provides the input for the estimation of needed corporate skills. Also, operational processes, such as a complaint process, can not be established effectively without being embedded in a holistic strategy and organisational structure, as they would risk missing the corporate goals and therefore not being sustainable [9]. Therefore, the competence analysis for hybrid producers is the missing link between the existing methods form which organisational change measures can be deducted (see fig. 3).

3.1 Anticipative Customer Requirements Analysis

First of all, the processes which the customer runs through with the obtained product are assessed by means of flow charts. In the B2B-sector this can be the handling of a machine or in the B2C-sector the use of the product (e.g. a bicycle) in everyday life. Then, critical incidents are identified, i.e. steps in the process where the customer has problems using the product. Also, the phases before and after use, e.g. acquisition and disposal or sale, are of high interest here as they provide optimisation potential for the customer. This is where the offering company can possibly provide further service solutions, providing it has the necessary competencies.

In order to find out, if a given solution would really fit the customer's needs, an analysis concerning customer values is conducted, the so-called Means-End-Analysis. Functionalities, instrumental values and terminal values are brought in a hierarchical order and assigned to possible solutions.

The third step, in this procedure, is to mirror the customer process with customer values, which is done by the Process-Empathy-Matrix. Here, gaps are identified, i.e. process steps where values are not met (indication for marketing or service improvement) or values which could be addressed in certain process steps in form of new services. At this point concrete service ideas are articulated and prepared for design.

3.2 Complaint and Improvement Management

The main task of the complaint process is to identify problems and to allocate them to distinct causes. As soon as the corresponding processes and responsibilities are pin-pointed, counter-measures are defined and tested.

Such a process needs to be prepared by Service-Blueprints, so that the right people can be triggered with the right information. In larger companies complaint or improvement agents might be of use when processes are multi-fold and structures not easily comprehended. But not only complaints have to be managed but also lessons-learned by customer-contact staff and positive feedback.

Service-Blueprints arrange all necessary measures in flow-charts whereby all roles and responsibilities are allocated to swim-lanes. By this, the interfaces of all relevant departments becomes transparent. Such Blueprints can be used to model IT- workflows, in form of reference-processes, and serve for system integration. Another benefit is that potential bottlenecks of information can more easily be identified and by means of re-modelling eliminated.

3.3 HyPro Competence Matrix

The HyPro Competence Matrix proves as the missing link between the assessment of customer requirements (step 1) and the improvement of service processes (step 3). In order to fulfil requirements, the organisation needs to know which corporate skills need to be improved. But also, on the other hand, corporate skills in form of processes and structures need to be defined, in order to enable improvement processes.

The competence matrix for hybrid producers relates customer requirements to corporate skills and assesses their correlations. Input are the requirements from the Anticipative Customer Requirement Analysis or any customer workshop and the HyPro Competence Catalogue. This catalogue comprises aspects of process and structure organisation as well as leadership and staff issues. It has been established by three phases of data collection and reduction and will be explained in the following chapter.

The customer requirements are prioritised and weighted according to the importance of the requirements for the customer. The corporate skills of hybrid producers are also assessed regarding their degree of realisation in corporate activities. Then, in expert workshops with key-account managers, the impact of corporate skills on the fulfilment of customer requirements is assessed by correlations, similar to the Quality Function Deployment method.

Form the resulting impact, visualised in a scatter plot with a portfolio (see fig. 4), the experts can see which requirements are satisfied with solid skills and therefore pose a **competitive advantage**. This occurs when hybrid producers skills are mature and also perceived by the customer.

Unstable customer advantages are achieved by requirements perceived as fulfilled by the customer but where the skills in the organisation have not reached a maturity to always guarantee this fulfilment. The solution here would be to establish these skills by means of organisational development and change management.

Corporate advantages are achieved when skills are implemented in corporate activities but do not help if they are not seen as such in the eyes of the customer. Customer communication and marketing are needed in order to enhance these benefits for customer solutions.

The least favourable case occurs when neither skills are up to scratch nor customers requirements are fulfilled and therefore mean a **competitive disadvantage**. In this case companies have two options: either neglect of this business field as it would pose too high an effort in order to reach requirement fulfilment or massive build-up of the lacking corporate skills requiring professional change management.

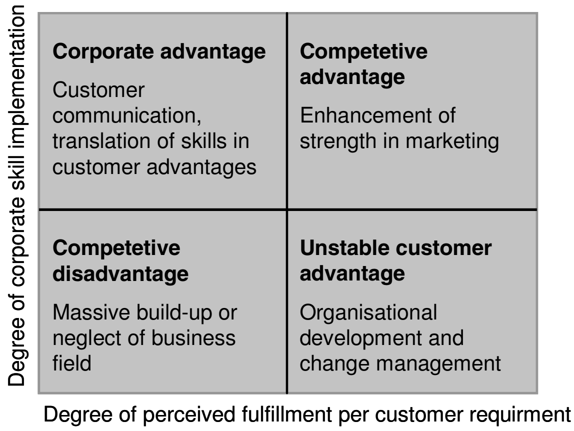


Figure 4: HyPro Competence Matrix

In the scatter plot, now, it can be seen which customer requirements need to be fulfilled better in order to achieve competitive advantages. Per requirement, the main corporate influence factors can be identified in the correlation matrix. As the development of corporate skills often requires organisational development or change management, project plans can be deducted by prioritising the least fulfilled customer needs and by paring them with factors of highest influence.

4 METHODOLOGY FOR COMPILATION OF HYPRO COMPETENCE CATALOGUE

First of all, **expert workshops** were conducted in three medium sized companies in order to identify core skills that hybrid producers need regarding special processes of product and service development and delivery as well as organisational settings in order to give the processes the adequate framework and communication between all participating actors. Scientists from three different German research institutes were asked to use the brainstorming method to name all skills hybrid producers need to have in order to meet the requirements resulting from the degree of immateriality of SPS and close customer interaction as well as individualisation.

65 requirements were named and then clustered by **qualitative content analysis**. From this, 7 categories of requirements resulted: development processes (combination

of products and services), corporate structure (strategic, cultural and structural orientation), business processes (enabling lessons learned and improvement), market & marketing (pricing of SPS and market knowledge), staff qualification and motivation (service orientation), competition observation and customer interaction / integration

In a second step, **staff interviews** in the corresponding firms were conducted in order to find out how they describe skills according to the named requirement categories. the Staff interviewed mainly came from departments with strong customer interaction and also from producing parts of the company, in order to secure the reflection on vital organisational interfaces. They were then asked how well the skills, on a scale of 1 to 10 were currently manifested in the company and, also on a scale of 1 to 10 how important they are. From the highest discrepancies, the most relevant fields of action for strategic servicification were deducted.

The discrepancy analysis resulted in a catalogue of corporate skills for hybrid producers and are explained in the following:

Organisational structure includes all aspects of formal structure: service departments' degree of independence, job specification and formal institutions such as regular meetings and support tools. Especially, interfaces between service departments and other departments are of high interest.

Development and business processes focus on all customer relevant procedures, integration of customer in development, interaction with customer during delivery and management of complaints and lessons learned. Also the company internal interaction plays a vital role as all steps can have an influence on customer satisfaction, whether direct or indirect.

Leadership issues deal with the enabling of the afore mentioned skill categories, meaning the core processes have to be controlled and role model functions taken on an exercised by the management. Goal definitions with all relevant staff concerning SPS and their contribution hereto as well as incentive management in order to raise commitment are aspects here.

These three categories and their requirements remain to be finally verified by the implementation in a diagnosis tool currently developed at the Fraunhofer Institute for Production Technology. The diagnosis tool aims at comprising maturity stages. The presented systematic for strategic servicification is not limited by the fact, that this tool is under development as it will be used for an initial self-check of organisational readiness for hybrid producers, whereas the systematic focuses on the post-initial-phases, where the companies already follow strategies to become hybrid producers. This is where the HyPro Competence Matrix becomes of interest but works without the definition of categories but with the complete requirement catalogue.

5 SUMMARY

This article describes which framework is needed to develop service activities in alignment with a companies strategy. First, the Aachen Quality Management Model is presented which describes how corporate skills and strategic orientation need to be linked and focus on customer's requirements. Second, three perspectives are needed in order to achieve the highest possible overlap of the above mentioned factors and therefore the desired level of quality for service solutions. These three perspectives are provided with methods, so that latent customer requirements can be assed, the corresponding companies' skills identified and improvement measures be taken. Corporate skills are here defined as processes and structures as well as leadership that enable companies to align all their service relevant activities in one strategy. As the HyPro Competence Matrix provides the missing link between existing methodical approaches, the development of the skill criteria is described. By means of expert and staff interviews a broad data base was created, which was then, step-by-step, structured by cluster and content analysis. The HyPro Competence Matrix enables the derivation of different strategies in order for the company to enhance skills by focussing on customer requirements. Means of organisational development and change management are resulting implications and can be interpreted by the interpretation of the resulting portfolio.

The final verification of the skill categories is still subject to further research as the implementation of all resulting development and change activities remains to be conducted on a large scale basis, i.e. with more companies than those included in the research so far.

6 ACKNOWLEDGMENTS

The support of the German Federal Ministry of Education and Research (BMBF) as well as the Project Management Agency – part of the German Aerospace Center (PT-DLR) is greatly acknowledged. We extend our thanks to our colleagues at Mercator School of Management, Department of Management & Marketing under Prof. Schmitz.

7 REFERENCES

[1] Glauner, C., 2007, Normen und Standards in der hybriden Wertschöpfung. Vortrag des DIN Deutsches Institut für Normung e.V., Congress: 10. Aachener Dienstleistungsforum.

[2] Masing, W., 2007, Handbuch Qualitätsmanagement, Hanser, München.

[3] Kano, N., 1984, Attractive Quality and Must-be Quality. In: Hinshitsu: Journal of the Japanese Society for Quality Control, Vol. 14 – No. 2., p. 39-48.

[4] Schmitt, R., Betzold, M., Hense, K., 2007, Das Aachener Qualitätsmanagement Modell. In: Pfeifer, T., Schmitt, R., Masing-Handbuch Qualitätsmanagement, Carl Hanser Verlag, München, p. 39-68.

[5] Scheuing, E. E.; Johnson, E. M., 1989, A proposed Model for new Service Development. In: The Journal of Services Marketing, Vol. 3 – No. 2, p. 25-34.

[6] Lorenzi, P.; 2004, Service Scout – Dienstleistungsbedarfe antizipativ erkennen und in Netzwerken systematisch erfüllen, Cuvillier, Göttingen.

[7] Zeithaml, V.; Bitner, M. J.; Gremler, D. D., 2006, Services Marketing – Integrating Customer Focus Across the Firm, McGraw-Hill, Boston.

[8] Petroni, A., 1998, The analysis of dynamic capabilities in a competence-oriented organization. In: Technovation Vol. 18 – No. 3, p. 179-189.

[9] Schuh, G.; Klotzbach, C.; Gaus, F., 2007, Designing services based on intelligent press-die-systems. In: 14th CIRP Conference on Life Cycle Engineering.

From product to service orientation in the maritime equipment industry - a case study

Detlef Matzen and Tim C. McAloone
Section for Engineering Design and Product Development, Department of Management Engineering, Technical University of Denmark, Denmark

Abstract
In the shipping industry, operational performance of ships and their equipment is crucial to business. Suppliers of machinery and equipment are aware of this situation and see business development potential in setting up service systems that are dedicated to ensuring the performance of their products in operation.
In this paper we present a case study of a shipping equipment manufacturer that is currently shifting business focus from manufacturing towards services delivery.
Using a modelling scheme to differentiate and categorise different development tasks within the frame of business development towards service oriented business, the case delivers insights into the broader context and product related parameters influencing the options and requirements for service system development.

Keywords:
Product Service System; Service Design; Case Study

1 INTRODUCTION

This paper presents the first findings of research in close collaboration with a Danish company that is supplier of equipment to the shipbuilding industry and currently is initiating a shift of business perspective, from product orientation to combined product/service orientation.

In the shipbuilding industry, like in many other industrial sectors, the continuing market globalisation both opens opportunities in terms of a rising number of potential customers and represents threats, due to the growing number of competitors worldwide. In general the Danish shipbuilding industry traditionally relies on the longevity and high technical and functional qualities of their physical products creating competitive advantage. These quality parameters are increasingly under pressure in current global markets, where many competitors offer functionally comparable components at substantially lower prices, and where quality differences are not readily visible to customers. Furthermore, the supplier companies' position as sub suppliers in the supply chain to contracting shipyards traditionally leads to compromises for these companies, as the shipyard often makes decisions based upon low cost, in order to keep their own sourcing expenses (and therefore the initial cost of the resulting ship) as low as possible.

Some suppliers now see the potential of providing support services for their equipment (and related installations) for two reasons. One is the obvious potential business opportunity, as the total cost of ownership (TCO) of virtually any equipment category is much higher than the installation cost, and as service provider the manufacturer has potential of diverting part of the resulting economic transaction value back to the manufacturing company [1]. TCO thinking is gradually becoming evident in some ship owners' acquisition activities, which also acts as a driver to service provision. The other reason is the ability to optimise the use phase performance of the delivered products by e.g. training of operators, close performance measurement or preventive maintenance, thus delivering higher value to the end customer of the equipment – in this case usually shipping companies or operators.

The challenges connected with shifting business perspective from product manufacturing to service delivery are manifold, and the following list represents only some of the important aspects [2] that should be understood and considered:

1. understanding of products' life phases and activities
2. identification of valuable service offers
3. development of delivery networks
4. development of internal delivery systems
5. marketing of service offers to (new) customers
6. altering the practices of users and customers
7. dynamically adapting and improving service offers

We focus in this paper on the first two items in the above list, namely the identification of potential service activities, based on an analysis of use phase transformations in relation to the core company product. Although only a single case is presented, the study is part of a larger collaboration project with a number of companies in the maritime sector. The aim of the research effort is to apply different development modelling techniques in concrete service development projects and then analyse the effect and future potentials of the modelling techniques in industrial use.

2 THE CASE COMPANY

The case study, on which this paper is build, was conducted within the frame of a larger research cooperation project with a number of enterprises in the maritime equipment and shipbuilding industry in Denmark.

2.1 Framing the case study

The case company manufactures equipment for installation in ships, maritime constructions and the land based construction industry. In 2006 the company had 550 employees generating a turnover of roughly € 100 million. One of their business areas is water-based fixed fire fighting (FF) installations on ships, setting the context for the case investigation. The maritime FF business area is chosen due to the obvious options of offering inspection services, which are mandatory for safety equipment on ships as regulated by IMO in the SOLAS convention [3] and enforced by national authorities and class societies representatives around the globe.

Organisation

The business unit is organised following a traditional pattern, with departments for (regionally and application divided) sales, contracting and technology, supported by manufacturing, supply chain and the service and parts departments, which are shared with the other business units in the enterprise. The Service and Parts department (S&P) of the company has been the focal part of the organisation throughout the case study, currently in the process of restructuring and developing the company's approach to after sales services and business.

Sources of the case study description

The case study was conducted within the frame of a concept development project, where the authors have developed a number of potential concepts for a service offer portfolio throughout the autumn and winter of 2007. Information was gathered through 7 meetings in the form of informal interviews and workshop sessions with department and project managers of the FF business unit. Throughout the study, the following employees of the company have been informing the researchers:

- Managers of S&P, Contracting, Sales and Technology
- Project Managers within Contracting and S&P development.

The majority of information was sourced from the partners in the S&P department.

2.2 Current product and contracting process

FF installations are currently sold mainly through shipyards, which compile and negotiate the so called makers list with their client, the prospective ship owner company.

The contracted FF system's general specifications are defined in negotiation between the FF salesman and the shipyard. Based on the resulting contract, the contracting department project manager designs and configures the system and sources the components from sub suppliers and the manufacturing department.

The various sub systems are assembled and tested in the company manufacturing plant, then dismantled and shipped to the contracting shipyard for installation on board the ship. After the system is completely installed on board the ship by the shipyard personnel, the project manager organises a final compliance test and commissions the system. Commissioning is usually supervised by representatives of the shipyard, ship owner and the classification society registering the ship.

The FF system shipped from the company includes pump units, feed water and pressure tanks, control systems, nozzles and sensors, depending on application, size and integration with other safety systems on board. The shipyards installation process is guided by a manual, which also defines specifications for components the shipyard is sourcing locally, such as piping and pipe brackets.

After the system is commissioned and the ship released for its test voyage, responsibility is transferred to S&P, who are organising maintenance during the guarantee period of usually 1 year and supply spare parts to customers. The guarantee period of ships is usually concluded with a technical check of the ship by the ship owner and select sub suppliers, not always including the FF company's S&P department.

2.3 Use phase requirements

Being a safety installation, FF systems are not released or operated on a regular basis. Daily use is therefore limited to crew drills, technical inspections, component or system tests and preventive maintenance of the system by the ships crew or shore personnel, which is to be conducted according to the manufacturer's instructions following weekly, monthly, quarterly, yearly and five-yearly cycles. If the system is actually released in either a test or an actual emergency situation, the system needs to be checked and 'restarted' properly in order to function as intended.

In principle, the system manufacturer thus defines how, when and by whom the system's operational state is ensured. Nevertheless the operator or crew do not always comply with the defined maintenance rules, leading to both safety hazards and the risk of detention of the ship if the system proves non operational in random or periodic checks by class representatives or port authorities. E.g. the ClassNK PSC report 2007 [4] documents that about 20% of all detainable deficiencies arise from faulty fire safety measures on board. The FF system is one of the measures in that statistical category.

As the fixed FF system is one of many safety installations onboard, and as the type and design of the system installed by the case company is one of many alternatives in the industry, noncompliance with the defined maintenance rules is not necessarily discovered nor corrected in time, resulting in the earlier mentioned risk situations.

3 CONCEPT MODELLING FOR SERVICE ORIENTATION

The authors postulate that services and products are bound together by the use activity [5], that services are delivered or executed in a so-called 'transformation system' (TS) [6], and that both products' and services' business aspects are based on value relations to the customer. The objective of this clarification is to allow the conceptualisation of new services to happen as a systematic pursuit of new solutions as it is known from engineering design.

3.1 The transformation process as design object

The transformation system model of Hubka and Eder [6], which is illustrated in Figure 1 below defines a transformational activity, which transforms an operand (a combination of material, energy and information) into a new state by the application of operators, which are categorised into technical systems, human operators, information or knowledge systems and management or goal systems. The process takes place in space and time within an active environment, which is influencing the transformation and its outputs. Trough the execution of the transformational process (TP) value is created for one or more actors, traditionally the user initiating and supervising the TP.

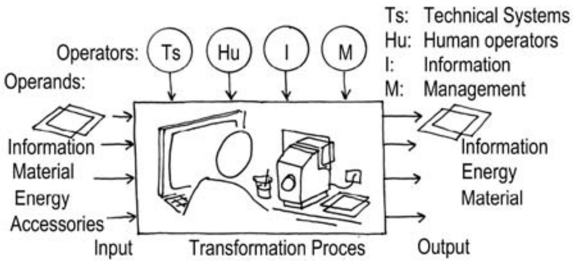


Figure 1: The transformation process model (redrawn from [6]).

In product-oriented business, the manufacturing company codes functionality necessary for sufficient performance into the technical system (Ts). The prospective product life of the Ts is analysed to find the important TP's which in turn set the requirements for the design of the Ts.

For the FF case this corresponds to the specification of e.g. necessary spray patterns for the nozzles or correct flow and pressure capacity for pump units.

When shifting to service-oriented business, the provider regards the ability of executing the TP as his delivery to the customer, thus widening the object of design to include the other operator categories and possibly the TP itself.

For the FF case that corresponds to e.g. defining whom will be allowed to execute inspections, how exchanged components are returned to the manufacturer for test and reuse and how the validity of inspection certificates is ensured.

The design object is therefore not only the Ts but also delivery systems for other operators and operands, and possibly a redefinition of parts of the product life and the inherent TP's.

3.2 Setting the product life TP's into context

As mentioned earlier, the raison d'être for the product and transformation is the generation of value for actors, predominantly the customer organisation.

Originating from a product life thinking approach such as described by Olesen [7], the dimension of the social construction [8] must be added, which set's the context for single life phases of the Ts and defining the life phase system network.

Setting up the network for single life phases and on a more detailed level for single TP's being part of that life phase, a method similar to Customer Value Chain Analysis [9] can be applied. As different actor networks are formed in relation to the different TP's, considerations on the network constitution, the relations, in terms of operator input and output and creation of value for the involved actors must be repeated for all relevant constellations.

The actor networks emerging in the case study include the shipyard currently responsible for installing the FF system, the shipping company contracting the ship, workshops and other partners delivering concrete services and class societies responsible for the classification of the vessel. Within those organisations a number of subdivisions must be defined, as not all employees within the single organisation have aligned interests.

As an example, the ship-owner's operation office is highly interested in information on the performance of the crew on board and thus will be welcoming the option of checking to what extend the crew executes their inspection tasks correctly in compliance with the manufacturers instructions. The crew on the other hand will usually oppose being put under surveillance in that respect.

Obviously the interests and value perceptions regarding the TP's of the FF system are differing fundamentally, and it is therefore important to relate the effects of reorganising the product life setup according to whom will be influenced, and whether the proposed changes can represent a higher value for the different actors.

3.3 Structuring a solution concept space

Having defined the requirements for modelling the single TP's and liking these to the multitude of actor networks formed, gives a foundation allowing a structured identification of potential development areas and links to the actors for whom the resulting servicing activities would be of value.

Unfolding the relations between transformations in the product life and the related value creation for identified actors is illustrated in Figure 2. The different identified TP's are arranged in a 'near chronological' order from left to right. In the actual product life, the processes will not occur in the modelled sequence, and there will also be differences in the frequency of occurrence, rendering some of the TP's more influential than others. The development team must be aware of these differences and prioritise their importance.

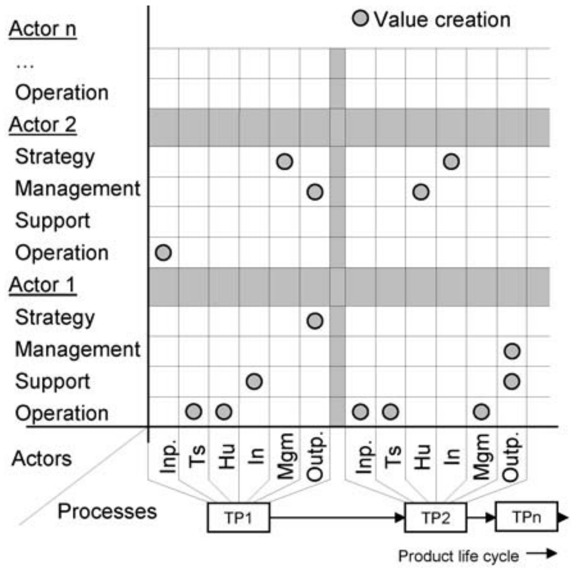


Figure 2: Unfolding potential value propositions related to product life processes.

The other dimension of the matrix specifies the identified actors in the use phase systems. Organisational actors will usually be subdivided according to functions or tasks, as different and sometimes conflicting interests exist. In general there will be divisions between operational, support, management and strategic functions of the company, as indicated in Figure 2.

For every crossing between actor and process operator/operand the design team can now investigate and decide on whether and why there can be proposed value creation from the prospective activity. In this way the matrix can guide the identification of important activity areas and deliveries that should be considered part of a service offer.

4 A SERVICE ORIENTED APPROACH FOR BUSINESS DEVELOPMENT

4.1 Service oriented case development potentials

Based on the knowledge gathered from the analysis of the current situation, the researchers have applied the modelling scheme described in section 3. Using the modelling scheme, a palette of interlinked development concept proposals have been synthesised, supported by the domain insights into the technology, application area and company competences supplied by the case study contacts within the company.

The development concepts range from proposals for further development of the company's nozzle programme to knowledge and certificate management for the ships operator, all related to the optimisation of the FF installations service life. Examples of concept elements are:

- Training programmes for key crew personnel, supporting the correct use and maintenance of the system
- Communication technology and systems enabling remote assessment of the crews self inspection performance on the ship
- Certification management systems enabling documentation of the maintenance history and regulatory compliance of single installations

As the examples show, the potential concepts fall into the modelling scheme's 4 categories of process operators (technical, human, information and management). The concepts are also oriented towards actors in different levels of both the provider and customer organisations:

- The nozzle concept supporting the physical inspection work of crew and service personnel.
- The training programme improving the crews knowledge about the FF system thus lowering the risk of faulty operation.
- The communication technology enabling performance measurement of the customers crew.
- The certification system documenting and communicating the customer's ships operational quality towards external actors like e.g. class society representatives.
- Feedback loops from inspections and via communication technology enable the company to identify future tasks due to the direct access to maintenance and use information.

Currently the company is in the process of evaluating the feasibility and market potentials of the synthesised palette of development concepts.

5 DISCUSSION AND CONCLUSIONS

Through the case study project, a large number of promising and realistic options for service delivery have been generated. The service oriented approach of analysing and synthesising solution concepts on the level of use activities, modelled utilising the transformation system model, enabled the team in identifying these potentials. By connecting the identified processes with their corresponding actor networks, insights into the value creation possible by service delivery are created. The approach thus is capable of opening and structuring the solution space.

On the other hand, the propositions can not be evaluated without further deep analysis of the single concepts, linking of related and synergetic tasks and evaluation of feasible delivery and communication channels.

In the case study it was identified that the company, although able to deliver high value services, only has limited access to the ship-owners representing the target customer of the service offer. This is predominantly due to the company's current position in the supply chain, where there is little possibility of approaching the prospective service system client prior to the contracting phase, and the service offers benefits thus can not be actively sold to the clients.

At this point, the approach seems well suited for identifying and structuring the solution space in service system development and can guide in the identification of open tasks in relation to building scenarios for future integrated service offers.

6 REFERENCES

[1] Wise, R. and P. Baumgartner, 1999. Go Downstream: The new profit imperative in manufacturing. Harvard Business Review, 77(5): p. 133-143.

[2] Matzen, D., A.R. Tan and M.M. Andreasen. 2005. Product/service-systems: Proposal for models and terminology. in Proceedings of 16. Symposium "Design for X". Neukirchen, Germany.

[3] IMO, International convention of the safety of life at sea, 1974 - SOLAS. 2004, IMO.

[4] Annual Report on Port State Control. 2007, Nippon Kaiji Kyokai: Tokyo. p. 44.

[5] Matzen, D. and M.M. Andreasen. 2006. Opportunity parameters in the development of product/service - systems. in Proceedings of DESIGN 2006. Dubrovik.

[6] Hubka, V. and E. Eder, 1988. Theory of Technical Systems. Berlin: Springer-Verlag.

[7] Olesen, J., 1992, Concurrent development in Manufacturing - based on dispositional mechanisms. DTU, Lyngby. p. 154.

[8] Bijker W.E., Editor. 1987, The Social Construction of Technical Systems: New Directions in the Sociology and History of Technology, MIT Press: London. p. 83-103.

[9] Donaldson, K.M., K. Ishii and S.D. Sheppard, 2006. Customer Value Chain Analysis. Research in Engineering Design, 16(4): p. 174-183.

Knowledge Generation as a Means to Improve Development Processes of Industrial Product-Service Systems

Seliger, Günther[1], Gegusch, René[1], Müller, Patrick[2] and Blessing, Lucienne[3]
[1] Assembly Technology and Factory Management, TU Berlin, Germany
[2] Engineering Design and Methodology, TU Berlin, Germany
[3] Université du Luxembourg, Luxembourg

Abstract
In business-to-business and also business-to-customer markets there is a tendency towards combined offers of products and services, which are sold in one package to fulfil customer needs. These combinations of products and services are called product-service systems (PSS) or industrial product-service systems (IPS^2) in case of industrial application. Feedback approaches for knowledge which can be captured during the delivery and use phase of IPS^2 are discussed in this paper. This knowledge is notably valuable for the development process of IPS^2. The focus is on knowledge generation systems capturing information related to products and processes anywhere available. These, so called Virtual Life Cycle Units (VLCU) collect operation data of products and services and generate knowledge about the systems behaviour or usage. Both development and operation benefit of captured information and generated knowledge to improve systems development and operation projects. The role of the knowledge generation, enabled by the VLCU, in the IPS^2 development process will be described in detail using different PSS examples and business models.

Keywords:
Industrial product-service systems, manufacturing system, development processes, knowledge generation

1 INTRODUCTION

In business-to-business and also business-to-customer markets there is a tendency towards combined offers of products and services, which are sold in one package to fulfil customer needs. These combinations of products and services are called product-service systems (PSS) [2] [3] or industrial product-service systems (IPS^2) in case of industrial application. The adaptation of products and services to continuously changing technical requirements, application areas and user demands is crucial to select the right PSS business model (e.g. an availability-oriented PSS offer). This requires new integrated methods, process models and tools for planning, development, delivery and use of PSS to exploit their full potential and to ensure the competitiveness of these systems.

Knowledge feedback approaches for PSS development are focused in this paper. Knowledge can strategically be collected during the delivery and use phase of a PSS. This knowledge is notably valuable for the development process of a PSS and for its delivery. Different loops of feedback will be characterized and discussed. First, customer feedback integration, and second, knowledge generation systems capturing information related to products and processes. Such systems collecting operation data of products and services anywhere available to generate knowledge about the systems behaviour or usage are called Virtual Lifecycle Units (VLCU). Developers can benefit of information and knowledge provided by VLCUs to improve their systems in further development projects. Furthermore, they have the ability to request missing information by designing a proper VLCU integration during the PSS development.

1.1 Basic ideas of product-service systems

In the area of high cost machinery PSS are sold instead of stand-alone products or services to exploit earlier unused economical and technical potentials or to enhance the value for the customer, cf. [1]. As in other branches, customer needs are not any longer simply reduced to a single need for product ownership. The lifecycle-orientation of PSS leads to aspects like availability, flexibility of system reconfiguration or sustainability. This requires the partial substitutability of products by services and vice versa. Those aspects are considered in a PSS scope. Additionally, proper feedback mechanisms to capture important or missing lifecycle information have to be implemented.

In research product-service systems are often linked to sustainable development or eco-design [3]. For instance, smart product (e.g. car) sharing strategies combined with supporting services could be used as an example to demonstrated potential of dematerialisation to face lesser resources.

Potentials to differentiate from competitors and to implement anti-piracy measures [4] lead to new business models. These help to bind customers for a longer period than typical with traditional product purchase and they offer new ways for economic benefits.

Social aspects are also often linked into PSS by various authors. Terms like ‘social responsibility’ or ‘green business’ become new drivers to focus on new forms of system characteristics, which seem to be relevant in PSS marketing.

1.2 Research in PSS-related areas

Close to product-service systems we find technical disciplines like 'Service Engineering (SE)' [5] [6], 'Integrated Product and Service Engineering (IPSE)' [7], 'Functional Sales (FS)' [7] or 'Functional Product Development' (FPD) [8].

The research project SFB/TR29 'Engineering of Industrial Product-Service Systems', funded by the German Research Foundation (DFG), focuses on the integrated planning, development, delivery and use of products and services [9]. The central idea in this research project is that the potential of product-service systems can only be fully exploited if the processes and methods for the development of products and services are adjusted to each other. The integrated view on products, services and their immanent software assemblies in all of the four phases planning, development, delivery and use is considered as an essential characteristic of PSS, cf. [1].

In this paper the research topics and approaches of two project partners of the SFB/TR29 will combined: The 'Development processes of Product-Service Systems' will be linked to the 'Automated Knowledge Generation'.

2 KNOWLEDGE EXCHANGE BETWEEN PSS DEVELOPMENT AND DELIVERY PROCESSES

As in every development process knowledge about the systems lifecycle (ramp-up, use, optimization, etc.) is vital to develop effective, highly efficient systems, cf. [10]. In contrast to traditional product purchase PSS business models allow the PSS provider a closer contact to its PSS. This offers PSS providers opportunities to learn more about their product-service systems. A PSS business model which implements for instance availability-oriented product use (product possessed by the provider and 'used' by a customer + availability ensuring service delivered by the provider) gives them a platform to collect information which earlier was only accessible for the customer or by feedbacks of service personnel. VLCUs are an enabler to support direct feedback additionally to feedback from technicians and customers, see Figure 1.

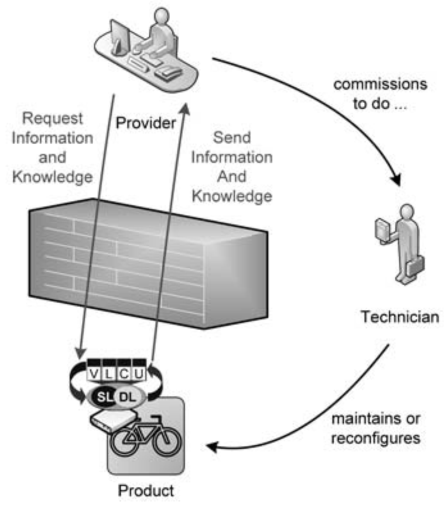


Figure 1: Direct feedback from a product enabled by a VLCU.

2.1 Floating borders of the development process

A closer look shows that there are two different types of knowledge and information exchange.

1. Knowledge and information 'circulating' within the delivery process of a PSS used to manage the delivery process.
2. Knowledge and information about the delivery process flowing into following or ongoing development processes.

Additionally the borders between the PSS development and delivery phase are dissolved by two other aspects:

1. The close contact between the provider and its PSS enabled by applied services (executed by technicians or technical devices).
2. The continuous customer integration already during development: The customer integration is often mentioned as very important for PSS development [11], cf. Figure 5. It is important to gather customer needs and to make customer activities (as external factor, cf. [12]) during the systems use visible to the designer of a PSS [2] [13].

Figure 2 visualises the overlap of development and delivery and the information flow. The PSS is designed and later introduced into the market. During the systems operation (delivery and use phase) feedback on the operation can be used for reconfiguration (extension or a down-grading of service contents or product modules), redesign or further development of the PSS during its lifecycle. Thus, the PSS development phase does not end with the preparation of the production and operation documentation as e.g. in the general approach of VDI 2221 [14]. Production and consumption, in contrast to product manufacturing and purchase, are not clearly separated, caused by the characteristic of services that production and consumption is simultaneous [2].

The dynamic character of product-service systems finally determines a kind of 'fuzzy-backend' [15] of the development process. Simplified it is shown in

Figure 3, which is an adoption of the model for Integrated Product Development by Andreasen and Hein [16].

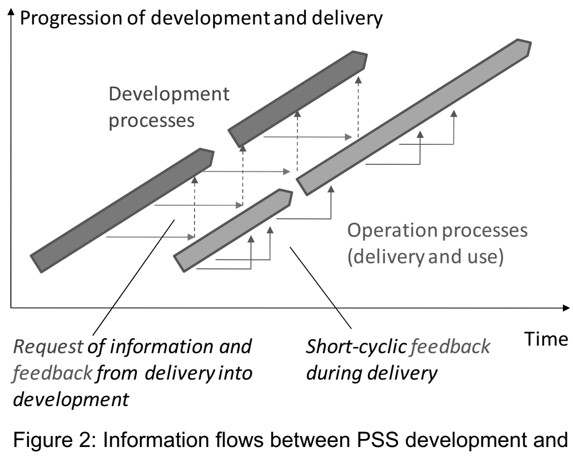


Figure 2: Information flows between PSS development and operation (delivery and use).

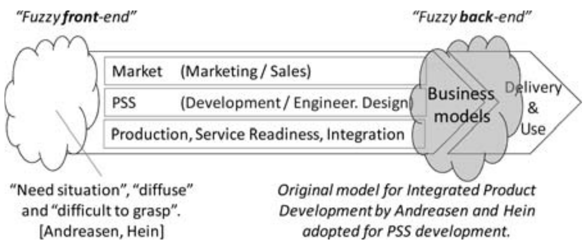


Figure 3: Integrated PSS development incl. fuzzy back-end.

Despite the 'fuzzy back-end', the development process has to be well coordinated and PSS releases have to be consistent after every reaction on (cyclic) effects which 'return' from the lifecycle of a particular PSS (Figure 4). According to our view different types of information flows are characterized by Figure 5. Next to mechanisms to capture customer feedback technical solutions, especially VLCUs, can be used to generate a continuous flow of information from the delivery and use phase into further development or system adoption/reconfiguration. This can be information of service personnel states, service quality and time or information on machine operation details (e.g. processed vibration or load details). Designers of PSS profit by using these conditions to request missing information based on 'feed-forward' mechanisms to enhance the PSS further development. 'Random' feedback and requested information together build a broader basis for efficient PSS design.

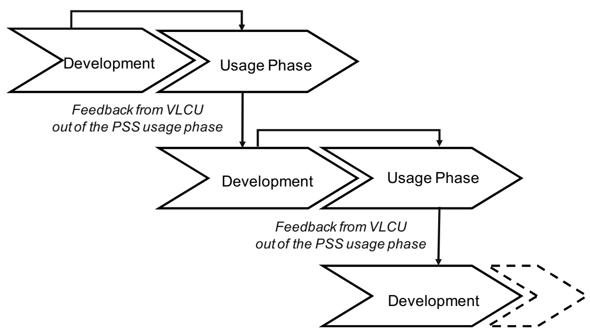


Figure 4: Example of feedback waterfall.

Both variants of information feedback (customer feedback and requested information feedback enabled by technical devices like VLCUs) has to be planned smartly in order to support decisions on how to react to findings gained from delivery processes. Reactions, for instance, can be marginal short-term changes in the coordination of delivery processes or result in heavy impacts on the development processes of a product-service system.

In the following sections there will be a focus on the capabilities of VLCUs to show which information can be requested by PSS designers and how additional (unexpected) knowledge can be generated. Two brief examples will be used to describe situations motivating this approach. One detailed example demonstrates the influence of different PSS business models on the PSS.

3 KNOWLEDGE GENERATRION DURING OPERATION

3.1 Fundamentals about the VLCU

Effective and efficient adaptation can help to reduce resource consumption by e.g. extending the product's life span and by supporting usage-oriented business models. As manufacturers and PSS providers are confronted with increasing demands for product and process availability, reliability and safety, the assessment, prediction, diagnosis, monitoring and control of past, current and future product and process behavior is desirable. Moreover these data and information deliver potential sources to find knowledge like inferences about conditions, wear or quality aspects and deliver so knowledge about the behavior and usage of their products.

Adaptation is facilitated by product accompanying information systems which are integrated into products and components and that are capable of acquiring, processing, and communicating relevant product and process data and information and generate knowledge during the entire life span [18].

The VLCU is a concept for a PSS accompanying information system for knowledge generation and automation, see Figure 6. A VLCU acquires via sensors or IT-Documents data and information from PSS operations, wherein the term operation includes products and services. This may include technical, economical, environmental and social product and process

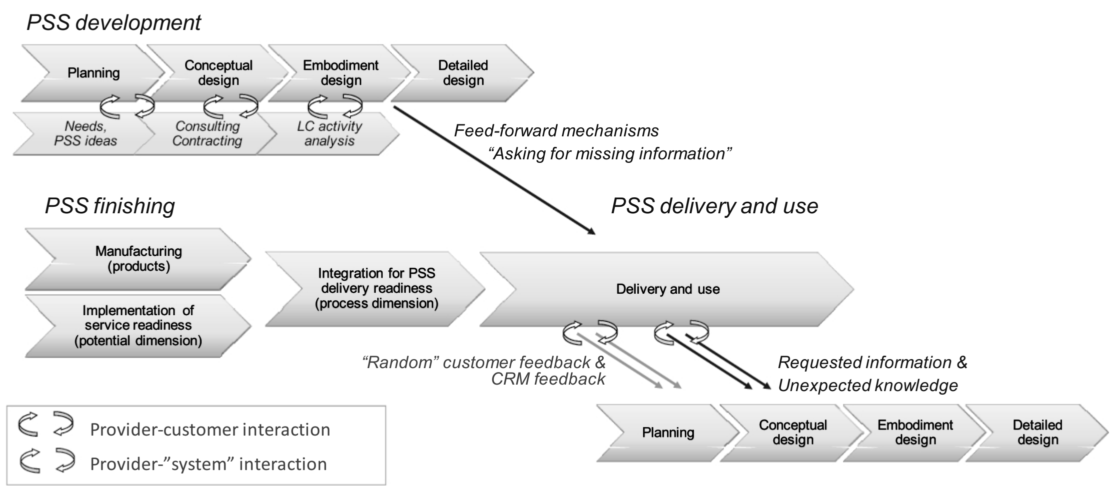


Figure 5: Stream/Cycle of PSS development, finishing, delivery and use – Classification of feedback mechanisms.

attributes and parameters [17] e.g. location, utilization, efficiency, emissions, condition, malfunctions and failures.

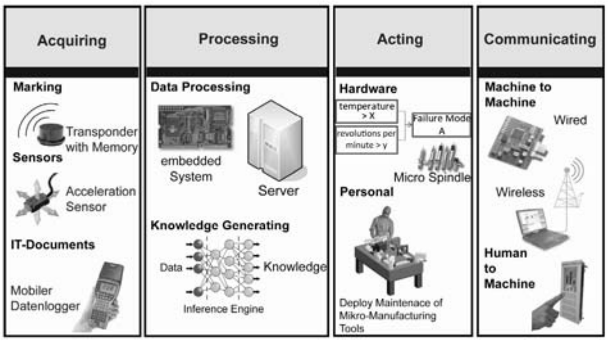


Figure 6: Modular concept of a Virtual Life Cycle Unit.

The acquired data and information will be processed with the objective to generate knowledge, e.g. inferences between product or process parameters and failures. Acquired and processed data, information and knowledge is being communicated for further evaluation or to inform the PSS management. With the help of condition prognoses the need of maintenance can be identified and automatically deployed. It is also possible to set up so called “failure modes” to prevent damage, which will be triggered if generated knowledge forecasts this on a running process. These data, information and knowledge can also be used to assist processes of the value creation chain such as development, redesign, production, recycling and disposal, and supporting processes such as logistics, quality management, controlling and sales. VLCUs improve usage-oriented business models like PSS including enhanced maintenance services.

The described architecture can be realized as a distributed network of embedded systems. More powerful and more processing power requiring algorithms can run on servers, executing the data and information evaluation [18].

In the PSS use phase the VLCU is able to acquire potential information about the usage behavior, resource needs, services and workers. This sets a base for the search of inferences, e.g. worker qualification, service efficiency or flowcharts regarding the user demands and requirements. This knowledge is invaluable for cost and resource reducing in service planning or redesign of the PSS. The goal is to increase the use-productivity of resources in PSS by finding inferences about adaptations between different usage phases.

3.2 Data acquisition in PSS

The problem of acquiring and processing data from complex machines can be solved by concentrating on standard components. Standard components like bearings, gears, compressors, pumps, dampers, filters, hose lines or pneumatic components are integrated into various more complex products, e.g. assembly systems, ground conveyors or industrial robots. By focusing on the assessment of data on standard components of machines, the development effort for the assessment of complex products is distributed technically and economically on many applications. Also, less overall expertise is needed in order to develop a VLCU system for complex products or components, because each subsystem can be examined almost independently of the others. The solution space is significantly reduced for each VLCU designer, Figure 7.

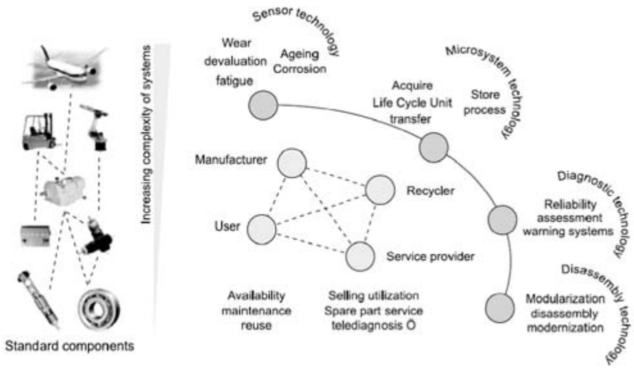


Figure 7: Products, standard components, interest groups, and business areas [19].

Maintenance processes on those standard components are a ‘door’ for data acquiring of a service. This enables a direct connection to the standards component, failure prognoses and to detect critical system parameters, inferences between services on a component and its condition, e.g. remaining life time or failure rate.

For VLCU implementations there are four major groups of services of a PSS, as shown in Figure 8. The indicators for those services are almost similar and basing on simple dimensions like time, amounts, evaluation levels and qualification classes. These information are in most instances already acquired in service documentary by the technician and the customer. Combining these information with the product condition data of the standard components, enables a failure diagnosis and prognosis with influences of the service applied to the product. Further, this facilitates to find inferences between standard components and services.

The PSS helps organization profits of knowledge of the use phase. This enables to set up the best service at the right time, location and with the optimal tools.

Figure 8: Services applicable to products in a PSS and relevant criteria for the knowledge generation.

3.3 Key steps of the automated knowledge generation

The knowledge generating procedure is divided into eight steps:

1. **Data Acquisition:** acquire data from sensors on the product and product related information, e.g. datasheets
2. **Data Cleaning:** remove noise and inconsistent data
3. **Data Selection:** select the part of the data that are relevant for the PSS
4. **Data Integration:** combine multiple data sources to information
5. **Data Transformation:** transform the data into a suitable format (e.g. a single table, by summary or aggregation operations) and store them in a database for further processing and access
6. **Data Mining:** apply machine learning and machine discovery techniques, e.g. inference engines
7. **Pattern Evaluation:** evaluate whether the found patterns or inferences meet the requirements (e.g. interestingness)
8. **Knowledge Presentation:** present the mined knowledge to the PSS provider (e.g. visualization)

The PSS provider receives new identified knowledge and is enabled for an effective and efficient redesign or reconfiguration of his PSS to changing customer needs and requirements.

To facilitate a high profit of a VLCU the designer has to take care about product and process related information. Sensors on process related or system critical components enable to generate information and knowledge. The PSS designer has specially to integrate such considerations into his design.

4 PSS (RE)DESIGN INCLUDING KNOWLEDGE GENERATION IN THE USE-PHASE

A product used in daily life has been chosen to demonstrate the capabilities of VLCU applications. The real demonstrator bases on a bicycle equipped with sensors and other VLCU components. This example is transferable to other mobile systems in the areas of transportation or construction e.g. cars, ships, cargo etc., but also to stationary systems like manufacturing system influenced by the dynamics of their environment.

4.1 Motivating examples

Especially systems running in processes with not clearly described boundary conditions (e.g. mobile systems as road construction machinery) or systems implementing new technologies (micro-manufacturing systems) have huge potential for VCLU use in a PSS business model and for PSS further development:

Mobile construction machines: The execution of numerous experiments covering all potential operation situations (e.g. machine-soil pairs with different soils all over the world) is often too expensive and a lot of findings are gained randomly in the field. A VLCU can help to build up a broad database to investigate inferences of different operation influencing parameters to design enhanced operation surveillance systems to raise the systems operation quality.

Micro-manufacturing systems: A down-scaling from macro-manufacturing is not in all cases possible. Environmental influences affect for instance the repeatability accuracy in manufacturing processes, but which factors are relevant and which are not is yet sometimes unclear. ‘Asking’ selectively for information on interfering factors which is missing in development can support designers to develop better systems (quality evolution). Especially for systems including modern technologies which are not to their full extent ‘understood’ by the designers this could be helpful to build up specific design, implementation or use guidelines.

In both cases/examples the customer finally gets product-service systems of better quality based on better controlled technologies and delivery processes, if the provider uses options like VLCUs to turn the close contact to its PSS into a design benefit.

4.2 Bicycle in a PSS business model

To provide people a bicycles-based mobility PSS is commonly offered by several companies around cities in Europe, e.g. ‘Call-A-Bike’ (Germany) [19], ‘Vélo’v’ (Lyon, France) [20] or ‘Velotaxi’ [21]. The availability of mobility is offered, the products are located around the inner cities and services are done in time intervals. However, there is no product accompanying information system installed and damage or knowledge for redesigns of the bicycles depends on comments of the customer or by technicians frequently inspect the product to check its condition.

A VLCU installed on those products can deliver valuable knowledge and information about the product condition and the usage of the bicycles, see Figure 9. With a few data interesting information can be generated, e.g. with data of acceleration and damper deflection it could be identified if the bike has been used as designed on regular streets or on dirty roads, e.g. in the wood. This knowledge can be used e.g. in a redesign approach to assemble tires for dirty roads or doing a constructive adjustment to a mountain bike. By adapting the bike to the required usage the provider benefits of more customers and a higher customer satisfaction. This example shows how a VLCU enables an enhanced design by providing information and knowledge of the use phase.

In functionality-oriented PSS the provider is interested to offer mobility as function to the customer. Breakdowns can be avoided by VLCUs condition prognoses. This enables condition based maintenance of the bicycle, which enhances the use-productivity of each component and saves resources and money by this.

In availability-oriented PSS, e.g. Call-A-Bike, the provider is able to organize his maintenance crew on the basis of the product condition and arrange reconfiguration due to usage demands. The customer benefits are a product designed on his needs, more safety by having condition prognosis of critical components integrated and more features, e.g. fun-factors like damping information on his ride.

In result-oriented PSS, the provider is interested to deliver the requested option to the customer in the right time in the right place. This requires fully functional bicycles and location information to direct the customer to the bicycle or vice versa and to the desired destination. Conditioned diagnoses and GPS location information provided with the help of the VLCU, by data from sensors on the bicycle combined with worldwide available data about traffic information, enable to provide mobility to the user with highest efficiency and effectiveness.

Figure 9: Conceptual VLCU application on a bicycle for use in a PSS business model.

5 SUMMARY AND PROSPECTS

In the PSS design phase the PSS designer has to check what the problems for the PSS use phase might be – where the critical points are – so that the VLCU is able to generate knowledge about the interesting points. The classical development model has to be enlarged to take into account the VLCU design and ways of information request and knowledge feedback. Further, the feedback loops from the use phase, delivered by the VLCUs set up a cascade of an ongoing enhanced design. To teach the designer adequate monitoring and knowledge generating tools, a convenient overview about sensors, algorithms, related information and indicators of components and processes is required and has to be composed.

6 ACKNOWLEDGEMENTS

We thank the German Research Foundation (DFG), www.dfg.de, for funding the research project SFB/Transregio 29, 'Engineering of Product-Service Systems', www.tr29.de, in which our research is embedded.

7 REFERENCES

[1] Meier, H., Uhlmann, E., Kortmann, D., 2005, Hybride Leistungsbündel, Nutzenorientiertes Produktverständnis durch interferierende Sach- und Dienstleistungen. wt Werkstatttechnik online Jahrgang 95 H. 7/8, S. 528-532.

[2] Matzen D., Tan A. R., Andreasen M. M., 2005, Product/service-systems: Proposal for models and terminology. Beiträge zum 16. Symposium Design for X, Lehrstuhl für Konstruktionstechnik, Technische Universität Erlangen, 27-38.

[3] McAloone T. C., Andreasen M. M., 2004, Design for Utility, sustainability and societal virtues: Developing Product Service Systems, Proceedings of the International Design Conference.

[4] Schuh, G., Friedel, T., Gebauer, H. 2004, Fit for Service: Industrie als Dienstleister.Carl Hanser Verlag München Wien, ISBN 3-446-22665-6.

[5] Lindahl M., Sundin E., Sakao T. and Shimomura Y., 2005, An Application of a Service Design Tool at a global Warehouse Provider, Proceedings of the 15th International Conference on Engineering Design (ICED'05), Melbourne.

[6] Tomiyama, T., 2001, Service Engineering to Intensify Service Contents in Product Life Cycles, Research in Artifacts, Center for Engineering, The University of Tokyo.

[7] IPSE project, 1/2008, http://www.ipse.se/index.html

[8] Functional Product Development, November 2007, http://www.ltu.se/tfm/fpd/about/articles/welcome?l=en

[9] SFB/TR29, 1/2008, Engineering hybrider Leistungsbündel, Dynamische Wechselwirkungen von Sach- und Dienstleistungen in der Produktion. http://www.tr29.de

[10] Pahl, G., Beitz, W., Feldhusen, J., Grote, K. H., 2007, Engineering Design, A Systematic Approach. Third Edition, Springer-Verlag London Limited, ISBN 978-1-84628-318-5.

[11] Matzen, D., Adreasen, M. M., 2006, Opportunity Parameters in the Development of Product/Service-Systems. International Design Conference (DESIGN'06), Dubrovnik

[12] Bullinger, H.-J., Scheer, A.-W. (eds.), 2006, Service Engineering, Entwicklung und Gestaltung innovativer Dienstleistungen. 2nd Edition, Springer Berlin Heidelberg, ISBN 10 3-540-25324-6.

[13] Matzen, D., McAloone, T. C., 2006, A tool for conceptualizing in PSS development. Design for X, Beiträge zum 17. Symposium, Lehrstuhl für Konstruktionstechnik, TU Erlangen, ISBN-10: 3-9808539-4-2.

[14] VDI, 1987, VDI guideline 2221, Systematic Approach to the Design of Technical Systems and Products. VDI-Verlag GmbH, Düsseldorf.

[15] Müller, P., Schmidt-Kretschmer, M., 2008, Challenges in PSS development processes – New paradigms, new development methodology. Proceedings of the 1st International Seminar on PSS, Bochum, Germany.

[16] Andreasen, M. M., Hein, L., 2000, Integrated Product Development. Institute for Product Development, Technical University of Denmark, Copenhagen.

[17] Seliger, G., Gegusch, R., Odry, D., 2/2007, Competitive Products and Processes. CIRP Competitive Manufacturing '07, Stellenbosch, South-Africa.

[18] Buchholz, A., Middendorf, A., Ray, P., Reichl, H., Seliger, G., 2004, Increased Reliability through Assessment of Standard Components with Life Cycle Units. Electronics Goes Green 2004+, Berlin, Germany.

[19] Call-A-Bike, www.callabike.de, 2008-31-01.

[20] Maussang, N., Sakao, T., Zwolinski, P., Brissaud, D., 2007, A model for designing product-service systems using functional analysis and agent based model. Proceedings of the 16th International Conference on Engineering Design (ICED'07), Paris.

[21] Velotaxi, www.velotaxi.de, 2008-31-01.

Modelling Approach for the Integrated Development of Industrial Product-Service Systems

Ewald Georg Welp[1], Horst Meier[2], Tim Sadek[1] and Katja Sadek[2]

[1] Institute of Product and Service Engineering, Chair of Engineering Design, Ruhr-University of Bochum, Germany

[2] Institute of Product and Service Engineering, Chair of Production Systems, Ruhr-University of Bochum, Germany

Abstract

Shifting business focus from offering mere physical products to offering IPS^2 an innovative strategy to meet customer requirements is revealed. The objective of this paper is to present a modelling approach for integrated development of IPS^2-concepts in early development phases. The characteristics of the IPS² concept modelling approach are deduced from a case study on a fictitious micro-manufacturing task. The proposed modelling approach abolishes established and mostly diffuse borders between products and services to assess systems behaviour already in early phases. Furthermore, applying the concept of modularisation to services in order to reduce complexity induced by specific IPS²-characteristics is discussed.

Keywords:

Product and Service Distinction; IPS² concept model; Modular Service Architecture

1 INTRODUCTION

The high degree of complexity in modern manufacturing systems and the broad variance and strong dynamics of changing customer demands forces enterprises to shift their business focus from only designing and selling physical products to offering integrated Industrial Product-Service-Systems (IPS^2) [1]. These combined solutions generate value for customers and suppliers alike during the entire delivery/use phase. Gain value on both sides systems flexibility throughout the entire IPS^2-lifecycle is an important aspect. Hence, the operational support in the shape of integrated modelling is essential to develop IPS^2 efficiently. Due to the fact that services exhibit a high degree of intangibility classic product modelling approaches can not be transferred indiscriminately to the service sector or to the integration of products and services [1].

Against this background an integrated modelling approach for IPS² is needed to enable value creation for customers and suppliers alike. In this context especially the early development phase gains major importance [2]. Moreover, a suitable methodology to cope with increasing complexity during these development phases is needed.

2 OBJECTIVES

The main objective of this paper is to highlight the relevance of IPS² concept modelling as well as its potentials along the entire IPS²-lifecycle. To do so, chapter 3 links different business models to IPS² using a case study. Additionally, an IPS² basic structure is deduced and specified. Based on this a change of product and service perception which is necessary for IPS² concept modelling is discussed in chapter 4. Chapter 5 focuses on IPS² concept modelling and introduces the IPS² concept modelling approach. Possibilities of modularisation under IPS² circumstances are discussed in chapter 6. The paper concludes in chapter 7.

3 COMPREHENDING IPS²

3.1 IPS² case study focussing business models

The following case study targets at a customer demanding the manufacturing of rotationally symmetric µm-parts using wire-cut EDM while simultaneously achieving maximum manufacturing productivity. Consequently, there are two problem aspects to consider: i) the technological aspect and ii) the aspect of manufacturing productivity.

Technologically the manufacturing process on a wire-cut EDM manufacturing facility is enabled by integrating a rotating spindle [3]. The supplier's solution comprises the rotating spindle and 'attached' services and does not include the entire wire-cut EDM manufacturing facility. Furthermore, a minimum level of qualification (qual=XY) is required to enable the operating personnel to manufacture µm-parts. This implies two alternatives of operation either by the supplier's personnel or by customer's personnel. In the following the technological problem is subordinated to the productivity oriented problem where manufacturing productivity is defined by units per time.

According to the classification of PSS proposed by [4], three different business models - **product oriented, use oriented and result oriented** - can be distinguished. Alternatives like the mere selling of the rotating spindle (‚pure' physical product oriented) or the mere selling of µm-parts (‚pure' service oriented) are excluded from the aforementioned business models.

Applied to this demand a **product oriented business model** comprises purchasing the rotating spindle by the customer. The production facility including the rotating spindle is considered to be owned by the customer and is operated by the customer's own personnel. An improvement of manufacturing productivity can only be achieved by increasing output quantity, because the time of utilisation of the manufacturing facility including the rotating spindle is assumed to be constant. If the customer's own output quantity is limited and the capacity of the production facility exceeds this quantity, procuring of external production orders will be necessary. In this case the IPS²-supplier offers the service of procuring external production orders and for instance additional services like production planning and scheduling.

Résumé: A product oriented business model requires few modifications of the physical product due to weak interdependencies between the product and added service.

The comprehensive integration of products and services is reflected in another business model which prompts the customer to refrain from purchasing the rotating spindle. Diametrically opposed to the product oriented business model customers choose a demand-oriented temporary utilisation of the manufacturing equipment. In this case achieving an increase of manufacturing productivity depends on decreasing the time of utilisation of the rotating spindle. Thus, the supplier retains ownership of the rotating spindle, enabling him/her to make it available to different customers. According to the PPS classification given by [4] this solution can be characterized as a combination of a **use and result oriented business model**. Due to the demand-oriented provision of the rotating spindle technical as well as organisational differences between customers are likely and have to be considered already in early phases of IPS²-development. Consequently, different levels of operator qualification as well as differences between wire-cut EDM production facilities have to be compensated for either by product or by service artefacts. The success of this business model is heavily dependent on the degree of flexibility of the entire product-service bundle. The definition of the degree of flexibility, supplier's guarantee of the rotating spindle's integrity at system delivery time and other legal issues form the main part of the contract between customer and supplier.

Résumé: The increasing complexity of use and result oriented business models requires the integration of products and services already in early phases of IPS²-development.

While the aforementioned business models consider **'what'** needs to be fulfilled to meet customer requirements the question of **'how'** this will be done is answered by a model of the IPS² itself [5] (see Figure 1).

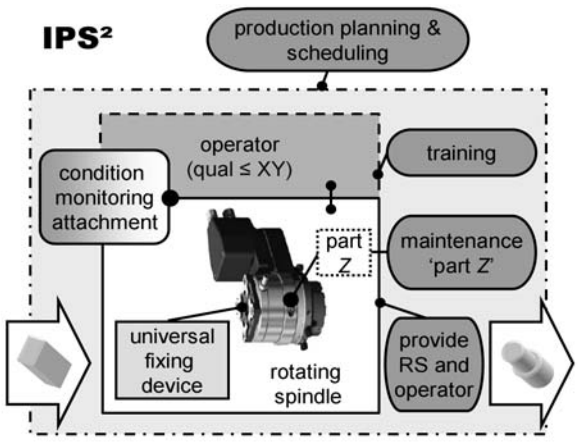


Figure 1: Product and service artefacts of the 'IPS²'-solution.

To meet customer requirements of "productivity dominated manufacturing of rotationally symmetric µm-parts" an adaptive mechatronic system including integrated services is developed. As illustrated in Figure 1, the *rotating spindle* is the core of the technological system. It contains a *part z* as a representative of the variety of different parts which require maintenance after manufacturing a certain output quantity. To enable the integration of the rotating spindle into different customers' production facilities it is equipped with a *universal fixing device*. Linked to the physical product, the supplier also provides the possibility of employing his/her operator depending on the level of customers' operator qualification and further surrounding conditions.

Apart from *maintenance* services the temporary availability of the manufacturing system induces services like *provision* of technical and human artefacts as well as *production planning* and *production scheduling*. Additionally, *training* services are also part of the IPS² to enable the customer to improve operator qualification during the delivery/use phase.

To monitor the condition of technical artefacts and to observe the way of operating the production facility during delivery and use a *condition monitoring attachment* is applied. As seen in Figure 1 the condition monitoring attachment is figuratively placed between physical product, human operator and service. As subsequently shown this fact represents the inconsistency of definitions of products and services [6], [7].

3.2 Deducing an IPS² basic structure

To deduce universally valid findings from the presented case study an IPS² basic structure is defined (see Figure 2).

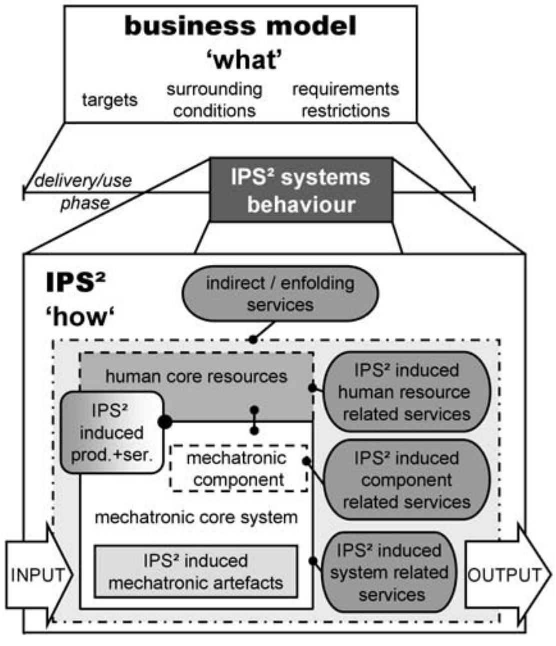


Figure 2: IPS² basic structure.

As shown in Figure 2 the entire delivery/use phase of an IPS² is taken into account in the IPS² basic structure. Applicability to each of the aforementioned business models is ensured. Depending on individual types and unique characteristics of business models certain elements within the IPS² itself are detailed, emphasized, or neglected.

As the superordinate artefact of the IPS² basic structure the **business model** defines and covers the entire delivery/use phase. It therefore contains abstract constructs like customers' and suppliers' targets as well as a representation of possible contradictions between their preferences. Surrounding conditions and their respective changes as well as requirements and restrictions are also inherent parts of the business model. Summarizing, the business model defines, in accordance with [4] the architecture for products, services and information flows. Furthermore, descriptions of the various business actors and their roles are included into it in order to define their potential benefits.

As already implicitly prescribed in the business model an IPS² possesses characteristics and properties which serve to meet customers' as well as suppliers' demands. Flexibility as mentioned in the IPS² case study, constitutes such a characteristic. It is needed to make a product available to different customers. The IPS²'s target-oriented potential of reacting to internal and external changes is primarily enabled by a *partial substitution* of product and service artefacts [8]. By exchanging product components with service artefacts and vice versa the supplier is able to adjust the IPS² to changing surrounding conditions during delivery and use. Versatility granted by this partial substitution is mirrored in the **IPS² systems behaviour**. Thus, the IPS² systems behaviour expresses the link between the business model (answer to 'what') and the IPS² itself (answer to 'how'). The IPS² systems behaviour enables to assess the interplay of product and service artefacts as well as the reaction to anticipated changes of surrounding conditions during the entire delivery and use phase.

The **IPS² itself** includes a technical transformation process to change inputs into outputs. Following the understanding of IPS² as industrial business to business solutions, the technical basis is composed of a mechatronic core system and an associated human core resource. The mechatronic core system includes its mechatronic components and supplementary IPS²-induced artefacts. In this context the phrase *'IPS²-induced'* is related to extraordinary product or service artefacts which are not necessarily needed to fulfil basic functions of the product oriented business model but gain relevance in connection with use and result oriented business models. Furthermore, the basic IPS² contains services related to single mechatronic components or to the entire mechatronic system. Services which solely connected to human resources are also defined. A strict distinction between basic and IPS²-induced services is not necessary due to fuzzy distinctions between obligatory and additional functions. Around the core system and the integrated services indirect and enfolding services exist which are provided almost independently from the core system. This approach of allocating services is comparable to the line of order penetration described by [9].

Integrated 'product+service' components, such as a condition monitoring attachment or augmented reality tools, are purely IPS²-induced and require special attention as discussed in the following chapter.

4 CHANGING PRODUCT AND SERVICE PERCEPTION

4.1 Distinguishing products and services

According to [6], every product involves services such as sales, delivery and support. Moreover, [6] also points out that every service involves physical product artefacts in order to generate benefit. Against this background it is difficult to provide a clear distinction between 'service' and 'product'. Integrated 'Product+service' components represent the weakness of established definitions while these components can not be explicitly classified as being part of one or the other discipline.

Based on the IPS² basic structure the question arises if it is still necessary to distinguish between products and services to describe and assess IPS² systems behaviour during early phases of IPS²-development. Against the background of integrating disciplines there might be no need to stick to this distinction.

According to the scenario's IPS²-induced 'product+service' component, condition monitoring is identified as a significant systems artefact which contains neither mainly product nor service related properties and characteristics. As illustrated in Figure 3 the objective of assessing systems behaviour during IPS² concept modelling is initiated by defining an IPS²-function, in this case *'monitor system's condition'*.

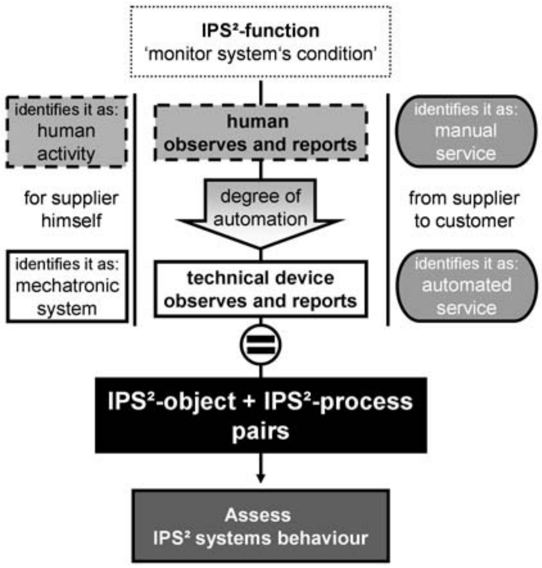


Figure 3: Supplier and customer perception of products and services.

Depending on the level of automation a pure manual solution (human observes and reports) or a pure automated solution (technical device observes and reports) correspond to the bottom respectively to the top of this scale. The application of the supplier's and the customer's perception onto these solutions can be taken from Figure 3 as well. Depending on the addressee either a supplier, who carries out the IPS²-function himself, or a customer to whom a supplier offers the identical IPS²-function, will perceive similar solutions differently. On the one hand the supplier identifies them either as a human activity or as part of the mechatronic system depending on the degree of automation. Thus, the supplier perceives the solutions as products. On the other hand a customer perceives a human or a technical device which observes and reports the system's condition. The customer identifies these as a manual service or as an automated service respectively. Thus, a customer perceives the exact same solutions not as part of products but as services.

Résumé: Products and services mainly differ depending on the observer's point of view but not in their inherent properties. As an answer to the aforementioned question it can be stated that it is not necessary to distinguish between products and services regarding modelling and assessing IPS² systems behaviour. Their related properties and characteristics do not change just by switching the addressee's perception. Furthermore, to constitute partially substitutable products and services within an integrated conceptual model the reorientation of product and service perception is necessary.

4.2 Introducing IPS²-object + IPS²-process pairs

Viewing products and services from a different angle, it becomes obvious that in order to fulfil IPS²-functions, two types of artefacts are needed in general. As illustrated in Figure 3, **IPS²-objects**, for instance a human resource or a technical device, constitute a physical structure. Furthermore, **IPS²-processes**, for example an algorithm for observing and reporting, need to be determined in order to reach a desired system state.

To model IPS² systems behaviour processes have to be applied to objects (see Figure 3)**.** Only the combination of both artefacts can yield the desired problem solution and the functional behaviour of the system. The introduction of this new paradigm renders the separation of product and service modelling irrelevant for the engineering of IPS²-concepts in early phases of development. By defining **IPS²-object + IPS²-process pairs**, products and services can be developed in an integrated concept model. The connection of multiple IPS²-object + IPS²-process pairs represents the IPS² systems behaviour along the entire delivery/use phase. Regarding the IPS² basic system this corresponds to the replacement of product and service artefacts by IPS²-object + IPS²-process pairs. A deliberate distinction during the development is mirrored by emphasizing either state-oriented IPS²-objects (physical product related) or activity-oriented IPS²-processes (service related) within these pairs.

The terminology used to regard state-oriented and activity-oriented artefacts is deliberately chosen to reflect concepts, which have been ascribed with a variety of different meanings in an interdisciplinary context. This serves to underline the diversity of the element pairs. In an IPS²-context the term 'object' which is also used in the context of the object-oriented paradigm of IT-technology [10] does not refer to the instantiation of a class. As expressed by the prefix 'IPS²-', IPS²-objects represent entities of technical and human elements. The states of these entities can be generated, destroyed or transformed by applying a process [10]. The term 'process' in combination with the prefix 'IPS²-' can therefore be considered as the rule for transformation. Such a transformation rule can be applied to technical as well as human elements, components, or systems when detailing an IPS². Distinguishing between logical and physical processes is also a sensible approach when modelling IPS² systems behaviour [Quelle Jansen].

5 INTEGRATED IPS² CONCEPT MODELLING

Following the general understanding of the term 'model' [11], the IPS² concept model denotes a consciously constructed aim-oriented representation of an original. To describe elements of the IPS² basic structure and their respective interrelations on a medium level of abstraction, a concept model serves the purpose of enabling both analysing and synthesising steps. Therefore, the main objective of a concept model can be summarized as the primary transformation of a cognitive construct into a real physical product respectively service solution. The term "early phases of product development" for IPS² is used synonymously to the understanding of this term established in product engineering.

However, regarding IPS² the traditional understanding of integrated development needs to be expanded to take IPS²-induced flexibility and aligned potentials of partial substitution into account. Consequently, interdependencies between products and services which are implicitly defined in the business model need to be concretised in the conceptual phase in order to determine the basis for IPS²-flexibility. Furthermore, this development phase serves the purpose of initiating potentials of IPS²-object + IPS²-process pairs for the partial substitution along the entire delivery/use phase. Therefore, the early phase of development starts with determining customer requirements and ends with the combination of IPS²- object + IPS²-process pairs to the IPS² systems behaviour. Thereby the IPS² itself is connected to the superordinate business model with its targets, surrounding conditions, requirements and restrictions. This marks the initial model to assess technical as well as economic system states and their transition to IPS² systems behaviour along the entire delivery/use phase.

Hence, IPS² concept modelling needs to be supported by an appropriate modelling methodology which considers superordinate aspects such as iterative approaches and a high degree of flexibility in the process of development.

Building on dissolving distinctions between products and services, the IPS² concept model is generated by defining and merging four modelling layers [12] (see Figure 4).

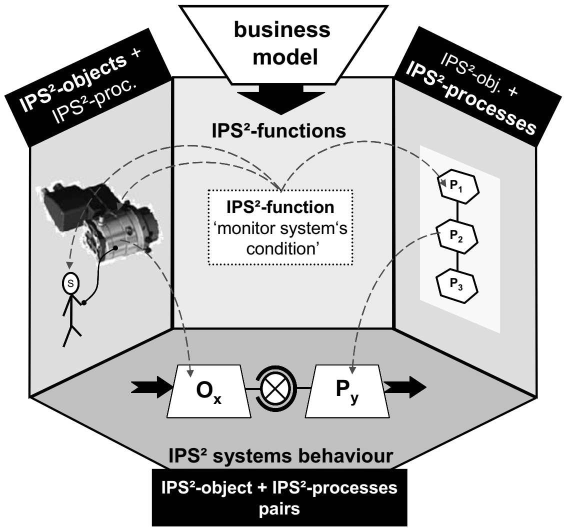


Figure 4: IPS² concept model.

As shown in Figure 4 the IPS² concept model is build by combining the IPS²-functions layer, the IPS²-object dominated layer and the IPS²-process dominated layer. The integration of IPS²-objects and IPS²-processes into pairs in order to formalise IPS² systems behaviour is done on the homonymous layer. The content of the business model is not modelled on a separate layer but is included into the IPS² concept model via context elements [12]. These superordinate elements are used independently from the level of abstraction and can be used on each modelling layer.

Accordingly, modelling IPS²-functions, as a development activity, describes desired IPS²'s characteristics on a high level of abstraction. An IPS²-function equals the uncommitted, abstract representation of intended changes within the IPS². Considering the entire delivery/use phase the interconnection of IPS²-functions through relationships is the initial step towards formalizing IPS² systems behaviour. Deriving IPS² systems behaviour from IPS²-functions directly is not possible although generally desired.

Starting from the change in paradigm regarding the dissolution of the distinction between products and services, concretising IPS²-functions refers to the IPS²-object dominated layer as well as to the IPS²-process dominated layer. In an IPS² concept model, aspired IPS²-functions are partially defined by forming and concretizing IPS²-object + IPS²-process pairs. An IPS²-object hereby serves as a physical, material or human point of reference, as previously defined. Complementary to this, an IPS²-process marks a structural connection of an initial state and a final state by following processing instructions. By applying an IPS²-process to a physical, intangible, or human object, the incremental functional behaviour of an IPS²-object + IPS²-process pair is modelled.

The linkage between IPS²-objects and IPS²-processes as well as the interconnection of IPS²-object + IPS²-process pairs takes place at the IPS² systems behaviour layer. Integrating the interdisciplinary layers to generate a model suited for simulation in the early phase of IPS²-development requires a consistent, universally valid, reference value which needs to be determined.

Résumé: The temporal fragmentation of products and services into IPS²-objects and IPS²-processes during the early phase of development constitutes more than just renaming established nomenclature. To meet the aim of deducing IPS² systems behaviour from the required IPS²-functions it is crucial to overcome information deficits in the process of classic product development and to introduce dynamic components.

6 MODULAR DESIGN

6.1 Linking IPS² concept modelling with product and service design

To enable the integrated development of product and service artefacts IPS^2 concept modelling needs to be connected to embodiment and design of products and services.

As exemplified in the case study a multitude of possible combinations of human and technical resources as well as applied processes is needed to deliver and use IPS^2 efficiently. This multitude of possibilities entails a high degree of complexity due to the required flexibility. This complexity needs to be dealt with by using suitable strategies to manage IPS²-variants.

Modularisation of products and services constitutes such a strategy which is already applied to reduce complexity in product engineering. However, the role of modularisation and modular design regarding the reduction of the degree of complexity in an IPS²-context needs to be investigated. Modular product architectures and their transferability to the service sector are to be discussed.

6.2 Discussing a modular service architecture

Product modularity arises from the division of an artefact into independent components and allows companies to create product variety by decoupling, reusing, exchanging or extending standardized elements. The modularisation of products reduces the development efforts, time and costs by using already existing modules. Product failures can be easily and quickly eliminated by simply replacing the defect component by an equivalent product module. In this context modular product architectures are defined as a scheme in which each combination of physical artefacts implements a specific set of functional elements. The interactions between these combinations must be precisely defined and are of fundamental importance to the primary functions of the product. The more a product component differentiates in its physical and functional use from other components the higher is the attainable level of modularity within the product architecture [13].

The initial step of evaluating the transferability of the concept of modular product architectures to the service sector consists in discussing the following questions:

- What are the fundamental characteristics of modular service architectures (MSA) and to what extent can modular product architectures serve as a prototype for service architectures?
- How can service architectures developed for service design be transferred to the IPS² concept modelling approach in order to derive IPS²-process dominated pairs?

The above mentioned transferability of modular service architectures is heavily influenced by the characteristics of services. Services are defined as a heterogeneous, mainly immaterial and perishable activity or process. They are offered by a company or an institute and are consumed simultaneously to their production [14]. An additional distinguishing characteristic of services is the necessity of integrating an external factor, provided by the service recipient, into service provision. Such external factors can consist of persons, goods, rights and information, etc.. Furthermore, transferability is also influenced by the principles of modularisation. Within the scope of software engineering [10] the term module is defined as a logical and physical unit performing a precisely defined function within an overall context. These modules are linked with each other through standardized in- and output-interfaces.

Modular systems are subject to the rules of decomposition, abstraction, nested hierarchy, loose coupling and limited access [13]. The term decomposition defines the segmentation of a system into its subsystems. With the aid of nested hierarchy these subsystems can be subdivided into further part systems to the point of elemental functions or to the limits of knowledge. Regarding modularisation the principles of loose coupling and limited access play a decisive role concerning the design of in- and output-interfaces. Loose coupling describes the request for strong dependencies between each element within one module. The principle of limited access requires non-internal relations between modules defined with the aid of public attributes. In this regard it is essential that the more the description of the module interface differs from its internals and the less those internals are revealed by in- and output information the merrier modules can be used and exchanged independent of a particular context [15].

Based on the previous descriptions of services and the principles of modularisation modular service architectures are understood as overall service systems that are built up on the aforementioned principles. They consist of unitized service modules which are linked to each other in compliance with the rules of loose coupling and limited access. Due to the process-oriented notation of the modular service architecture, services are visualized under consideration of their intangibility as well as their integration of external factors.

Therefore, modular service architectures are used in order to shorten service design times and costs by the aid of the aforementioned modularisation mechanisms. In this regard the ability to respond to changes at the point of sale increases. Moreover, the recombination of service modules allows customized service configurations which require minimum design efforts.

Thus, it becomes apparent that the approach of modular product architectures is transferable to the service sector. Apart from IPS² concept modelling where product and service just depend on the addressee's perspective, special service characteristics have to be taken into account in modular IPS² service design. Especially with regard to the partial substitution in an IPS^2-context the use of such modules can be of particular help. Following this concept, the advantages of modularisation such as managing complexity as well as decreasing development costs and time can be realized during the IPS^2 design process.

7 CONCLUSION

According to IPS² which focus on discipline's integration a fictitious case study linked to business models is presented to show that deducing an IPS² basic structure is crucial to enfold IPS²-characteristics for concept modelling. This structure basically consists of three integral parts.

The discussion on product and service perception shows that no distinction between elements of both disciplines is needed in the context of IPS² concept modelling. With the introduction of IPS²-object + IPS²-process pairs a comprehensible way to integrate product and service artefacts within an integrated IPS² concept model is presented. Integrated product and service modelling during the early phase of IPS²-development is essential to deploy abilities of IPS² as well as to assess IPS² systems behaviour. As part of linking IPS² concept modelling with product and service embodiment and design as well as to reduce complexity of the system modularisation within IPS² is discussed.

8 ACKNOWLEDGMENTS

Financial support from the German Science Foundation (DFG) through SFB/TR29 is gratefully acknowledged.

9 REFERENCES

[1] Meier, H., Uhlmann, E., Kortmann, D., 2005, Hybride Leistungsbündel, Nutzenorientiertes Produktverständnis durch interferierende Sach- und Dienstleistungen, wt Werkstattstechnik online, Springer VDI Verlag.

[2] Welp, E.G., Sadek, T., Müller, P., Blessing, L., 2008, Integrated modelling of products and services – The conceptual design phase in an integrated IPS² development process, Proceedings of CIRP Design Synthesis, April 7 – 8. Delft, Netherlands, April 2008. *(under review)*

[3] Uhlmann, E., Oberschmidt, D., Geisert, C., Stelzer, C., 2007, Interrelationship among rotation speed and process result in Wire Electrical Discharge Grinding (WEDG) as wear indicator. In: Proceedings of the 2nd International Conference on Micromanufacturing - ICOMM 2007, 10-13 September 2007, Greenville, South Carolina, USA, pp. 40-45.

[4] Tucker A., Tischner U., 2006, New Business for Old Europe – Product-Service Development, Competiveness and Sustainability, Greenleaf Publishing, Sheffiled.

[5] Soderborg, N.R., Crawley, E.C., Dori, D., 2003, OPM-Based Definitions and Operational Templates, in: Special Issue on conceptual modelling and System Architecture, Communications of the ACM, pp. 68 ff.

[6] Tan A.R., McAloone T.C., Andreasen M. M., 2006, What happens to integrated product development models with product/service-system approaches?", 6th Integrated Product Development Workshop IPD, Magdeburg.

[7] Mont O., 2004, "Product-service systems: Panacea or myth?", Doctoral Dissertation, Lund University, Lund.

[8] Karger, M., Richter, A., Sadek, T., Strotmann, W., 2008, Flexibility of Industrial Product Service Systems – An Assessment Based on Concept Modelling, Proceedings of 3rd International Conference on Business Market Management, St. Gallen, Switzerland. *(to be published)*

[9] Fließ, S., Kleinaltenkamp, M., 2004, Blueprinting the service company. Managing service processes efficiently, Journal of Business research 57,H. 4, pp. 392-404.

[10] Rumbaugh J., et al, 1990, Objektoriented Modelling and Design", Prentice Hall.

[11] Jansen, S., Welp, E.G., 2005, "Model-Based Design of Actuation Concepts: A Support for Domain-Allocation in Mechatronics", Proceedings of the 15th International Conference on Engineering Design (ICED'05), Melbourne.

[12] Welp E.G., Sadek T., 2008, "Conceptual Design of Industrial Product-Service Systems (IPS²) based on the Extended Heterogeneous Modelling Approach", Proceedings of International Design Conference (DESIGN 2008), May 19 - 22 Dubrovnik, Croatia. (under review)

[13] Erixon, G., 1998, Modular Function Deployment –A Method for Product Modularisation, pp. 50-57.

[14] Matzen, D., Tan, A. R., Myrup Andreasen, M., 2005, Product/Service-Systems - Proposal for Models and Terminology, 16. Symposium "Design for X", Neukirchen,.

[15] Böhmann, T.; Krcmar, H., 2003, Modulare Servicearchitektur, in: Bullinger, H.-J., Scheer, A.-W., Service-Engineering – Gestaltung und Entwicklung innovativer Dienstleistungen, Springer-Verlag, Berlin.

Integrating function model and activity model for design of service

Tatsunori Hara[1], Tamio Arai[1] and Yoshiki Shimomura[2]

[1] Department of Precision Engineering, School of Engineering, The University of Tokyo, Japan

[2] Department of System Design, School of System Design, Tokyo Metropolitan University, Japan

Abstract

Manufacturing industry needs to supply services more than materialized products to customers recently. Early literature on service marketing lacks the perspective on functional meaning of service activities to the customer. This paper focuses on presenting a framework so that we comprehensively deal with customer value, function, and activity. In this paper, we describe service blueprint of activities by BPMN. The above method is implemented on a CAD system, and the followings are found through an example: separate function models are integrated through related activities; and interactions between the customer and the provider are incorporated.

Keywords:

Service/Product Engineering; Service Blueprint; CAD

1 INTRODUCTION

Manufacturers are required to supply customers with services beyond simple material products. In this context, the engineering target that needs to be analyzed and designed is shifting from simple products to Product-Service Systems (PSS) [1]. We have been researching Service/Product Engineering (SPE) to develop PSS since 2002 [2] [3].

In marketing field, customer value and human activity are main topics in service system. However, literatures of this research lack the perspective on effect and meaning of activities to the customer. In traditional engineering field, effect and meaning of artifacts can be commonly represented by function. Hence, we have mainly modeled service contents as a set of function to change states of the receiver. This paper focuses on integrating the function model and the activity model so as to evaluate the service according to the customer's point of view.

The rest of this paper consists of the following: Section 2 defines service and describes problems in service design. Section 3 introduces early research on service activity and gives notation of service activity. Section 4 illustrates a method of integrating function model and activity model. Then, Service Explorer, which is our computer-aided design tool, is applied to an example case in Section 5. Section 6 discusses and Section 7 concludes the paper.

2 MOTIVATION

2.1 Service to be targeted in this study

We define service as a contract between a service provider and a service receiver to change the state of the receiver [2] [3]. This definition is broader than typical definitions in the traditional management and marketing fields (e.g. [4] [5] [6] [7]). Our definition of service is quite similar to Hill's definition [8]: A change in condition or state of an economic entity (or thing) caused by another. Emerging service research (e.g. [9] [10]) also starts with discussion regarding this broader perspective. According to our definition, most business activities are services, including selling physical products. Services to be targeted in this study correspond to PSS (designed to change the state of the receiver), while a pure service (that is comprised only of human activity) is called a service activity.

2.2 Problem in service design

Figure 1 shows elements of service to be analyzed and designed: customer value, function of service, and service activity. We point out two problems in service design from the markering and engineering point of view :

The gap between customer analysis and service activity design (marketing point of view)

In marketing field, customer analysis and service activity design are the focus of the design process. While there is considerable research on service process development in the reference literature [7] [11] [12], service activities are often directly related customer needs with little discussion of the effect and meaning of such activities (Figure 1).

The separation of function design and service activity design (engineering point of view)

We represent service contents as a set of function and entity to change states of the receiver [2] [3]. Influences by physical product and/or human are aggregated in the functional form. Although service greatly involves human elements and service activity is a means of actualizing function, the relationship between function and service activity is not apparent in the early literature (Figure 1).

The objective of this study is to solve these two problems by presenting a framework so that we comprehensively deal with customer value, function, and activity. We develop a method to correlate function model from engineering field and service activity model from marketing field. In the following section, we explain a graphical representation of service activity in marketing field and specify its notation.

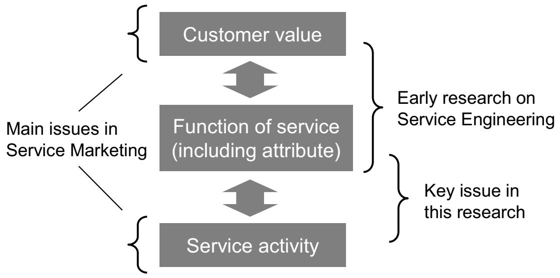


Figure 1 : Key issue in this research: filling the gap between function and activity of service

3 REPRESENTATION OF SERVICE ACTIVITY

3.1 Service Blueprint

The service blueprint [13] [14] and the service map [15] are the most famous techniques used by marketers to describe service activities sequentially and visually. In the service blueprint and the service map, service activities are arranged according to two lines: the line of interaction around which the customer and the service provider interact; and the line of visibility that separates 'onstage' (visible) from 'backstage' (invisible) activities performed by the provider.

The service blueprint is an effective technique for analyzing and designing the delivery of services on paper prior to actual delivery, but some of the literature [16] [17] points out that the blueprint of a service is likely to be an operating manual for the service, not a way of depicting the customer perspective. Service mapping is a refined, but still task-oriented version of the blueprint. Since managerial discretion is mostly over the service process, the design of this process is usually from a manager's perspective and is task-oriented [17] [18]. The service blueprint also only weakly connects customer value and service activity, as was pointed out in Section 2.2.

In this paper, we follow the concept of Shostack's blueprint, but improve the notation and superimpose functional meaning onto the service blueprint.

3.2 Service blueprint by BPMN

We use BPMN (Business Process Modeling Notation) [19] [20] for the blueprint (as shown in Figure 2) so that service activities become design objects in the blueprint. This is because Shostack's blueprint notation is not consistent in the early literature. The blueprint specifies the service delivery process and interactions between the customer and the provider that are represented by BPMN message flow. A group of BPMN message flows between the receiver and the provider and related activities represents a service encounter [11] [21].

In the following section, we discuss the relationship between function and service activity. Then we present a method to integrate function-attribute model through service activity according to the relationship.

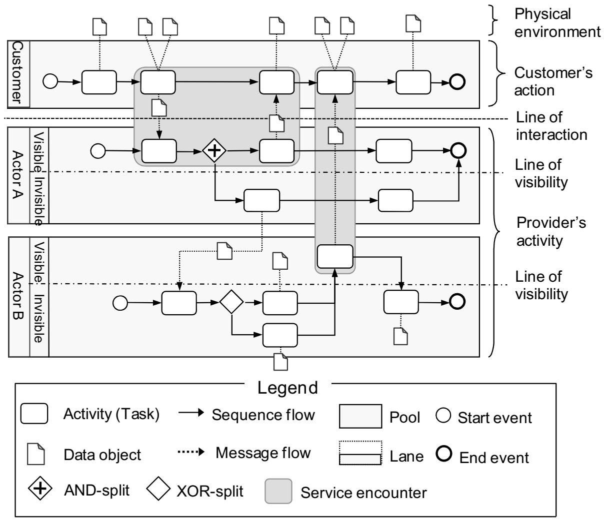


Figure 2 : Service Blueprint described by Business Process Modeling Notation (BPMN)

4 INTEGRATING FUNCTION-ATTRIBUTE MODEL THROUGH SERVICE ACTIVITY

4.1 Relationship between function and activity

The function is defined as a purpose in the mind of humans and can be realized by the system through the provision of certain behaviors by the structure. The behavior is the causal relationship or structure among a set of related states [22]. In this paper, we define a set of service activities as a human behavior and a function can be performed by the set.

4.2 Method to integrate function-attribute model through service activity

Figure 3 shows a framework to integrate function-attribute model and service activity model according to the relationship between function and activity in Section 4.1. This framework consists of four layers: receiver state layer (Figure 3 (a)), function layer (Figure 3 (b)), activity layer (Figure 3 (c)), and attribute layer (Figure 3 (d)). Using this framework, functions are deployed into activities on the activity layer and are integrated through related activities. Moreover, activities on the activity layer are reviewed from the functional viewpoint.

Step 1) Determining primitive service

First of all, primitive customer actions in receiving a service are arranged on the activity layer. Scenario using persona [2] and service script in marketing field [11] are useful to obtain these customer actions. Concurrently, state parameters of the customer are enumerated on the receiver state layer according to the scenario and a set of RSPs (Receiver State Parameters) [2] [3] are identified from them. Through this step, customer value that the service targets is determined.

Step 2) Constructing functional structures of service

After identifying the customer value as RSPs in the previous step, functions of service that affect each RSP are described in tree structure [2] [3] on the function layer (Figure 3 (b)). By detailing upper functions and constructing a functional structure, customer value (represented through RSP) is related to concrete functions on the function layer. Therefore, designers can perform a static evaluation of customer satisfaction based on performance of functions. However, the function structure includes little information with regards to the service delivery process. Thus, the ways in which the lowest functions are completed are not apparent. In the next step, separate function structures are integrated through the activity layer from the following viewpoints: the service delivery process, the relationship with customer, and consistencies with other functions.

Step 3) Constructing service delivery process

In this step, service activities are described according to the service blueprint as introduced in Section 3.1. First, the basic steps performed by the customer on the activity layer are partially chosen from the scenario presented in the step (1). Then, based on several functional structures constructed in step (2), actors of service activities are specified and arranged as BPMN pools on the activity layer. Mechanical/information system that involves collaboration with customer also allocated as actors. Then, the customer actions are modified and service encounters are constructed in consideration with functions on the function layer. The lowest functions of them are also deployed into series of activities on the activity layer (Figure 3(c-2)). Such activities include not just the provider activities but the customer actions.

After laying out activities, all activities on the service blueprint are organized to ensure the totality of the delivery process. Examples of operation to organize activities on the service blueprint are as follows: classifying into visible/invisible activity; adding/deleting activity, aggregating similar activities, and dividing into plural activities. Activity on the service blueprint can not be always related to functions on the function layer (e.g. support process) (Figure 3(c-3)). Such activity is understood as activity that does not affect the state of the customer, while it involves executing the service. Finally, sequence flow, message flow, and other BPMN flow figures are added. Such figures include branch on condition, exception handling, and so on.

This framework contributes to not just service delivery process are organized, but separate function structures for RSPs (Figure 3(b)) are integrated through the activity layer.

Step 4) Describing entities that comprise the service

Entities such as resource, processing object, and product of activities are arranged on the attribute layer (Figure 3(d)): not only entities from functional viewpoint but also entities from the viewpoint of physical environment surround customer. Entities also have their own attribute parameters. Hence, as with visibility of activity, entities on attribute layer can also have visibility to customer (Figure 3(d-2)).

5 DEMONSTRATION

5.1 Implementation on Service CAD system

We extended the functionality of Service Explorer [2] [3] [23], which is a computer-aided design system, based on the method in Section 4. The latest Service Explorer [23] has the capability of representing service activities and correlating

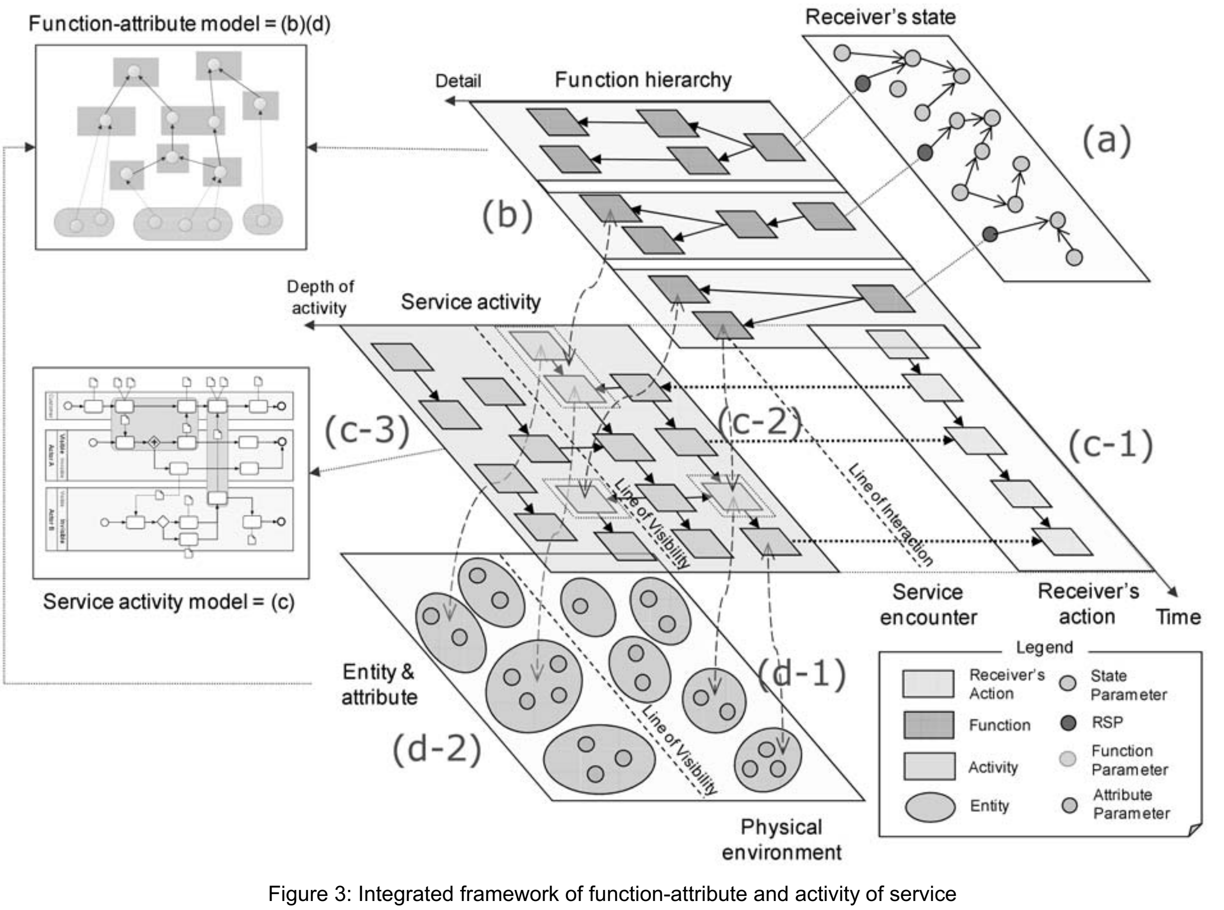


Figure 3: Integrated framework of function-attribute and activity of service

them to functions. BPMN Modeler [24], provided by the Eclipse Project [25], is adopted for Service Explorer as the activity editor. Using Service Explorer, managers, marketers and engineers can work together to improve services while designing service. Service Explorer is capable of both reviewing a service and designing a new service, can help visualize services, can evaluate services and can simulate services.

5.2 Elevator operating service

Let us apply the aforementioned method and tool to an elevator operating service that transports its customers. Stakeholders of the service are as follows: elevator user, elevator company, and security company. The elevator user is the end receiver of the service and the final customer.

First, we specified the basic actions of the elevator user when going to a shop using an elevator and identified RSPs: the security and safety of the service; time and effort to move; availability of the service; and comfortable environment. Second, four view models corresponding to these RSPs were described. Then, service activities were described in service blueprint. Screenshots of Service Explorer are explained in followings.

Representation of service activities

Figure 4 depicts partial service blueprint of the service. The blueprint involves four actors: a elevator user, an elevator, an elevator company, and a security company (not depicted in Figure 4). The elevator is arranged as BPMN pool because it interacts with the user frequently. The blueprint in Figure 4 mainly shows the user actions of using elevator and service activities that occur when an earthquake or fire is detected. These activities are processes for providing functions toward the RSP concerning security and safety of the service. The portions describing the activities of the elevator and the elevator company are split up into tow sections. Activities in one section are visible to the user, while activities in the other are invisible.

These service activities are performed from begining to end via interactions between actors (Figure 4(a)(b)). In an emergency, these activities start when the elevator experiences an emergency stop. The user calls on the intercom and the elevator company receives the call. The staff of elevator company watches the user while checking images from the monitoring camera in the elevator. Then, the staff of the security company dash to the elevator and cope with the emergency after being contacted by the elevator company.

As shown in Figure 4(c), data objects connected to the user actions comprise physical environment surround the user, while data objects connected to message flows represent contents of message among actors.

Cross-functionality of the activity editor and the view editors

Figure 5 depicts collaboration of function-attribute model and activity model on Service Explorer. Figure 5 (a)(b)(c)(d) represent view models corresponding to four RSPs. The icon on the lower right of activity node means the activity is related to function(s) in view model. For instance, the activity of emergency stop (Figure 5(e)) is related to the activated functions (responding to the emergency qulickly and certainly) and the RSP ‘security and safety of service’ as shown in Figure 5(f). This information appears in tooltip as shown in

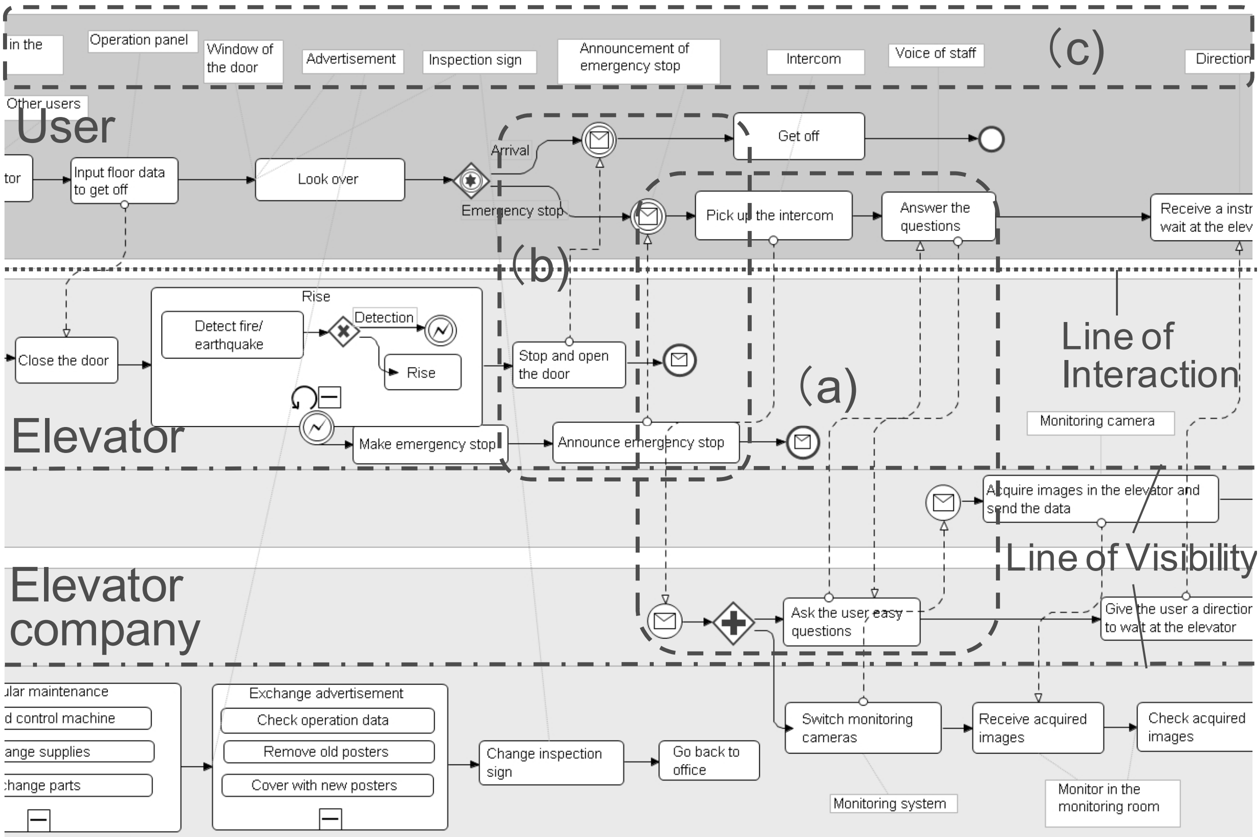


Figure 4: Service Blueprint of elevator operating service (partial)

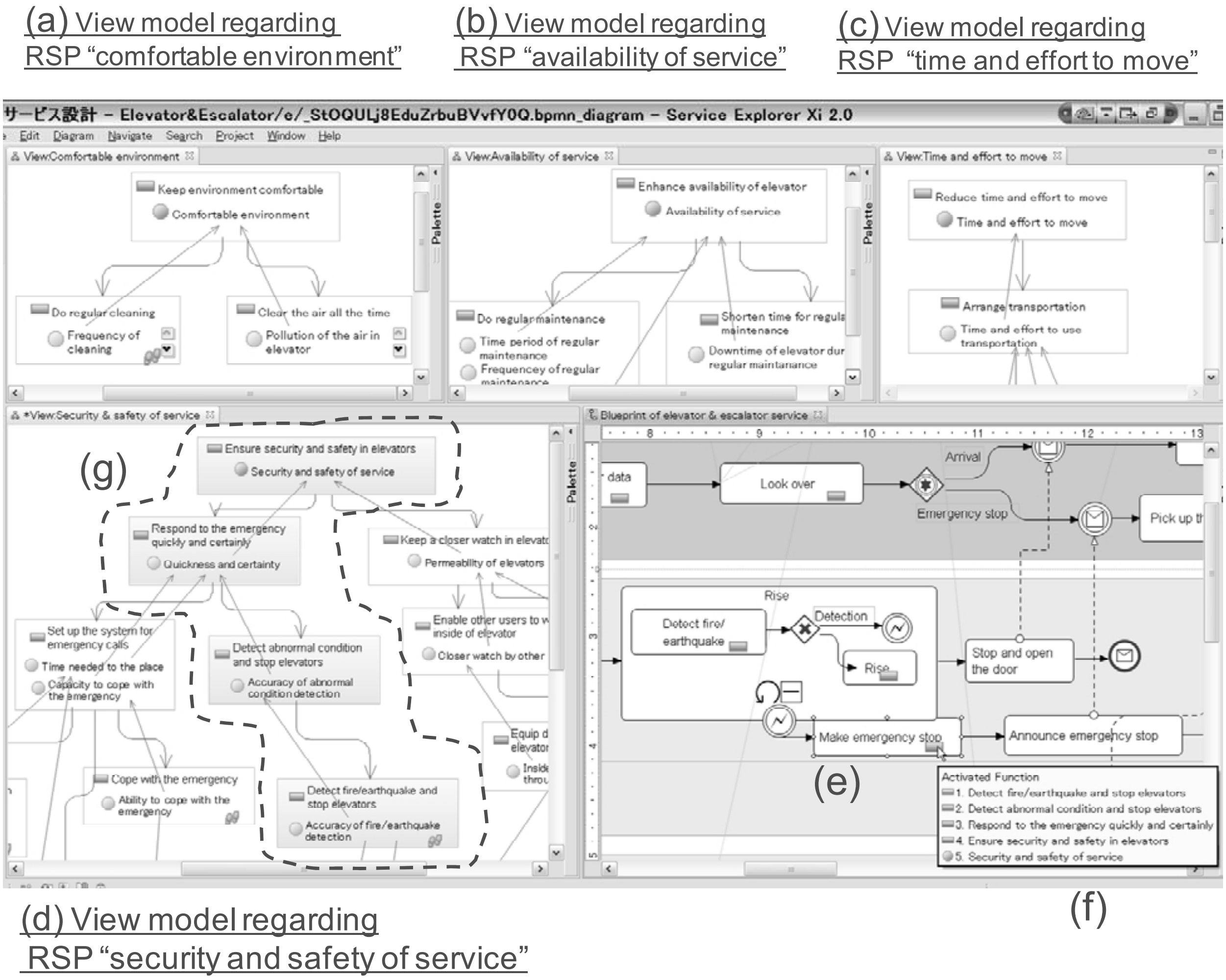


Figure 5: Collaboration of function-attribute editor and activity editor

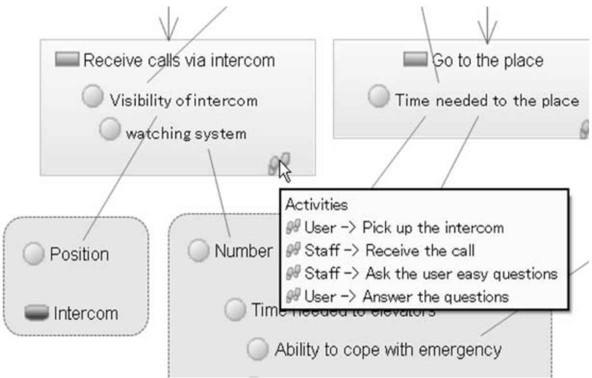


Figure 6 : Information of function actualization

Figure 5(f) and also highlights such RSP and functions in view model as shown in Figure 5(g). On contrary, functions in the view model are related to activities and behaviors as shown in the right part of Figure 6. We understand from Figure 6 that the function of receiving calls via intercom involves the user actions of picking up the intercom and answering the questions. The function requires the user actions other than activities of the service provider to be completed.

Thus, by combining the view model and the service blueprint, it is possible to analyze and design the service activities while confirming their influence on the user. The cross-functionality of the editors is an advantage of our tool.

6 DISCUSSION

Service blueprint described by BPMN

BPMN becomes widely used as standard notation of business process. Using BPMN, therefore, service blueprint gets to have a wide range of application and we can represent service delivery process and collaboration with customer on the computer. Furthermore, the aspect of service delivery is emphasized in business process by introducing marketing concepts such as service encounter, visibility of activity, and physical environment surround customer.

However, description of messaging on message flow is not adequate for now. Effect on receiver (receiver state) is essential of service.

Development of simulation method

Our method is not a simulation method, however, can help in the development of a simulation using the satisfaction indices of receiver, the functional model of service and the process model of service. The next stage of this study is to develop a simulation method that includes both aspects of function and service activity. Moreover, any simulation method developed should include costs and uncertain human elements. By evaluating designed solutions early, a PSS can be designed iteratively.

7 CONCLUSION

Customer value, function, and activity are key elements to design service offering like PSS. This paper presented a framework to correlate such elements; especially for function model and activity model. Our method and tool serve as a hub for existing studies in both marketing and engineering.

For engineering field, by incorporating information of activity into function model, the value of both physical products and service activities throughout product lifecycle can be well balanced. For marketing field, the gap between customer value and activity can be filled by incorporating functional meaning and effect to customer in activity model.

Future research will include the development of methods to specify customer value, to improve services, and simulate services.

8 ACKNOWLEDGMENTS

This research was partially supported by the JSPS Research Fellowships for Young Scientists.

9 REFERENCES

[1] Tukker A., Tischner U., 2006, New Business for Old Europe, Greenleaf Publishing

[2] Sakao T., Shimomura Y., 2007, Service Engineering: A Novel Engineering Discipline for Producers to Increase Value Combining Service and Product. Journal of Cleaner Production, Special issue on sustainable production and consumption - making the connection, 15(6), 590-604.

[3] Arai T., Shimomura Y., 2004, Proposal of Service CAD System -A Tool for Service/Product Engineering-. Annals of the CIRP, 53(1), 397-400.

[4] Sasser E., Olsen R.P., Wyckoff, D., 1978, Management of service operations, Allyn and Bacon, Boston.

[5] Zeithamal V.A., Parasuraman A., and Berry L., 1985, Problems and Strategies in Services Marketing, Journal of Marketing, 49 (Spring), 33-46.

[6] Fitzsimmons, J.A., Fitzsimmons M.J., 2001, Service management: Operations, strategy, and information technology. Third edition: McGraw-Hill

[7] Lovelock C.H., Wright L.K., 1999, Principle of Service Marketing and Management, Prentice-Hall

[8] Hill, T.P., 1977, On Goods and Services. The Review of Income and Wealth, 23(4), 314-339.

[9] Vargo S.L. S.L., Lusch R.F., 2004, Evolving New Dominant Logic for Marketing, Journal of Marketing, Vol. 68, pp.1-17.

[10] Chesbrough H., 2006, A research manifesto for services science, Communications of the ACM, Vol. 49, No. 7, pp.35-40.

[11] Fisk R.P., Grove S.J., John J., 2003, Interactive Services Marketing, Houghton Mifflin Company.

[12] Ramaswamy. R., 1996, Design and Management of Service Processes, Addison-Wesley, Reading, MA.

[13] Shostack G. L., 1982, How to Design a Service, European Journal of Marketing, Vol.16, No.1, pp.49-63.

[14] Shostack, G. L., 1987, Service Positioning through Structual Change, Jounal of Marketing, Vol.51, pp.34-43.

[15] Kingman-Brundage, J., Technology, 1991, Design and Service Quality. International Journal of Service Industry Management, 2(3), 47-59.

[16] Brooks, R. and Lings, I., 1996, A Hierarchy of Customer Satisfaction, The Inadequacies of Service Blueprinting, in J. Beraçs and J. Simon (eds.) Marketing for an Expanding Europe, Vol II, Proceedings of the 25th Annual Conference of the European Marketing Academy, Budapest, Hungary,147-64.

[17] Pires G., Stanton P., 2004, The Role of Customer Experiences in the Development of Service Blueprints, ANZMAC 2004 Conference.

[18] Kingman-Brundage, J., 1989, The ABCs of Service System Blueprinting, in Designing a Winning Service Strategy, eds. M. Bitner and L. Crosby (AMA, Chicago) .

[19] BPMN Information (http://www.bpmn.org)

[20] Havey M., 2005, Essential Business Process Modeling, O'Reilly Media, Inc.

[21] Shostack G.L., Planning the service encounter, In J. Czepiel & M. Solomon & C. Surprenant (Eds.), The Service Encounter, Lexington, MA: Lexington Books, (1985), pp.243-254.

[22] Umeda Y., Ishii M., Yoshioka M., Shimomura Y., and Tomiyama T., 1996, Supporting Conceptual Design based on the Function-Behaviour-State Modeler. Artificial Intelligence for Engineering Design, Analysis and Manufacturing, 10(4), 275-288.

[23] Hara T., Arai T., Shimomura Y., and Sakao T., 2007, Service/Product Engineering: a new discipline for value production. In Proceedings of the 19th International Conference on Production Research (ICPR-19), Valparaiso, Chile, CD-ROM.

[24] SOA Tools Platform Project (http://www.eclipse.org/stp)

[25] Eclipse.org (http://www.eclipse.org)

Remanufacturing of Products used in Product Service System Offerings

Erik Sundin[1], Johan Östlin[1], Anna Öhrwall Rönnbäck[1], Mattias Lindahl[1] and Gunilla Ölundh Sandström[2]

[1] Department of Management and Engineering, Linköping University, Sweden

[2] Department of Machine Design, School of Industrial Technology and Management, Royal Institute of Technology, Sweden

Abstract

As a product service system provider it is important to consider its benefits and drawbacks. Connecting the product service system with a remanufacturing system has a good potential of being economically and environmentally beneficial. This paper elucidates the case of three different remanufacturers and how their relation with their core provider affects their business. Products sold as a part of a product service system have great potential of being remanufactured in an efficient manner. This is for example due to large possibilities to plan the remanufacturing operations and to achieve pre-information about the cores coming in to the remanufacturing facilities.

Keywords:

Life Cycle, Service Engineering, Reverse Logistics, Case Study

1 INTRODUCTION

Remanufacturing is an industrial process where worn-out/broken/used products referred to as cores are restored to useful life. During this process, the core passes through a number of remanufacturing operations, e.g. inspection, disassembly, component reprocessing, reassembly, and testing to ensure it meets the desired product standards. This could sometimes mean that the cores need to be upgraded and modernized according to the customer requirements [1-3].

Remanufacturing is a rather large business in the United States [2], but also in Europe, where the industry has been growing lately due to its profitable business and an increase of company environmental awareness. Furthermore, environmental legislative pressure from the European Union (EU), such as the launching of the WEEE (Waste of Electric and Electronic Equipment) [4] and ELV (End of Life Vehicle) [5] directives. From an institutional viewpoint, the effects of theses directives can become a significant driver for the remanufacturing industry. How these directives are implemented will have a significant effect on the remanufacturing industry.

Besides to follow the current legislation and directives there are various motives for product remanufacturing e.g. increased profitability, ethical responsibility, secured spare part supply, increased market share and brand protection [6]. Furthermore, remanufacturing has also been shown to be environmentally preferable in comparison with other end-of-life treatments, since the geometrical form of the product is retained and its associated economic and environmental values preserved [3, 7, 8].

Material flows are an important factor for the overall remanufacturing system [9]. A traditional view on these closed-loop supply chains is that they encompass two distinct material supply chains: the forward and the reverse. Generally, the forward chain concerns the flow of physical products from manufacturer to customer, while the reverse chain describes the flow of used physical products from customer, then acting as supplier, to the remanufacturer. These flows are then 'closed' by, for example, the remanufacturing operation. One of the major differences between the 'forward' and the 'closed' supply chain is that the customer frequently acts both as a customer for remanufactured products and as a supplier of cores to the remanufacturing company [10].

Another means of closing the material flows is to focus on offering product/service combinations instead of selling physical products; a condition for this is that the hardware used in association with the service is reused, either as is or after it has been remanufactured or recycled. In research, to offer product/service combinations are referred to as have a product service systems (PSS). Moreover, the production and processing of raw material is decreased when the hardware is remanufactured. Thus, by having a strategy for both PSS and remanufacturing of products, both economic and environmental benefits can be achieved. Having a remanufacturing perspective on PSS, and vice versa, with a connection to product design, is a new way of looking at these concepts. This could give the manufacturing/ remanufacturing company new market shares for its products. The concept of PSS is briefly described in section 4.

In comparison to manufacturing, remanufacturing has some general characteristics that complicate the supply chain. For example, a company retrieves used products (cores) from the suppliers of cores, these suppliers are normally the end customers but it can also be scrap yards, core brokers or incurrence companies. As for end customers there is a major difficulty to assess the number and the timing of the returns. Another complicating issue is that the quality of the used products is usually not known [9, 11-13].

However, not much research has been published about how to successfully connect PSS offers with remanufacturing. In order to make the strategy of PSS more economic and environmentally beneficial these PSS offering companies should consider the remanufacturing aspects in more detail. A part of these with focus on the product design of the physical products used in the PSS offerings has been conducted [14, 15]. However, there is still much research needed in order to have a better effectiveness in the remanufacturing and reverse supply chains.

2 AIM

The aim of this paper is to explore how remanufacturers manage to operate the remanufacturing of product used in product service systems. This means to elucidate problems between the stakeholders and provide solutions to how to achieve a win-win-win situation between the PSS provider (often the original manufacturer), remanufacturer and PSS user.

3 METHODOLOGY

The methodology used for this research has been a mix of studying existing research in the area of remanufacturing and product service systems. To retrieve empirical data the authors have mostly used semi-structured interviews with the companies and also performing field studies and reading of company documents.

4 PRODUCT SERVICE SYSTEMS

The importance of services becomes larger as our society and its economic capacity matures. Service activities, for example, are the source of core value in the tertiary industry. In addition, the secondary industry has recently been cited as more interested in services (e.g. [16 and 17]). Examples of new businesses in the secondary industry include copy machine manufacturers, who sell the service of copied papers by the number of photocopies instead of copy machines (such as Fuji Xerox in Japan) and engine providers, who sell working flight hours instead of aircraft engines (such as Volvo Aero in Sweden).

Manufacturing companies bundling these new kinds of product-services offerings have put many different names on their offerings, such as the 'the diamond package' or 'the full-service bundle'. Meanwhile, researchers at both universities and companies have tried to name these packages in more general terms. Product Service Systems (PSS) (e.g. [18 and 19]) is an often mentioned concept for this new type of business. Other mentioned and similar concepts are for example Functional Sales (e.g. [20]), Functional Products (e.g. [21]), Integrated Product Service Engineering (IPSE) (e.g. [22]), Integrated Solutions (e.g. [23]), Hybrid Products (e.g. [24]) and Service/Product Engineering (SPE) formerly called Service Engineering (e.g. [25]).

Mont [26] defines PSS as *"a pre-designed system of products, services, supporting infrastructure, and necessary pre-arranged networks that can fulfill consumer needs on the market and, at the same time, minimize environmental impact"*. In the PSS literature, there is sometimes a strong focus on the service content of PSS offerings and PSS has e.g. from the environmental community [27] been described as a way to reduce the environmental impact from products and to dematerialize the society. However, there will never be any 100% dematerialized PSS offerings. There will always be a need for at least some physical products and that is illustrated in Lindahl and Ölundh [28].

Figure 1 below shows the continuum of product-services that these concepts stand for, as well as the many different definitions of these concepts. In common to these above mentioned business concepts, developers are attempting to incorporate service activities into the design space, one which has traditionally been dominated by physical products in manufacturing industries (see e.g. [14 and 15]).

5 TYPES OF REMANUFACTURING COMPANIES

There are different types of companies that perform remanufacturing. These companies can be divided according to their relationship to the product manufacturer, i.e. the Original Equipment Manufacturer (OEM). These following three categories of remanufacturers were also described by Lund [29] and Jacobsson [30].

5.1 Original Equipment Remanufacturers

Firstly, there are certain OEMs which remanufacture their own products; these companies are also called Original Equipment Remanufacturers (OERs). In this case, it is the OEM/OER who remanufactures its own products arriving from service centers, tradeins from retailers or end-of-lease contracts. For these OEMs/OERs, the remanufacture of products is profitable, and they can offer their customers a wider price range of products. Furthermore, OEMs/OERs have all the needed information concerning product design, availability of spare parts and service knowledge. The remanufacturing process could be integrated with the ordinary manufacturing process or be separated from it. Also, the parts from the remanufactured products could be used in manufacturing, or the products could be entirely remanufactured. An example of this kind of company is FUJI Film, which remanufactures its single-use cameras in Japan at the same facility as it produces these cameras.

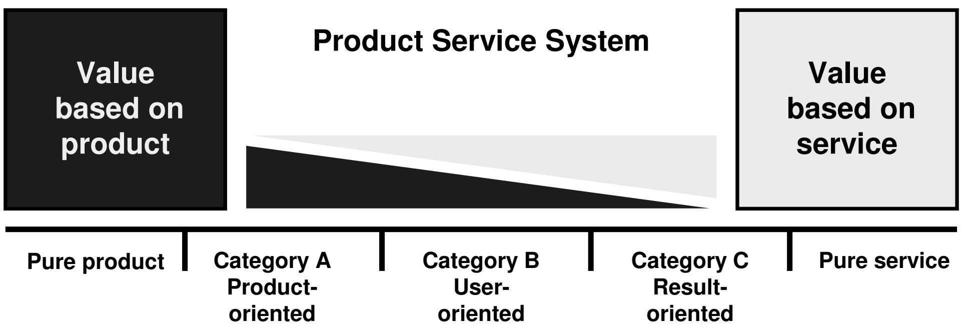


Figure 1. The product-service continuum. The picture shows the different categories of the continuum of product-service systems where a manufacturing company can choose to place their product/services (Based on Tischner et al. [31]).

5.2 Contracted remanufacturers

Secondly, there are remanufacturing companies that are contracted to remanufacture products on behalf of other companies. This means that the OEM normally owns the products, but does not need to perform the actual remanufacturing of them. Still, the OEMs have their products remanufactured and can offer them to their customers once again for a lower price. For the remanufacturer, there is likely to be a fairly consistent stream of business with fewer working capital requirements (e.g. work in progress) and risks, and the company can expect to obtain assistance from the OEM in terms of replacement parts, design and testing specifications, and even tooling [29]. An example of this kind of remanufacturer is Our-Way Inc., which remanufactures refrigeration compressors in Atlanta, USA [29].

5.3 Independent remanufacturers

Thirdly, there are many independent remanufacturers who remanufacture products with little contact with the OEM, and who need to buy or collect cores for their process. Sometimes, these companies are paid by the last owner or distributor to pick up discarded products [30]. These independent remanufacturers also often need to buy spare parts for their products that are to be remanufactured. The typical independent remanufacturer is a private corporation with ownership closely held [29]. Lund [29]] further states that this type of operation is an integrated one, in that it purchases cores, remanufactures them and markets them under its own name or for the private labels of others. Generally, exchange of experience between these remanufacturers concerning reprocessing to the OEM is minimal [30]. Furthermore, Hammond et al. [32] states that it is not likely that the relationship between independent remanufacturers and OEMs will grow in the future.

All remanufacturing firms, however, cannot be neatly placed into the categories above; some companies display a mix of these remanufacturing types. For example, an independent remanufacturer could have some of its products contracted with OEMs. In this study we have chosen a remanufacturer of each of these categories according to the following table (Table 1):

Table 1: Case study companies

Company	Type	Product	OEM
BT Products	OER	Forklift truck	BT Products
Wahlquists Verkstäder	Contracted	Filling machines	Tetra Pak
Storebro Maskin-renovering	Independent	Machinery	Various companies

6 RESULTS

As described earlier the option to remanufacture products during the lifecycle can be a source for increased competitiveness. Linked to the product a remanufacturer can be both independent, contracted and an OEM. The type of relationship to the product and the customer in a PSS has a high impact on the potential success of a remanufacturing process. In this paper, the pros and cons that can be found in industrial cases are presented.

6.1 OEMs – The BT Case

In the industrial setting the actor that has the potential to have the closest relation with the end customer is the OEM. For the OEM the relation with the customer can take many different forms, depending on the different business offers provided by the OEM. In the case for OEMs there are situations where the company leases, rents or uses other types of relationship based business offers that e.g. offers a wide range of after sales services. One frequently cited example of this is the Xerox copying machines and their PSS [8]. The studied case for this paper is the forklift truck (BT-industries) case. BT sells many of their forklift trucks on different rental programs. In this type of relationship, the linkage between the customer and seller is strong. The time range of this type of relationship makes commitment an important factor. A respondent puts it like this: *"for a successful relationship the customer has to feel that our products and services are reliable".*

The intensity can also vary dependent on the situation of the customer. In some cases, the product is rented to a customer for many years providing with regular servicing; other times, the products are rented for shorter periods, following for example seasonal demands. BT Industries, which provides many of its forklift trucks through rental programs, explains it to its customers as follows:

"Think of the advantages. You avoid all the risks of ownership. You don't consume capital – not even a deposit is required. We take care of the equipment and make sure it's always reliable, and we work with you to manage your fluctuating need – we call it capacity management. Rental from BT means flexibility for the future and all of your costs are predictable."

During the use phase, the seller is responsible for maintenance and repairs of the product; the maintenance technicians handle this commitment. By providing regular service and maintenance operations at the customer's location, a relationship is established through the service/maintenance personnel. Maintaining the relationship is an important factor of the BT policy of basing its service technicians at its major customers.

Another effect of a close relationship such as found in the BT case is that the service/maintenance personnel gain detailed information as to whether or not there is a need for a future remanufacturing operation. This information can make the return flow of products easier to control and provide information on the forklift truck status to the remanufacturing process. A high degree of control can for example be gained by high control of the products that are in operation at the customer, also called the installed base. By using information gathered from the installed base, the supply chain can be effectively coordinated. This information can be gathered by, for example, service/maintenance personnel as in the BT case or by software solutions integrated in the product (used by for example Xerox in photocopiers).

As a conclusion the OEMs are the actors that have the highest potential for control of the remanufacturing process. The OEMs might have detailed information on quantities and timing of returns according to contracts or information on the status of the product in use by the customer, through e.g. maintenance personnel or online monitoring systems. OEMs are also generally most favorable for gaining access to new components (spare parts) for a competitive price. Also, as an OER the remanufacturer can have access to all drawings of the products being remanufactured and also been given possibilities for upgrading the products according to the newest standards.

6.2 Contracted remanufacturers – The Wahlquist Verkstäder case

The contracted remanufacturers are a category of remanufacturers that are dependent on orders from other actors that sells remanufactured products. In this paper this category of companies are represented by Wahlquist Verkstäder (WV) that are affiliated with the filling machine manufacturer Tetra Pak. In this case WV is performing remanufacturing of filling machines according to orders from Tetra Pak. Here, Tetra Pak has chosen to outsource the remanufacturing to WV. The main reason for this is that the knowhow of remanufacturing and the cost structures of the WV are considered superior in comparison to a Tetra Pak managed remanufacturing process. In a contracted remanufacturing operation the major difference is that the actor performing the remanufacturing is separated from the control over a PSS. In comparison, the OEM normally has better position gain information about the products lifecycle compared to the contracted remanufacturer because they are closer to the customer in the supply chain.

According to Guide [33] there are generally a problem of balancing return of products suitable for remanufacturing and the demand for remanufactured products. In the case for the contracted remanufacturer WV, the difficulty of balancing returns and demand lies at the OEMs that are managing the PSS. Still, if the OEM has problems in matching supply and demand there can also be problems for the contracted remanufacturer mainly in fluctuating remanufacturing volumes due to uneven supply and demand from the OEM.

An additional problem of not being an OEM in the VW case is the low control of the quality of the incoming products that are subject for remanufacturing. Assessing the costs for the remanufacturing process can therefore be difficult according to quality of the returned product. This poses a risk for VW according to how the costs for remanufacturing are determined. In theory there are mainly two options to choose from.

The first option is to use a variable cost for the remanufacturing according e.g. the amount of the new components and man hours that has to be used in the remanufacturing process. This results in a disadvantage for the OEMs since they might have a difficulty in deciding to remanufacture a product or not. In a sense, the OEM takes the risks for assessing the cost for remanufacturing for a remanufacturing process that they have little control over.

The second option is to use a fixed price for the remanufacturing process. In this case all the risk is transferred to the remanufacturer. If a product demands extensive amounts of new material and man hours due to low quality of the incoming product it can result in total loss for the remanufacturing order. Vice versa, a high quality product can result in a high profit. In this situation the contracted remanufacturer becomes dependent on the decision from the OEMs whether to remanufacture or not.

6.3 Independent remanufacturers – The Storebro case

For the independent remanufacturers as in the Storebro case, it is common that remanufacturing of machines is done according to direct customer orders. In the Storebro case the machining product is apart of a manufacturing system (example in Figure 2). In OEM situations as in the BT case it is common that the product changes user in the lifecycle. In the BT case, when a product is returned for remanufacturing, it is mostly replaced immediately with a new or a remanufactured product. This is not the case in the Storebro situation where the products are remanufactured on orders from an end customer. This can result in some difficulties, say for example that the product to be remanufactured is a part of a system. In this situation the system will be non-usable for the time when the product is being remanufactured if no replacement product is used. This is not preferable for many situations; for example, a truck that needs to replace its turbo charger would have to stand still for a significant time while the turbo is being remanufactured. Another example is if a filling machine such as those described in the TetraPak case were to be sent to remanufacturing, the result could be that e.g. the complete process in a dairy would become inactive. Most surely, the expenses for not using the system (truck and dairy) are greater than the gains for the (potential) lower cost for remanufacturing. Still, this does not have to be the situation for a customer that can wait for the remanufactured component while it is being remanufactured. Therefore the potential for independent remanufacturers to perform remanufacturing in a PSS manner is limited to specific situations where products can be sent away for remanufacturing. There is also a possibility for independent remanufacturers to realize a PSS if remanufactured products can be delivered right away, for example, when the current product is sent away for remanufacturing as in a deposit based system [34] (this is not the situation in the Storebro case).

The just as for the other cases the cost of the remanufacturing is dependent on the quality of the core and how much new material and labor that will be needed in the process and this is difficult to determine before the remanufacturing is undertaken. In the Storebro case a fixed price is applied according to a pre estimated calculation. To assess the cost for remanufacturing a pre inspection of the machines are performed as early as possible. The inspections are generally done when a product is assembled and therefore it is difficult to estimate the reusability of individual components. This results in a risk of additional cost in the remanufacturing process.

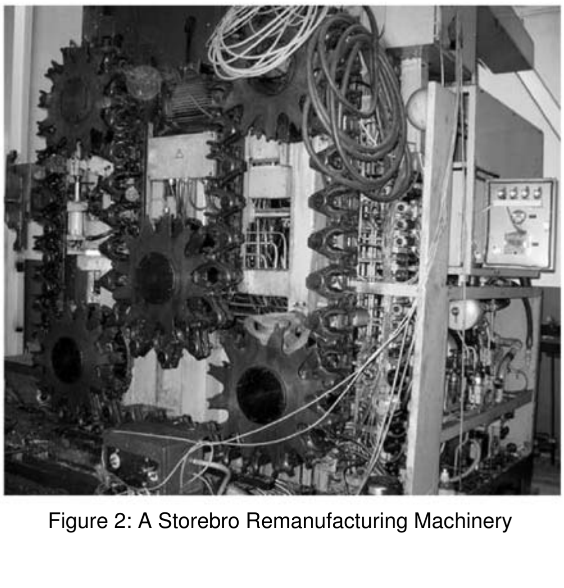


Figure 2: A Storebro Remanufacturing Machinery

One additional complicating factor for the independent remanufacturers is the competitive situation in regard to the OEMs. In many situations they are a direct competitor with the OEMs. This can for example result in difficulties to procure new replacement components for broken components since the only way is to procure these are through the OEM.

7 DISCUSSION AND CONCLUSION

According to this papers aim, this paper has illustrated and explained how three different remanufacturing companies manage to operate their remanufacturing of product used in PSS offerings. As shown, and also found in previous research, the relation that a remanufacturer has with its core provider is important for their business [34]. The more control they have over the development and use of their products, the easier they have to influence their ability to remanufacture their incoming cores.

Furthermore, as this study also has shown, products sold as a part of a PSS have great potential of being remanufactured in an efficient manner. This is e.g. due to the big possibilities to plan the remanufacturing operations and to achieve pre-information about the cores coming in to the remanufacturing facilities. This is especially the case when the PSS provider also is the OEM and remanufacturer as in the case of BT Industries.

8 ACKNOWLEDGEMENTS

The authors would like to thank all case study companies for participating in our study. This research was partially supported by the Swedish Governmental Agency for Innovation Systems (VINNOVA) which also should be acknowledged.

9 REFERENCES

[1] Seaver, W.B. *Design Considerations for remanufacturability, recyclability and reusability and reusability of user interface modules*. in *Proceedings of IEEE International Symposium on Electronics and the Environment (IEEE-94)*. 1994. San Francisco, CA, USA.

[2] Lund, R., *The remanufacturing Industry: Hidden Giant*. 1996, Boston, Massachusetts: Boston University.

[3] Sundin, E., *Product and Process Design for Successful Remanufacturing*, in *Production Systems, Department of Mechanical Engineering*. 2004, Linköping University: Linköping, Sweden.

[4] EU, *Directive 2002/96/EC of the European Parliament and the Council of 27 January 2003 on waste electrical and electronic equipment (WEEE).* Official Journal of the European Union, 2003.

[5] EU, *Directive 2000/53/EC of the European Parliament and of the Council of 18 September 2000 on end-of life vehicles.* Official Journal of the European Communities, 2000(L269): p. 34-42.

[6] Seitz, M.A. and K. Peattie, *Meeting the Close-Loop Challenge: The Case of Remanufacturing.* California Management Review, 2004. Vol. 46(No.2).

[7] Bras, B. and R. Hammond. *Design for Remanufacturing metrics*. in *Proceedings of First International Seminar on Reuse*. 1996. Eindhoven, The Netherlands.

[8] Kerr, W. and C. Ryan, *Eco-efficiency gains from remanufacturing - A case study of photocopier remanufacturing at Fuji Xerox Australia.* Journal of Cleaner Production, 2001. No. 9: p. 75-81.

[9] Guide, V.D.R.J., *Production planning and control for remanufacturing: industry practice and research needs.* Journal of Operations Management, 2000. 18: p. 467-483.

[10] Krikke, H., I. le Blanc, and S. van de Velde, *Product modularity and the Design of Closed-Loop Supply Chain.* Californian Management Review, 2004. 46(2): p. 23-39.

[11] van Nunen, J.A.E.E. and R.A. Zuidwijk, *E-Enabled Closed-Loop Supply Chains.* California Management Review, 2004. 46(2): p. 40-54.

[12] Guide Jr., V.D.R. and V. Jayaraman, *Product acquisition management: current industry practice and a proposed framework.* International Journal of Production Research 2000. 38(16): p. 3779-3800.

[13] Geyer, R. and T. Jackson, *Supply Loops and Their Constraints: The Industrial Ecology of Recycling and Reuse.* California Management Review, 2004. 46(2): p.55-73.

[14] Sundin, E. and B. Bras, *Making Functional Sales Environmentally and Economically Benficial through Product Remanufacturing.* Journal of Cleaner Production, 2005. Vol. 13(Issue. 9): p. 913-925.

[15] Sundin, E., M. Lindahl, and W. Ijomah, *Product Design for Product Service Systems - design experiences from three Manufacturing companies from Sweden.* Journal of Engineering Design, Submitted.

[16] Oliva, R. and R. Kallenberg, *Managing the transition from products to services.* International Journal of Service Industry Management, 2003. Vol. 14(No. 2): p. 160-172.

[17] Mathieu, V., *Service strategies within the manufacturing sector: benefits, costs and partnership.* International Journal of Service Industry Management, 2001. 12(5): p. 451-475.

[18] Goedkoop, M.J., C.J.G. Van Halen, H.R.M. te Riele, and P.J.M. Rommens, *Product Service Systems, Ecological and Economic Basics*. 1999, VROM: Hague, the Netherlands.

[19] Mont, O. and A. Tukker, *Product-Service Systems: reviewing achievements and refining the research agenda.* Journal of Cleaner Production, 2006. 14(17): p. 1451-1454.

[20] Ölundh, G. and S. Ritzén. *How do functional sales affect product development and environmental performance?* in *International Conference on Engineering Design, ICED 03*. 2003. Stockholm.

[21] Brännström, O., *Market Offer Development - Industrial experiences in a business-to-business context*, in *Department of Applied Physics and Mechanical Engineering, Division of Computer Aided Design*. 2004, Luleå University of Technology: Luleå, Sweden.

[22] Lindahl, M., E. Sundin, T. Sakao, and Y. Shimomura. *An Interactive Design Methodology for Service Engineering of Functional Sales Concepts – A potential Design for Environment Methodology*. in *LCE2006 - 13th*

CIRP International Conference on Life Cycle Engineering. 2006. Leuven, Belgium.

[23] Davies, A., *Moving base into high-value integrated solutions: a value stream approach,* Industrial and Corporate Change, 2004. 13(5): p. 727-756.

[24] Meier, H., *Lifecycle-based Service Design for innovative business models.* Annals of the CIRP, 2004. 53(1): p. 393-396.

[25] Arai, T. and Y. Shimomura, *Proposal of Service CAD System - A Tool for Service Engineering.* Annals of the CIRP, 53/1:397-400., 2004.

[26] Mont, O., *Clarifying the concept of product-service system.* Journal of Cleaner Production, 2002. 10(3): p. 237-245.

[27] Mont, O., Special Issue on Product Service Systems and Sustainable Consumption. Journal of Cleaner Production, 2003. 11(8): p. 815-933.

[28] Lindahl, M. and G. Ölundh. The Meaning of Functional Sales. in Life Cycle Engineering: Challenges and Opportunities: 8th International Seminar on Life Cycle Engineering. 2001. Varna, Bulgaria: CIRP.

[29] Lund, R.T., Remanufacturing Technology Review. 1983: p. 19-29.

[30] Jacobsson, N., Emerging Product Strategies - Selling Service of Remanufactured Products, in The International Institute for Industrial Environmental Economics. 2000, Lund University: Lund, Sweden.

[31] Tischner, U., M. Verkuijl, and A. Tukker, *First Draft PSS Review. SusProNet Report, draft 15 December. Available from Econcept, Cologne, Germany; TNO-STB*. 2002: Delft, the Netherlands, or www.suspronet.org.

[32] Hammond, R., T. Amezquita, and B. Bras, Issues in Automotive Parts Remanufacturing Industry: Discussion of Results from Surveys Performed Among Remanufacturers. Journal of Engineering Design and Automation, Special Issue on Environmentally Councious Design and Manufacturing, 1998. 4(1): p. 27-46.

[33] Guide Jr., V.D.R., Production planning and control for remanufacturing: industry practice and research needs. Journal of Operations Management, 2000. 18(467-483).

[34] Östlin, J., E. Sundin, and Björkman. M., Importance of closed-loop supply chain relationships. International Journal of Production Economics (accepted), 2008.

Integrated Design of Industrial Product-Service Systems

Jan C. Aurich[1], Eric Schweitzer[1] and Carsten Mannweiler[1]

[1] Institute for Manufacturing Technology and Production Systems, University of Kaiserslautern, Germany

Abstract

Technical services aim at enhancing the economical and ecological performance of industrial products. To systematically exploit their potentials for both manufacturers and their industrial customers, products and services need to be integrated. The resulting Product-Service Systems (PSS) can thereby be defined as customer life cycle oriented combinations of products and services, realized in an extended value creation network. To exploit the full potentials of PSS, product and service components of the PSS have to be designed integrated. Therefore, the presented approach covers the implementation of a PSS design process by selecting, combining and adapting appropriate process modules of existing product and service design processes.

Keywords:
Product-Service Systems; Design Process; Development Process

1 INTRODUCTION

Industrial customers are increasingly expecting to be provided with services such as maintenance, upgrading, operator trainings or process improvement. These services do not only contribute to keeping up existing product functionalities [1], but also provide additional ones along the whole life cycle [2]. Since in the past capital goods manufacturers have largely focused on design, realization, and distribution of high quality products, a gradual change of traditional manufacturing companies to producing service providers [3] that focus on customer solutions becomes necessary. To support this change and to provide the basis for realizing customer solutions in terms of benefit oriented Product-Service Systems (PSS), processes for product and service planning, design, and realization need to be integrated.

Based on a systematic planning of PSS that cover the specification of the properties of both the physical product as well as those of the corresponding services, product and service components of the PSS have to be designed integrated. Therefore, the presented approach covers the implementation of a PSS design process in an extended value creation network by selecting, combining and adapting appropriate process modules of existing product and service development processes.

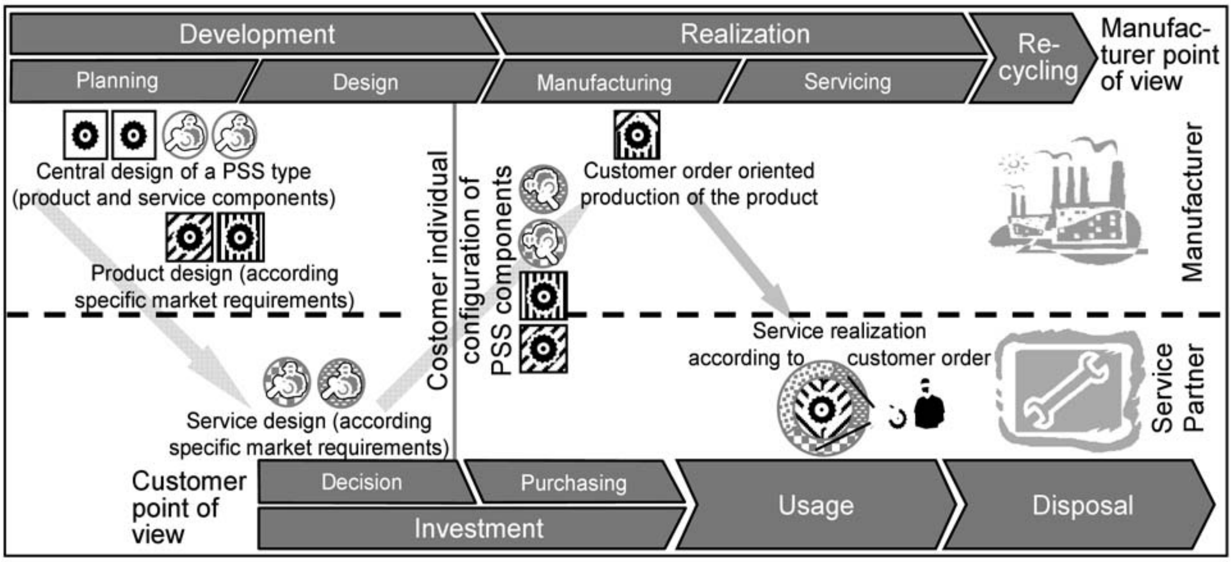


Figure 1: Design in PSS Life Cycle.

2 INDUSTRIAL PRODUCT-SERVICE SYSTEMS

Product-Service Systems can be defined as customer life cycle-oriented combinations of products and services, realized in an extended value creation network, comprising a manufacturer as well as suppliers and service partners [4], [5], [6]. In the capital goods industry, industrial PSS consist of a complex physical product dynamically enhanced along its life cycle by mainly non-physical services.

For design and realization of both product and service components, the product manufacturer relies on a network comprising multiple parts of suppliers and service providers. While the product core is produced at a limited number of production locations of the manufacturer, the services are provided at the place of product usage by the manufacturer's service branches or his independent partners in close cooperation with the customers.

With respect to designing and realizing industrial PSS, three dimensions need to be distinguished [7]:

- The product and service components together provide the customer with a certain set of expected functionalities that represent the result or product dimension of the PSS.
- PSS realization is based on different processes (e.g. product maintenance) that represent the process dimension of the PSS.
- The resources for servicing as well as for providing the manufacturer with continuous feedback are provided by the service network. It thus represents the information dimension.

3 PREPARATION OF THE DESIGN PROCESS

Establishing an integrated design process for PSS in manufacturing companies of the capital goods industry requires the fulfilment of corresponding organizational and operational prerequisites in terms of standardized processes and well defined organizational units and responsibilities [7]. Therefore, a process for systematic service design is introduced that upon its modularization represents a promising starting point for linkage with corresponding product design processes.

3.1 Analysis of existing Product Design Processes

In order to guarantee a high compatibility between product and service design processes, the existing product design processes have to be surveyed systematically. This analysis contains both the organizational and operational structures of the product design processes already implemented.

The description of the organizational structure aims at identifying the organizational units and persons involved. Design processes as well as additional preparative, post processing or supporting processes are described according to a predefined standard, too. Thereby, the specification of several tasks with clearly defined inputs and outputs of information and resources as well as the methods in support represents a main goal. An allocation of responsible persons to the so gained process modules as well as a description of their tasks, authority and responsibilities concludes the description.

3.2 Implementation of Service Design Processes

While systematic product design is well established in the industrial practice, the design of services is still predominantly performed with little or no systemization [8]. The specification of service design processes is thus described subsequently.

Service design processes aim at elaborating a standardized model based description of services. Based on a detailed analysis of documents existing in the value creation network, a common service model can be defined by scrutinizing existing services with respect to their result specifications, process descriptions, as well as the required resources. In order to guarantee a common standard, methods like e.g. morphological boxes, can be used to support the description. Since a systematic design of information exchange processes between the manufacturer, the service partners and the customers represents a main factor for the later provision of the manufacturer with continuous product and customer feedback, these processes have to be taken into account, too. Therefore, the information exchange completes the aspired service model.

Based on the results of analysis, the aspired service design process can be specified. Thereby, a standardized

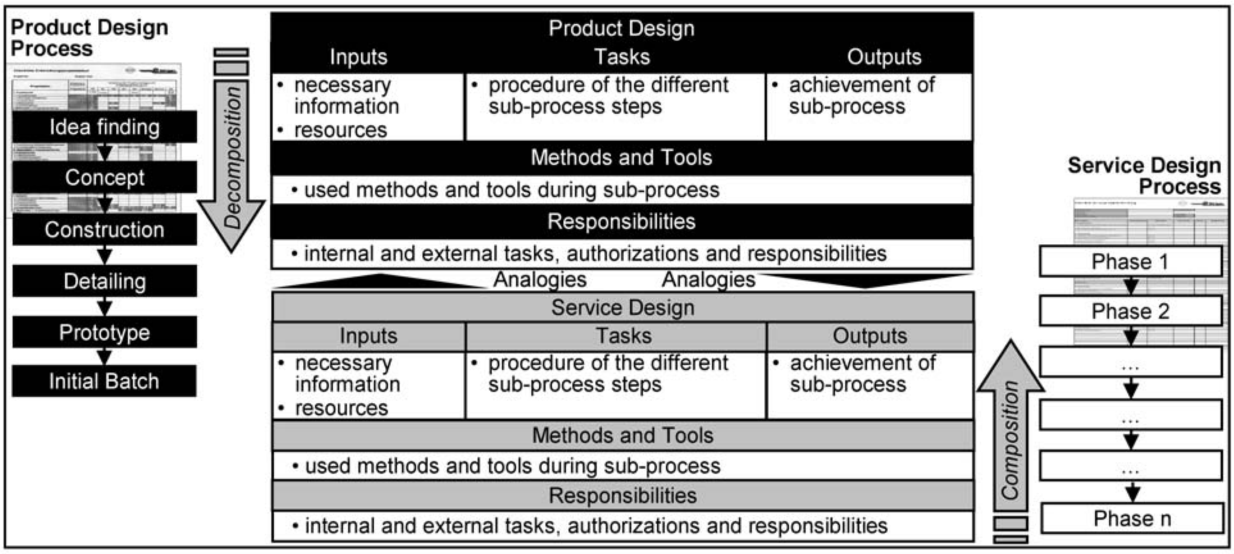


Figure 2: Implementation of Systematic Service Design Processes.

description of the output represents a first step. Thus, the uniformity of the product and service models in the entire value creation network is guaranteed. In a second step, the processes necessary for the generation of the service model are described according to the existing standards for process description. Based on this description, the process is modularized hierarchically in order to get a detailed description of sub-processes and several tasks. This modularization allows an analogy building with the product design processes analyzed in the previous step. Therefore, the existing product design process is also decomposed into several modules. The inputs and outputs of the resulting process modules represent standardized interfaces that allow their mutual linkage if the output of one process module matches the input of another one [4]. For each product design process module the corresponding potentials for describing the defined service product model are then evaluated and the activities and input / output information are altered and documented as necessary. The composition of the resulting process modules for service design leads to the required service design process (Figure 2). This procedure additionally ensures a comparable structure, documentation as well as a similar level of detail in the description of product and service design processes.

3.3 Implementation of a modular process library

In order to allow systematical usage of the resulting process modules for product and service design, a corresponding process library is implemented, which represents a common standard for the description of process modules and hence supports modelling of individual PSS design processes. Thus, a flexible and standardized configuration of the PSS design processes becomes possible.

The basic structure of the modular process library is provided, for example, by the Integrated Production Process Model (IPPM) [9], which represents a suitable framework. The single process modules of both product and service design processes contained therein are described according to a common standard. This allows a linkage of the process modules in case of agreeing input- and output information and resources. Possible linkages can also be regarded in the modular process library.

4 DESIGN OF INDUSTRIAL PSS

Based on the preparations described above, as well as the ideas for certain types of PSS generated in the previous PSS-Planning phase [7], the PSS is designed in the next step. The design phase aims at describing a new PSS-type according to the specified model structure.

4.1 Preparation

The preparation of the PSS design aims at providing the basics for the following design project. This contains both the definition of the design project goals and the establishment of a suitable project plan.

The definition of the goals to be fulfilled by the design project is based on the requirements specification generated in the previous PSS-Planning phase, that include a widespread description of the PSS properties. Thereby, properties of the basic type of the PSS and properties adaptable according to specific market requirements can be distinguished.

The organizational and operational structures of the design project follow the already implemented standards and structures of either product or service design projects. After dividing the design project into several phases (e.g. analysis of requirements, conceptual design, detailed planning and realization) the most important outcomes of the phases are determined. Based on this, separate tasks for designing the physical product on the one hand and the different services on the other hand are defined. Upon determining the project's milestones, the work plan is refined in a stepwise manner by means of the above process modules. In detail, the work packages are decomposed into separate tasks. Then, process modules for product and service design are selected from the process library in order to assure their standardized description. The assignment of tasks to be executed in the whole value creation network concludes the preparation.

4.2 Execution

The execution of the PSS design project aims at elaborating a model based description of the new PSS-type as well as the product and service models needed for its adaptation.

Therefore, the specified product and service design tasks have to be specified and linked. It becomes apparent, that an analysis of the existing interrelation between the different process modules represents the next step. To identify which outputs from one process module matches the input for another one, a modified Design Structure Matrix [10] (Figure 3) can be used.

		Service Design Process						
	Phases	Design		Modeling			Realization	
	Process Module	Mainten. Concept	Market Definition	Result	Process	…	Improve-ment	…
Product Design Process – Design	Concept	▼▲						
Product Design Process – Design	Layout	▼▲						
Product Design Process – Modeling	Drive Unit			▲	▲			
Product Design Process – Modeling	…							
Product Design Process – Real.	Proto-type						▼▲	

▲ Product design generates input for service design
▼ Service design generates input for product design
▼▲ Alternating correlation

Figure 3: Modified Design Structure Matrix.

Measures for bridging the resulting process interfaces between product and service design can then be specified in a pinpointed way to guarantee a systematic exchange of knowledge between the organizational units involved in service and product design processes. As a consequence, product and service design process modules can be

- Parallelized in case of independency,
- Sequenced in case of an unidirectional dependency, or
- Integrated in case of a bidirectional dependency.

Figure 4 summarizes the described PSS design cycle.

Upon defining the interrelations between product and service design, the execution of the corresponding tasks represents the next step. Since the definition of manufacturing processes

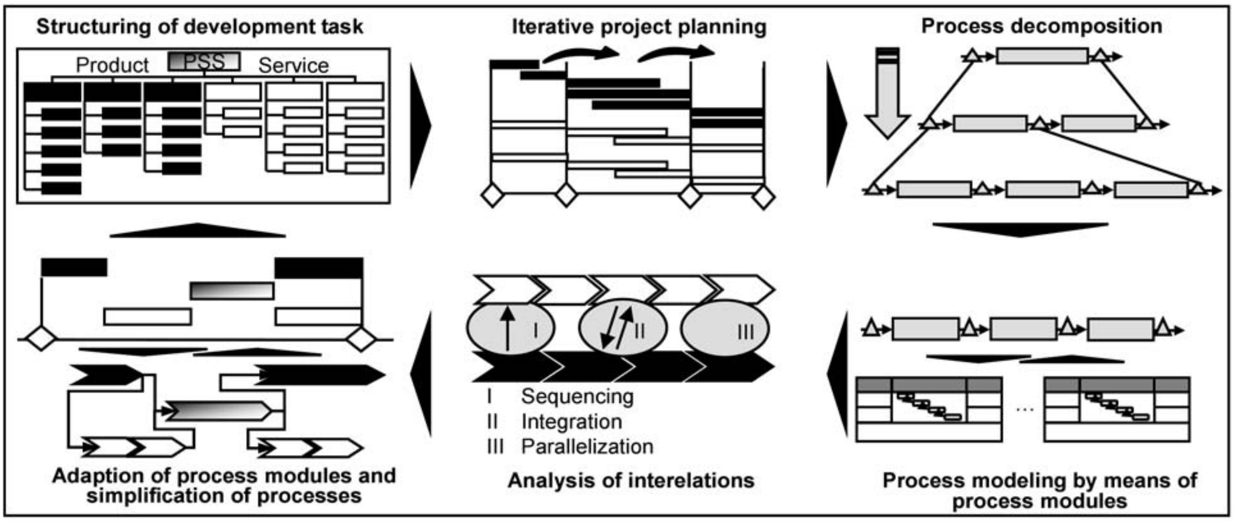


Figure 4: PSS Design Cycle.

for the product as well as servicing processes for the service components of the PSS thereby represents a major task, the organizational process library can again be applied in case the corresponding process module have to be specified beforehand.

As the main result of the design process, a complete documentation of the new PSS-type exists. While the product components can be adapted subsequently by the manufacturer at the place of production (e.g. according to a specific customer order), the service components are adaptable according to specific customer or market requirements either by the local service partners or branches by means of the specified service design process.

4.3 Postprocessing

To conclude the PSS design project systematically, a standardized documentation of the design project itself as well as the results in form of a detailed model based description of the new PSS-type has to be realized. The PSS design process ends with the preparation and realization of the worldwide market rollout.

5 CONCLUSIONS

The integrated design of industrial Product-Service Systems within the extended value creation network requires a systematic constitution of organizational and operational structures. To enable not only manufacturer but also branches or service partners to fulfil their specific tasks in view of PSS-Design, the structure of both PSS-Design processes and the aspired results has to be standardized for all of them. Therefore, the presented approach covers the implementation of a PSS design process, applicable in the whole value creation network, by selecting, combining and adapting appropriate process modules of existing product and service design processes.

6 ACKNOWLEDGMENTS

The research underlying this paper results from a project funded by the German Federal Ministry of Education and Research (BMBF) within the Framework Concept "Research for Tomorrow's Production" (fund number 02 PG 1031) and managed by the Project Management Agency Forschungszentrum Karlsruhe, Production and Manufacturing Technologies Division (PTKA-PFT).

7 REFERENCES

[1] Takata, S., Kimura, F., van Houten, F.J.A.M., Westkämper, E., Shpitalni, M., Ceglarek, D., Lee, J., 2004, Maintainance: Changing Role in Life Cycle Management, Annals of the CIRP, 53/2: 1-13.

[2] Aurich, J.C., Fuchs, C., 2004, An Approach to Life Cycle Oriented Technical Service Design, Annals of the CIRP, 53/1 : 151-154.

[3] Schuh, G., Friedli, T., Gebauer, H., 2004, Fit for Service: Industrie als Dienstleister, Carl Hanser Verlag, München Wien.

[4] Aurich, J.C., Fuchs, C., Wagenknecht, C., 2006, Life cycle oriented design of technical Product-Service Systems, Journal of Cleaner Production, 14/17: 1480-1494.

[5] Mont, O.K., 2002, Clarifying the Concept of Product-Service System, Journal of Cleaner Production, 10/3: 237-245.

[6] Baines, T.S. et al., 2007, State-of-the-art in product-service systems, Journal of Engineering Manufacture 221/10, 1543-1552.

[7] Aurich, J.C., Schweitzer, E., Fuchs, C., Life Cycle Management of Industrial Product-Service Systems. In: Takata, S.; Umeda, Y. (Ed.): Advances in Life Cycle Engineering for Sustainable Manufacturing Businesses, London: Springer, 2007, 171-176.

[8] Spath, D., Zahn, E. (Ed.), 2003, Kundenorientierte Dienstleistungsentwicklung in deutschen Unternehmen, Springer, Berlin, Heidelberg, New York.

[9] Aurich, J.C., Wagenknecht, C., 2003, Bausteinbasierte Modellierung unternehmensübergreifender Produktionsprozesse, Zeitschrift fuer wirtschaftlichen Fabrikbetrieb 98/12, 661 – 665.

[10] Eppinger, S.D., 1994, A Model Based Method for Organizing Tasks in Product Development, Research in Engineering Design 6/1: 1-13.

Evaluation of Product-Service Systems During Early Design Phase

Nicolas Maussang, Peggy Zwolinski, Daniel Brissaud

G-SCOP Laboratory, University of Grenoble, France

Abstract

The development of service engineering within the manufacturing systems domain requires a new approach to design such systems. A PSS design methodology that provides methods and tools to support engineering designers during the design phase is a necessity. One of the key issues during the development process is to assess and evaluate the solutions during the early design phase. The diversity of elements included into a PSS (physical objects, service units and their relations) must be taken into account during this evaluation. We propose in this paper a specific evaluation of PSS during the early design phase. This evaluation phase is part of a methodology that is being developed to design global and optimized sets of products and services. For this evaluation, criteria, economical and environmental factors are presented and applied on a case study.

Keywords:

Product-Service System, environmental evaluation, economical evaluation, design phase

1 INTRODUCTION

Service Engineering [1], Product-Service Systems [2], Functional Products [3]... All of these concepts emerged from a new selling approach in which the customer does not pay anymore to own the product but rather pay to have a use, result or functionality provided by a system. In those systems, physical objects and service units are involved together to realize technical functions and finally to deliver value to the customer. Those systems are complex and the elements are related to each others. Consequently, to facilitate the transition for companies to shift to this new business model, it is necessary to support engineering designers with tools and methods during the design process [4]. The development of such systems must not be separated between the product design and the service design. Both elements must be designed "simultaneously" within an integrated manner. For manufacturing systems, it leads to develop a global set that embeds products (or physical objects) together with the (technical) services and think about the whole life cycle of the system [5].

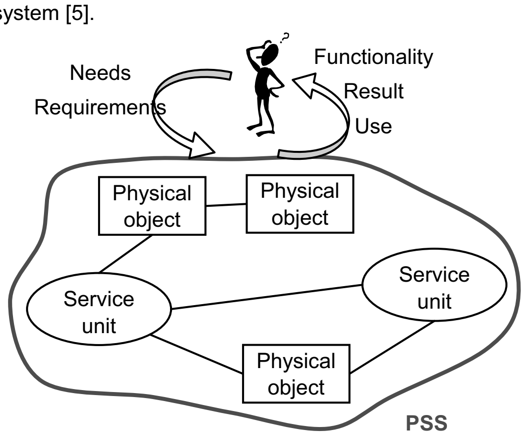


Figure 1: Elements of a Product-Service System

PSSs are defined as systems composed of physical objects, service units and relations between each others that ensure to the customer a result, a function or a use. Physical objects are the tangible entities that provide technical functions to the customer while service units are the organisational elements that ensure the availability of technical functions but also intangible functions to the customer (e.g. training). Service units can be technical (e.g. maintenance, upgrade, delivery) and also in contact with the customer (e.g. call centre, invoicing). Consequently, those systems deal with technical and organisational specifications. Technical specifications are due to the physical objects involved into the system that will provide a tangible output to the customer (e.g. to measure, to wash, etc.). The organisational specifications come from the necessity to establish the overall organisation between the physical elements and the service units. Service units are mainly technical skills labour that will bring an added value to the physical products (e.g. training, availability of assets, etc.). The organisational aspect leads to link the elements. For the customer, this is the whole system which fulfils his requirements (see figure 1). In this economy, it is a global set that is provided to the customer. From an external view (the customer's view), there is no more border between the products and the services of the system. It is only the "final result" brought by the system that has an importance for the customer. The basic idea of PSS is to fulfil the customers' expected functions by thinking at the end to a global system.

But the shift to this paradigm is questioning the existing engineering design processes. From an engineering point of view, is it possible to conclude that the physical objects involved into Product-Service Systems will be the same as the ones developed into a mass production paradigm? The answer is tricky. We assume that the engineering design process has to be modified in order to develop these systems and have a global optimised package of physical objects and service units.

The next section details the development of Product-Service Systems. Some elements that influence this development are detailed to show the need for reconsidering PSS design. On the other hand, the modelling and scenario-based implementation are described. The third section details the reason for an evaluation phase. Some elements about the criteria for evaluating PSS are detailed. To illustrate the evaluation proposed, a case study is used to implement the evaluation. Finally, discussion and conclusion will finalise the paper.

2 DEVELOPMENT OF PRODUCT-SERVICE SYSTEMS

2.1 What are the elements influencing the development process of PSSs?

Considering the PSS concept and the way to provide service instead of product, it seems that several elements have an influence on the success of service selling. By the way, these “external” elements will have to be taken into account during the process.

Business context

First of all, the economical context in which the service is provided (through PSS) will have an influence on its success. By comparing products to PSSs, it seems that in B2B the potential of PSS will be more understandable and more successful. Consider the example of a production line that has a specific technical task, e.g. freezing. The enterprises of this production line bought in the first case a fridge to fulfil its requirements. Considering that the production line must not be stopped due to a productivity requirement, if the fridge failed the line will be stopped. Consequently, the availability of the fridge has an influence on the whole availability of the line. Providing result/function instead of product seems to have a great potential. If the PSS developed enables the company to ensure the availability of its own production line, company will not hesitate to put in place a “freezing PSS” instead of the product (the fridge) alone. Generally, companies have more interest into productivity criteria and consequently can be more aware of Product-Service Systems advantages. Compared with the B2C context, the sense of belonging is stronger, because it shows a certain social success of the product owner (e.g. cars). That is the reason why selling services to private individual may be more risky in a short term period. In any cases, the company that will provide PSS instead of products will want to have some guarantees for this new business model. The risk of a service economy for a company is to have a critical mass of customers in order to be economically feasible. But on the other hand, to encourage customers to buy service instead of product, the economic criteria will be predominant. By the way, economical criteria for both companies and customers will have to be detailed to ensure the success of this new economy.

Elements of the system

As said previously, Product-Service Systems are made of physical objects and service units. By the way, the engineering designers must not be only focused on the products, but alternatives provided by service units must be considered carefully. By the way, the design process must lead to detail several solutions based on physical objects and service units. This is the main difficulty for the development of these systems because of the heterogeneity of the elements (physical elements vs. intangible services). It is thus necessary to have a methodology that takes into account those specificities. Moreover, the assessment of the design process will have to be realised on both elements. So, it is compulsory to find factors that take into account physical objects and service units.

Partnerships and internal organisation of the company

The identification of partners for implementing a PSS development is quite necessary. This is mainly due to the fact that companies enlarge their competencies by thinking about service units and not focus only on the product. In most of the cases, service units are put in place in order to provide a particular assistance to physical products to ensure the availability required by the customer. But to realise it, the company alone can not realise it. That is the reason why partnerships must be identified early in the design process to be able to propose alternative service units and finally propose more value to the designed system.

2.2 The deployment of functions

Customers’ needs are the primary elements necessary to develop systems (service, product or Product-Service System). Those requirements make up the technical specifications of the system. Once they have been detailed, the designer has to highlight the functions that will fulfil customers’ needs. This deployment is realised through the FAST diagram (Function Analysis System Technique). In this deployment, the main function is decomposed into several technical functions. These functions are then detailed into technical solutions. In this tree diagram description, the technical solutions can be either physical objects or service units. Several kinds of solutions can be envisaged. But this deployment does not indicate how these elements are linked between each others and does not give any indication on the general organisation of the system. Moreover, it is important to keep a “high-level” of description of the technical function and not to focus on a particular solution element. The aim of this method is to regard the overall system and to see which technical functions could be employed to realise the result expected by the customer. Afterwards, an internal modelling is required.

2.3 Modelling of a PSS through the description of scenarios

The modelling proposed is based on the functional bloc diagram of the Functional Analysis [6]. It enables engineering designers to map all the elements of a particular solution. Both physical elements and service units are detailed and linked. The links between elements are due to a particular relation between elements. The details of these relations into the functional bloc diagram is realised by a description of scenarios. As a link is assimilated to a particular action of an element regarding another one, the action oriented description of a scenario leads to show the relations between elements. It draws up the general organisation into the system. This organisation is influenced by the technical functions required to fulfil the function or result expected by the customer. Each technical function embeds elements and links. To conclude, this representation of Product-Service System permits to the engineering designers to develop PSSs by defining the elements needed and detailing the general organisation between elements. But when several solutions are envisaged, it is thus necessary to evaluate and compare solutions to choose the more relevant.

3 NEEDS FOR AN EVALUATION OF SOLUTIONS

When engineering designers develop a product, or service, or a product-service system, several solutions can be envisaged to fulfil customers' needs. Consequently, the evaluation phase is one of the most important phases. First of all, the elements included in a solution must respect all the customers' specifications. Secondly, regarding the Product-Service Systems within sustainable development, economical, environmental and social aspects must be taken into account. In this paper, the economical and environmental evaluations are addressed.

3.1 Which principle?

Eco-efficiency is one of the concepts used to integrate an economic and ecologic evaluation into the design process. Lehni [7] proposed 3 main objectives to improve the eco-efficiency:

1 Reducing the consumption of resources.
2 Reducing the impact on the nature.
3 Increasing product or service value.

It also exists checklists in order to consider the triple bottom line of sustainable development, as for example [8]. Regarding the ecological aspect, the most relevant study of the impact of a product, a process, a service or anything else is the Life Cycle Analysis approach. But, this evaluation is realised after the development phase and most of the time realised by environmentalists. So, few engineering designers are involved in this kind of evaluation [9]. This is the reason why it is necessary to evaluate the ecological aspect of a product during the early phase of the development. Some approaches are developed and can be implemented early in the design process. The "lifecycle bricks" [10] is one of the tool that can be used during the early design phase. It enables the designers to consider the environmental impacts of the components that are involved into the system. A single score or a multiple criteria view can be done with this method. In the paper, we focus on the use phase.

3.2 Factors and criteria

Environmental indicators

In LCA, 7 fields are identified to evaluate the ecological impact of a product or service: raw materials, hazardous substances, energy, air pollution, water pollution, soil pollution and waste. Moreover, data of the whole lifecycle must be gathered to realise the evaluation. That is the reason why the implementation of this analysis is realised at the end of the development phase.

For the environmental evaluation of PSS, indicators must give an overview of the system's ecological impact to the engineering designers. Three factors can be identified that influence the environmental impact:

- Material: the amount of raw material consumed and waste created at the end of life of physical objects give an idea about the impact on the (tangible) resources.
- Energy: for both physical objects and service units, energy is necessary to operate. Several kinds of energy can be identified, for example electricity, water and fossil fuel.
- Emissions: material and energy are the "inputs" of the system, but some outputs can also be quantified. In the emissions factor, CO2 emissions are mainly tackled. It is clear that others kinds of emissions can be addressed but it depends on the cases studied. In the case of PSS with an important transportation function, the CO2 produced is relevant of the emissions produced. But in other cases (e.g. chemical process), others emissions will have to be integrated into this factor as for example hazardous substances.

The material factor identifies the raw material necessary to manufacture physical objects. An object used for a certain period (a certain number of hours) is then scrapped at its end of life. Tangible objects are the main waste of PSSs. Consequently, this indicator gives to engineering designers an indication of resources necessary for the physical objects included into the system.

Energy is necessary for both physical objects and service units to operate. In this paper, this indicator tackles only the use phase. So, the energy necessary to manufacture physical objects is not taken into consideration. Several kinds of energy are identified: the electricity, the water, and the fossil fuel. These factors indicate the amount of energy required by the system to be run.

In this paper, the CO2 emissions are mainly addressed. When moves are required to repair, balance (or anything else), CO2 is produced. Moves are not the only aspects of CO2 emissions. Functioning of physical objects can also emit CO2 (or equivalent CO2 emissions). Consequently, it can also indicate the rate of emissions produced by physical objects during the functioning.

Economical indicators

If one thinks at the economical indicator, the first emerging criterion is automatically the purchase price. But behind the price, it is also necessary to regard the cost over the use phase of the system. Indeed, costs for maintenance and expendable can also be included. Consequently, 2 factors are identified for the economical indicator:

- Purchase/exchange: purchasing physical object is a non-negligible cost to take into account. The purchase price of physical objects can not be unlinked to its own life span. For example, probably a higher product price induces a more reliable product. Consequently, this factor has to be compared to the number of cycles of use. In addition, the replacement of components (i.e. the "expendables" of the system) must be included into this factor. One more time, life span can not be unlinked to the cost of components.
- Assistance: to ensure the smooth functioning of the system and ensure availability to the customer, maintenance actions (preventive or curative) will be necessary. The maintaining of the system leads to technical actions that have a cost. The assistance factor also includes intangible function like training, advising, etc.

These two factors highlight the fact that such systems have a capital cost due to the physical objects included to realise technical functions. Moreover, service units provide more added value by proposing new functions, or simply by maintaining a particular level of performance to ensure the system availability. The "assistance" factor deals with the inclusion of service units to maintain the functionality. But, like the emissions factor of the environmental indicator, the assistance factor is maintenance oriented in our case. But in other particular PSSs, it could be more related to the training of the customer, the support, or anything else.

Consequently, economical factors are linked to the robustness and reliability of products. Criteria like the rate failure and the use time have a direct influence on the purchase factor that can be assimilated to a redemption factor, as well as the assistance factor that deals with maintenance actions that can be realised on the object. The assistance factor also monitors the service units involved in the system.

Physical objects and service units criteria

Each of these factors takes into account several criteria from physical objects and service units. Regarding physical products, purchase cost is a predominant criterion. But also criteria relative to the robustness of the objects must be included. That is the reason why criteria like life span (use time span would be probably better), the probability of object failure, Time to Repair (T2R) must be considered into the evaluation functions. Considering maintenance operations, the less complicated the product and the cheapest these actions. But a complexity criterion is not easy to grasp, that is the reason why time to repair is a good indication of the object complexity for maintenance operations. Considering environmental criteria, energy consumption and emissions production must be identified to calculate the environmental impacts. Finally, the interpretation of a material factor is linked to the quantity of physical objects replaced during the maintenance actions. The mass of objects replaced is the relevant parameter of the function.

Service units are mainly composed of technical skilled people who realise technical actions (e.g. repair, replace, advise, train, etc.). Consequently, the salaries of people involved into the service are economical criteria that have to be taken into account during the evaluation phase. On the other hand, the ecological impact is mainly due to the moves and displacements of the service units. Moreover, in the case of a maintenance service as an example, the number of objects checked and repaired has to be integrated into the maintenance factor. Regarding environmental considerations, the energy consumption and emissions production are parameters identified. Finally, economical costs are also linked to the kind and number of operations realised and the estimated time for these operations.

3.3 Evaluation functions

In the evaluation phase, both economical and ecological indicators have to be implemented to compare the proposed solutions. In order to give an overview of these evaluations, it is necessary to express them for a specific time unit. This is the same principle that is used for LCA. LCA scores are given for a functional unit (e.g. to go from A to B, to transport 50kg). The unit used for PSS evaluation is a day of functioning. During a day, each physical object and service unit can consume material and energy, produce emissions while having a capital cost of physical component and a cost associated with the assistance actions.

Environmental evaluation

As said previously, the environmental evaluation can be split in 3 axes, which are the material used, the energy consumed and the emissions produced by the system. Each of them has to be expressed for a specific time value. Accordingly, each factor is expressed as the following.

- $\mathcal{F}$(Material)= $\sum_i mass_{obj\_i} * \%\_failure_{obj\_i} * \%replac_{obj\_i}$
- $\mathcal{F}$(Resources; Electricity)= $\sum_i KWh\_day_{elements\_i}$
- $\mathcal{F}$(Resources; Water)= $\sum_i l_{water}\_day_{elements\_i}$
- $\mathcal{F}$(Resources; Fuel)= $\sum_i l_{fossil\_fuel}\_day_{elements\_i}$
- $\mathcal{F}$(Emissions)= $\sum_i km/day_{vehicle\_i} * CO_2/km_{vehicle\_i} + \sum_j CO_2/day_{obj\_j}$

The material factor is expressed as the physical object mass multiplied by the percentage of object replacement on broken objects. It means that during their life span, a certain number of this object have to be replaced to ensure the functioning. It is expressed in kg per day of material replaced into the system. The resources factors (energy, water, fossil fuel) are the sum of the consumption of each element (either physical objects or service units). For the emissions factor, the sum of CO2 production per day per element gives an indication on the global system emissions. The emissions can be produced by the moves of a service unit or by the functioning of a physical object. Consequently, each factor is expressed per day of functioning. This is important to detail the number of days of use for physical elements in order to homogenise with the production emissions and resources consumption.

Economical evaluation

For the economical indicator, 2 factors are put forward: the purchase/exchange factor and the factor related to the assistance actions. Economical factors are expressed as follows:

- $\mathcal{F}$(Purchase)= $\sum_i \frac{Cost_{obj\_i}}{life\_span_{obj\_i}}$
- $\mathcal{F}$(Assistance)= $\sum_{technical\_actions} employees * salary$

The purchase factor is relative to the price to pay for buying components. Each component has its own life span (i.e. n days of functioning). This factor is relative to the depreciation of physical objects. It indicates, for one day of use (i.e. a theoretical 24 hours of use by customers) the cost of physical objects. Consequently, the sum of the daily cost object indicates the purchase factor of the system. About the assistance cost, this factor is relative to the number of people required to maintain the system, to advise or help customers. The size of service units is particularly important. As an example, a maintenance unit is based on the percentage of failures and the time necessary to fix those breakdowns. It is also necessary to introduce the operational time of technical people. For example, it is necessary to go to where failed objects are located before starting a reparation action. Consequently, for a 8 hours labour day, maybe 75% are devoted to the technical operations.

Finally, the purchase cost is expressed per day of functioning as well as the maintenance cost. The former depends on the capital costs of objects while the latter is determined by the number of maintenance actions that can be realised for an object.

4 COUPLING ENVIRONMENTAL EVALUATION WITH THE ECONOMICAL ONE

Even if an economic indicator was proposed on the one hand and an ecological indicator on the other hand, it is assumed that they must be coupled to have a whole indication on the solutions. This coupling must support decisions making during the design process. For example, how engineering designers could choose between a high reliability objects versus an efficient technical maintenance service? In that case, coupling the economical criteria with the environmental one can be a way to help engineering designers' choices.

4.1 Case study

The case study of Vélo'v, a renting bike system installed in Lyon (France) is used to illustrate this assumption. The case study is discussed to represent it during the design phase [11]. In this case study, a technical function ensures the smooth functioning of the system. Several elements are involved into this technical function. There are physical objects (bike, station, lock pad, etc.) and also service units (the maintenance unit, etc.). Those elements are linked, due to the actions that are necessary to start the maintenance process repair. To declare a failure, the customer must indicate it to the station when he brings back the bike. When the maintenance unit goes to the station, the broken bikes are indicated to the maintenance unit that can fix them. If the failure is too important, the bikes are brought to a workshop. Only little failures are fixed on site by the technical team. By the way, to be able to repair bikes, the maintenance unit must have an action on the station and on the lock pads that store bikes. All these actions and interactions between the elements can be modelled through a functional bloc diagram (see figure 2).

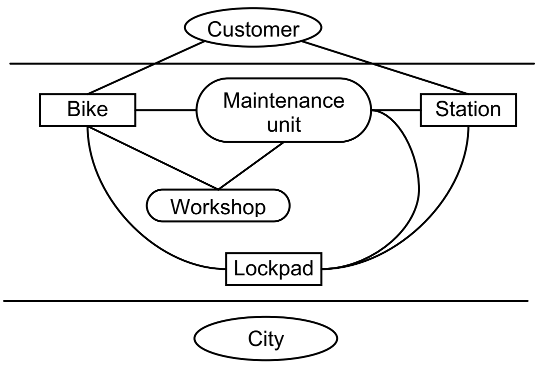


Figure 2: Partial Functional Bloc Diagram for the technical function "ensure the smooth functioning"

4.2 Linkage between economic and environmental evaluations

The partial bloc diagram represented figure 2 illustrates the links between the elements involved. The focus is particularly on the bike and the maintenance unit. To conduct the economical and environmental evaluations, some criteria must be quantified. In this evaluation, only criteria regarding the bikes and the maintenance unit are addressed. To illustrate the influence between the environmental and economical evaluations, 2 cases are envisaged (see table 1).

	Bike	
Criteria	Case A	Case B
Cost (€)	1500	1800
Mass (kg)	25	26
Span life (day)	800	850
Time to repair (hr)	0,5	0,5
% failure in one day	10	8
% replacement of broken bikes	20	15
Number of bikes	4000	4000

Criteria	Maintenance unit	
Salary (€/day)	60	
Fuel consumption (l/100km)	12	
CO2 emissions (g/km)	200	
% Time repair	75	
% Time transport	25	
Moves (km)	50	40
Number of technical people	33,33	26,67

Table 1: Bike and maintenance unit criteria

To size the maintenance unit, it is necessary to take into account the number of bikes and the time necessary to repair failed bikes. This calculation is based on an 8 hours labour day:

$$Nb\_people = \frac{Nb\_bikes * \%failure * Time\_to\_repair}{8 * \%\_repair\_time}$$

Moreover, the kilometres covered by a technical people to go to the repair location depend on the amount of technicians. That is the reason levels of moves differ. For these evaluations, formulas are detailed:

- $\mathcal{F}$(Material)= $Nb\_bikes * \%\_failure * \%\_replac * mass$
- $\mathcal{F}$(Fossil_fuel)= $Nb\_tech * moves * consumption$
- $\mathcal{F}$(CO2)= $Nb\_tech * moves * CO2\_emissions$
- $\mathcal{F}$(Purchase)= $\frac{Nb\_bikes * Cost}{Span\_life\_bike}$
- $\mathcal{F}$(Maintenance)= $Salary * Nb\_tech$

The variations in cases A and B are due to the robustness criteria of bikes. In case B, the bike is more expensive and heavier. But on the other hand, the percentage of failure is reduced as well as the replacement rate of failed bikes. Each case is evaluated according to the level of criteria detailed in table 1. The results are gathered into table 2.

In case B, the bike involved in the system is more expensive and consequently the purchase factor is higher. On the other hand, as the failure rate is lower, the factor for maintenance decreased, as well as the fuel consumption and CO2 emissions. It is mainly due to the fewer number of technical people required to repair bikes. Moreover, bikes B are heavier, but the material factor decreased because of the rate replacement that is less important than case A.

Factors	Case A	Case B	Cases B/A
Material (kg/day)	2000	1248	0,624
Fossil fuel (l/day)	200	128	0,64
CO2 (kg/day)	333,3	213,3	0,64
Purchase (€/day)	7500	8470,6	1,13
Maintenance (€/day)	2000	1600	0,8

Table 2: Evaluation results

4.3 Discussion

It is clear that this evaluation is only realised on specific elements of the system. The global performance of the system must be realised on all elements. In addition, the criteria for physical objects and service units will also depend on the kind of PSS and the kind of technical functions (e.g. transport oriented, transformation oriented functions, etc.).

Economical factors (purchase and maintenance) could be added because they are expressed as euros per day. So, if the comparison between case A and B is realised, the ratio between B and A is 1,06. It means that economically, solution A is better than B. But regarding the environmental factors, B is the best solution. It is clear that companies would compare at first the economical evaluation. A solution, to put at the same level economical and environmental evaluations, would be to give an “economical value” for environmental indicators. It can be realised for the fossil fuel factor and also for the CO2 emissions by regarding the CO2 market. Material also could be “translated” into economical factor by having the price of raw material.

5 CONCLUSION

This paper puts in place some factors that can enable engineering designers to evaluate Product-Service Systems during the design phase. Those economical and environmental factors use criteria relative to physical objects and service units. In the case studied, criteria relative to a physical object and a service unit enable the engineering designer to evaluate solutions envisaged during the design process. The objective of the evaluation is to aid the decision making during the design of Product-Service Systems. As elements differ by nature (tangible vs. intangible), those factors involving service unit criteria and object criteria, can bridge this gap and give a global overview of the system.

It is clear that the social evaluation of PSS would close the loop to assess the sustainability of solutions. But before bridging this gap, the implementation of a tool will be the next step of our PSS design research. The work will consist in coupling several methods used into the methodology in order to have an integrated global indicator that could be used by engineering designers during the whole design process.

6 REFERENCES

[1] Tomiyama, T., 2001, Service engineering to intensify service contents in product life cycles, EcoDesign 2001, Tokyo.

[2] McAloone, T.C. and Andreasen, M.M., 2004, Design for utility, sustainability and societal virtues: developing product service systems, Design 2004, Dubrovnik.

[3] Sundin, E., Lindahl, M., Sakao, T., Shimomura, Y. and Björkman, M., 2005, New engineering design for functional sales business, ICED'05, Melbourne.

[4] Maussang, N., Brissaud, D. and Zwolinski, P., 2007, Which representation for sets of products and associated services during the design process?, In S. Tichkiewitch, et al., eds. Advances in Integrated Design and Manufacturing in Mechanical Engineering II, (Springer).

[5] Westkamper, E., Alting, L. and Arndt, G., 2000, Life Cycle Management and Assessment: Approaches and Visions Towards Sustainable Manufacturing, Annals of the CIRP, 49/2: 501-526.

[6] Delafollie, G., 1991, Analyse de la valeur, (Hachette, Paris).

[7] Lehni, M., 2000, Eco-efficiency: creating more value with less impact, WBCSD (Geneva).

[8] Maxwell, D. and van der Vorst, R., 2003, Developing sustainable products and services, J Clean Prod, 11/8: 883-895.

[9] Millet, D., Bistagnino, L., Lanzavecchia, C., Camous, R. and Poldma, T., 2007, Does the potential of the use of LCA match the design team needs?, J Clean Prod, 15/4: 335-346.

[10] Gehin, A., Zwolinski, P. and Brissaud, D., 2007, Towards the use of LCA during the early design phase to define EoL scenarios, CIRP LCE 2007, Tokyo.

[11] Maussang, N., Brissaud, D. and Zwolinski, P., 2007, Common representation of products and services: a necessity for engineering designers to develop Product-Service Systems, CIRP Design 2007, Berlin.

Index of Contributors