Operations Research Ontology for the Integration of Analytic Methods and Transactional Data

Edrisi Muñoz, Elisabet Capón-García, Jose M. Laínez-Aguirre, Antonio Espuña and Luis Puigjaner

Abstract The solution of process systems engineering problems involves their formal representation and application of algorithms and strategies related to several scientific disciplines, such as computer science or operations research. In this work, the domain of operations research is modelled within a semantic representation in order to systematize the application of the available methods and tools to the decision-making processes within organizations. As a result, operations research ontology is created. Such ontology is embedded in a wider framework that contains two additional ontologies, namely, the enterprise ontology project and a mathematical representation, and additionally it communicates with optimization algorithms. The new ontology provides a means for automating the creation of mathematical models based on operations research principles.

Keywords Operations research · Enterprise wide optimization · Decision support systems · Knowledge management

E. Muñoz (✉)
Centro de Investigación en Matemáticas A.C., Jalisco S/N, Mineral y Valenciana,
36240 Guanajuato, Mexico
e-mail: emunoz@cimat.mx

E. Capón-García
Department of Chemistry and Applied Biosciences, ETH Zürich, 8093 Zurich, Switzerland
e-mail: elisabet.capon@chem.ethz.ch

J.M. Laínez-Aguirre
Department of Industrial and Systems Engineering, University at Buffalo,
14260 Amherst, NY, USA
e-mail: jmlainez@gmail.com

A. Espuña · L. Puigjaner
Department of Chemical Engineering, Universitat Politècnica de Catalunya,
Av. Diagonal, 647, E08028 Barcelona, Spain
e-mail: antonio.espuna@upc.edu

L. Puigjaner
e-mail: luis.puigjaner@upc.edu

J. Mejia et al. (eds.), *Trends and Applications in Software Engineering*,
Advances in Intelligent Systems and Computing 405,
DOI 10.1007/978-3-319-26285-7_12

1 Introduction

Process industries are highly complex systems consisting of multiple business and process units, which interact with each other, ranging from molecule to enterprise level. Therefore, in order to solve real world problems, it is necessary to develop algorithms and computational architectures so that large-scale optimization models can be posed and solved effectively and reliably. Hence, the collaboration among different scientific disciplines, namely process systems engineering, operations research and computer science, is highly important [1]. In this sense, enterprise-wide optimization (EWO) is a discipline related to the optimization of supply operations, manufacturing and distribution in a company [1]. A challenge in EWO consists in developing flexible modeling environments for the problem representation, which is the ultimate basis for reaching efficient decision-making. A key feature in EWO is the integration of information and decision-making along the various functions and hierarchical levels of the company's supply chain. Current information technology tools, such as data mining, allow a high degree of information flow in the form of transactional systems (e.g. ERP's) and efforts have been recently devoted to integrate transactional systems and analytical models (e.g. optimization and simulation). These transactional systems must interact with analytical models providing the necessary data to reach appropriate solutions. However, further development is still necessary to easily develop, build and integrate different enterprise models. Such integration should reflect the complex trade-offs and interactions across the components of the enterprise.

This work proposes the creation of an Operations Research (OR) semantic model as a step forward in capturing the nature of problems and technologies for decision making in the enterprise. Specifically, the whole process of decision making and the creation and classification of equations according to their structure are qualitatively represented in terms of OR principles. This OR ontology (ORO) is integrated with other two semantic models previously developed, namely the Enterprise Ontology Project (EOP) [2] and the Ontological Math Representation (OMR) [3], thus enhancing the functionalities of the original ontological framework (Fig. 1). The scope of these models comprises the representation of the real system for EOP, the mathematical representation domain for OMR, and finally the problem design representation for ORO.

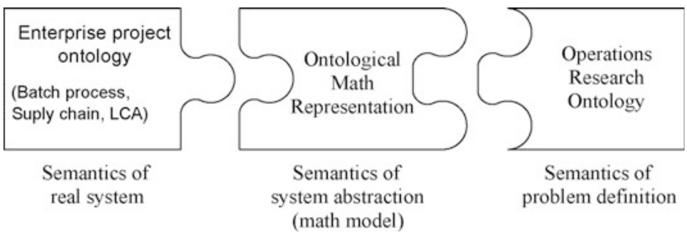

Fig. 1 Relationship among the semantic models with the represented domain

1.1 Operation Research

The Operations Research as science, provides analytical models (technical and mathematical algorithms) to solve decision problems. Operations Research as an art, provides creativity and ability of problem analysis, data collection, model building, validation and implementation for the scientific achievement of problems solution. Besides, the establishment of communication lines between sources of information and analytical models, is a key factor of success.

Thus, the decision making process involves building a decision model and finding the optimal solution. In this way, the problem model is defined as an objective and a set of constraints that are expressed in terms of decision variables of the problem. Decision models allow the identification and systematic evaluation of "all" the possible choices for problem solution. The basic components of an operations research system are decision choices, problem constraints and the objective criteria. The resulting system is an abstraction of the reality which considers the identification of key factors, such as, variables, constraints and parameters. We can reach feasible, unfeasible, optimal and suboptimal solutions, when a problem is solved. The quality of the optimal solution depends on the set of feasible options that are found by the problem modeling and problem solution process. There are different solution techniques such as linear programming (LP), integer programming (IP), dynamic programming (DP), mix integer linear programming (MILP), non linear programming (NLP), and mix integer non linear programming (MINLP) representing algorithms to solve particular classes of operations research models.

Models need to be supported by reliable data in order to provide useful information to the decision maker. Data collection may be the most challenging part of the model to be determined. While modeling experience accumulates, a key factor is the development of means to collect and document data in a useful way. If the data accommodates probabilistic data, then probabilistic or stochastic models are derived.

Basically the process for finding a solution (calculation method) in operations research for mathematical models is iterative in nature. Indeed, the optimal solution is reached by steps/iterations in which each iteration results in feasible and unfeasible solutions.

The phases of a study of operations research are: (i) problem definition; (ii) model building; (iii) model solution; (iv) validation of the model; and, (v) implementation of optimal results. *The problem definition* comprises three main aspects: (a) goal or objective definition; (b) decision alternatives identification; and, (c) identification of the limitations, restrictions and system requirements. *Model building* specifies quantitative expressions which defines the objective and constraints based on system variables. *The model solution* depends on the mathematical technique to be used. Additionally, it is recommended to implement a sensitivity analysis to study the behavior due to changes in the parameters of the system. *The validation of the model* consists of the adequate description of the real system by means of the created model by using historical data. Finally, *the implementation of optimal results* involves translating the results into detailed operating instructions.

2 Operations Research Ontology

Moreover, ontologies are emerging as a key solution to knowledge sharing in a cooperative business environment [4]. Since they can express knowledge (and the relationships in the knowledge) with clear semantics, they are expected to play an important role in forthcoming information-management solutions to improve the information search process [5]. Thus, a development of operations research semantic model which aims to support a formal and systematic process for problems model and design. In order to design a robust and reusable model, a systematic methodology based on a continuous improvement cycle has been followed [6]. This methodology is composed by 5 main steps, which are, plan, do, check, act and re-plan phases. Next, the main characteristics of the above mentioned phases are briefly described.

2.1 Plan Phase

Operation research models support a rigorous problem description using mathematical language (symbols, variables, equations, inequalities, etc.), by structuring the problem as a set of equations. The set of equations represent the objective of the problem and different restrictions and constraints of the system where the problem is found. The proposed semantic model aims to formally represent operations research models for engineering systems, giving special importance to the model construction, structuring and solution steps. Several sources of knowledge have been consulted [7, 8], and the result of the formalization has been validated with engineers currently involved in operations research and its application to the area of process systems engineering.

The resulting semantic model must formalize the design of any operations research problem (Fig. 2), allowing to:

- support the correct structure of operations research, objective function and constraints.
- formalize the meaning of system equations, simple equations and elements within those equations.
- formalize the relationship between system of equations, equations and elements of mathematical models.
- share knowledge with other ontologies for reaching a better system understanding.

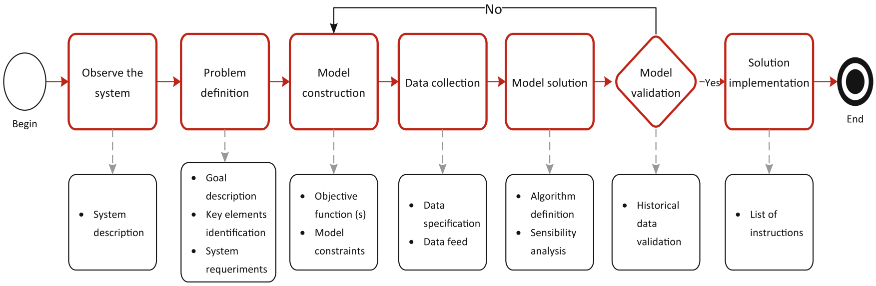

Fig. 2 Flow diagram of the operations research process

- analyze more efficiently the effects of changes in the problem when factors affect the system.
- strengthen decision-making based on operations research modeling.

2.2 Do Phase

The creation of a general operations research modeling ontology was performed using the principal components of the operations research domain gathered from the sources of knowledge and listed in a glossary of terms. Concepts such as Activity, Algorithm, Allocation, Alternatives identification, Assumed real system, Attribute, Break-even analysis, Collection of information, Cost-benefit analysis, Decision trees, Decision-making, Efficiency, Feasible solution, Goal setting, Implementation, Linear programming, Mathematical model, Model, are defined in Table 1, which presents the complete list of terms considered in this work related to the operations research domain.

On the basis of the above presented list, the concepts required to construct the operational research model results in ontological classes which are structured in a taxonomy, as shown in Fig. 3, representing the top classes of the model. These classes are the main building blocks for representing operations research as base for a robust decision making process. On the other hand, some verbs used in the operations research, such as, performs, achieves, has solution relates to, satisfies, establishes, specifies, among others, were used in order to stablish object properties.

The ontology was implemented in Protégé [9] using the Web Ontology Language (OWL). A key point of this model is its connectivity to the system reality. Such connection is achieved by means of the class “Term”, the modeling entity directly related to the engineering domain to be mathematically modeled. Quantitatively speaking, the ontological model contains 72 classes, 52 properties and 9 axioms (see Fig. 3).

Table 1 Basic concepts regarding operations research ontology

Concept	Description
Algorithm	Method to solve a problem through a defined, precise and finite set of instructions
Allocation	The activity of allocate the necessary mount of resources to each activity needed to develop it
Assumed real system	It is an abstraction of the situation resulting from the identification of key factors (variables, constraints and parameters)
Break-even analysis	Analysis of relationships to determine the size or transaction volume reaches a balance between losses and profits
Cost-benefit analysis	A method for finding the best result between benefits and costs, expressed in monetary value
Decision trees	A technique to analyze sequential decisions based on the use of outcomes and associated probabilities
Decision-making	The activity of select certain actions from several options rationally
Efficiency	Achievement of objectives with the least amount of resources
	Achievement of objectives at the lowest cost or other unintended consequences
Feasible solution	A set of values of the problem variables that satisfies the equations
Implementation	It implies the translation of the results into detailed operation instructions
Linear programming	It is mathematical algorithm or procedure by means of which a problem, formulated in terms of lineal equations, is solved optimizing an objective function, also lineal
Mathematical model	Is a model where the objective function and the constraints are expressed in quantitatively or mathematical way
Model	It is an abstraction of the assumed real system that identifies relevant relationships of the system in the form of an objective function and a set of constraints
Model construction	It specifies quantitative expressions for the objective function and the constraints in function of the constraints variables
Model validation	It is presented when the performance of the model is similar to the performance of the real system by using historical data
Operations research	A scientific discipline, which applies theory, methods, and special techniques in order to search the solution to decision making problems related to the conduction and coordination of operations (or activities) within an organization
Optimal solution	A feasible solution that optimizes the objective function
Optimization	Optimization attempts to answer a problem choosing the best among a set of items
Phase	The problem creation is divided in compulsory phases, which are mainly situation examination and information collection
Problem definition	It is defined by three main tasks: (1) Goal description; (2) Identification of the system decision choices; and (3) Identification of limitations, constraints and requirement of the system
Model solution	It depends of the complexity and nature of the model. For the solution some mathematic, simulation or heuristic, or any combination must be

(continued)

Table 1 (continued)

Concept	Description
	used. Even more, some sensitive analysis should be performance in order of study the behavior of the parameters in the system
Risk analysis	Analysis that weighs the risks of a situation to include chances to get a more accurate assessment of the risks involved.
Sensitive analysis	Methodology that allows you to examine the behavior of a result concerning controlled variations of a set of independent variables
	Calculation of the effect of an exogenous variable produces in another variable
Simulation model	It is a model where the system is expressed in basic or elemental modules and integrated bye the “if/then” logic relations
Stochastic model	It is a model where the data is unknown and probabilistic approaches are used in order to obtain data
Suboptimal solution	A feasible solution which is not the optimal solution
System	It is a set of parts or related items organized and interacting to achieve a goal. The systems receive (input) data, energy or environmental matters and provide (output) information, energy or matter
Unfeasible solution	A set of values of the problem variables that do not satisfy the equations

2.3 Check Phase

The language and the conceptuality were checked by experts in the domain to ensure that the model matched the users’ requirements for the design and development of operations research models. Basically, the ontology was checked by members of the research groups specialized in optimization approaches applied to process system engineering (two professors, three associated researchers, two Ph.D. students and two informatics engineers). In addition, the consistency of the model was checked using ontology reasoners, such as, Pellet Incremental and Hermite reasoners from Protégé. The reasoning time for checking the consistency of the classes and of the model was 0.624 and 0.764 CPU’s, respectively, in an intel-core 2 at 2.83 GHz, resulting in a successful compilation. Finally, the checking and validation of the ontological framework consisted in verifying that the operations research models could be successfully instantiated in the ontology. This step comprised a preliminary approach to check the usability of the model.

2.4 Act Phase

This phase comprised all the actions on the model necessary to repair defects and implement suggestions made during the previous phase. Specifically, as an

Fig. 3 Extract of the taxonomy of the operations research ontology

example, the use of the class "Element" was appropriately tuned, in order to specify that the real meaning regards a "Mathematical Element" and unnecessary classes and properties were eliminated after debugging. Finally, the use of the current model and all the formal changes made to the ontological model were recorded in the project documentation.

2.5 Re-planning Phase

The results of this phase are included in the previous subsections, which present the final structure of the pursuit project. However the project has been re-planned several times adding new functionalities to the operations research domain. The whole final semantic model can be inspected and downloaded at http://bari.upc.es/cgi-bin/samba/smbdir.pl?group=&master=&host=neo&share=guest03&dir=\ORO&auth=OT8COzZ,fGcDEionFyQjFg8=.

3 Proposed Framework

In this section, the procedure of the whole framework is presented. As a first step, the actual state of the process is captured by the instantiation of the Enterprise Ontology Project which contains an integrated representation of the enterprise structure, ranging from the supply chain planning to the scheduling function, thus comprising activities related to the operational, tactical and strategic functions. As a next step, by means of the instantiation of the Ontological Math Representation, the various mathematical models (mathematical elements) already established or the design of new ones are translated into a semantic representation as mathematical expressions in order to capture the mathematical meaning of enterprise domain elements. This model relates existing classes belonging to the enterprise domain ontology to mathematical elements to understand and translate the system abstraction in equations. Finally, this work proposes the Operations Research Ontology in order to, on the one hand, formalize and support the processes of: (i) problem abstraction, (ii) analytical model building, (iii) problem solution, (iv) verification and (v) deployment of the best solution; and on the other hand, allow automated representation of the results of this whole process in standardized formats (e.g., the so-called mathematical programming formats MPS and NL).

The workflow diagram of the proposed framework is illustrated in Fig. 4. The first phase aims at reaching a formal conceptualization of the real system of the process industry under study. This step encompasses the standardized semantic description of the system using the enterprise ontology project (EOP), and the definition and acquisition of the required dynamic and static data. The second phase pursues the future formalization of the mathematical equations describing the system abstraction using the ontological mathematical representation (OMR), this

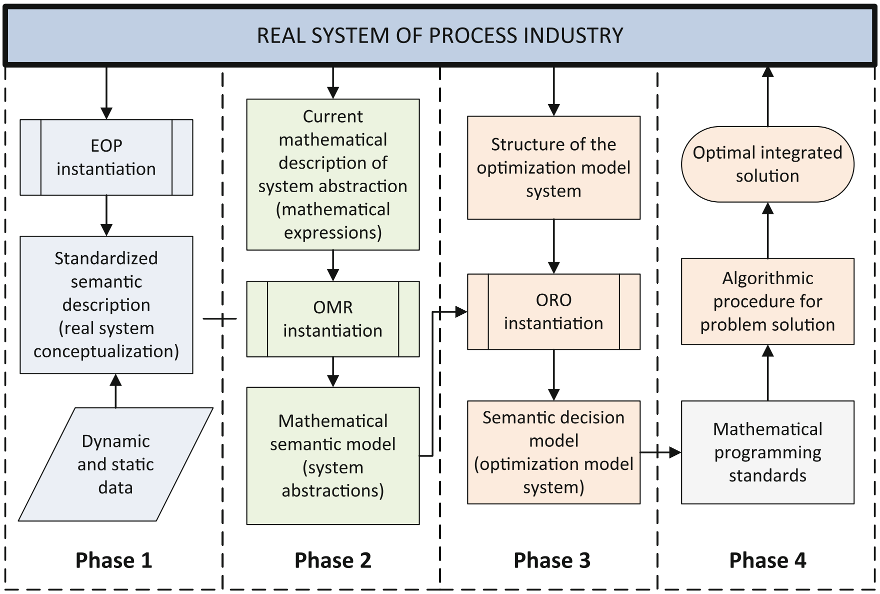

Fig. 4 Flow diagram of the design of problem system construction framework

phase results in a potential mathematical description of the entire system. The OMR can capture both mathematical expressions already in use and new developments for the system conceptualization. Thirdly, the structure of optimization model system along with the mathematical semantic model are the basis for instantiating the operations research ontology (ORO) and obtain a semantic decision model for optimization purposes. ORO has the task of designing the structure and the equation system in order to define a certain problem following the operations research guidelines. Finally, the mathematical programming standards are applied to the semantic decision model and the problem is solved to reach the optimal integrated solution, which assists managers in making the decisions to be deployed in the real system.

4 Case Study

To demonstrate the functionalities of this integrated ontological framework, a case study presented by Muñoz et al. [2] is considered. It consists of an integrated supply chain network planning and scheduling problem. In this work, the processes encompassing the problem definition, the problem formulations, as well as the solution procedure are represented within the "Operations Research Ontology". Therefore, the whole process of operations research has been applied to the problem solving supported by the knowledge management framework.

Table 2 Extract of "Problem definition" in the case study

Goal definition	Key elements identification
Maximize the economic performance of the whole supply chain structure	Direct cost parameters, such as production, handling, transportation, storage, and raw materials
Determine the assignment of manufacturing and distribution tasks to the network node	Indirect expenses, associate with the capacity utilization
Determine the amount of final products to be sold	Prices of the final products in the markets
Determine the amount of transported material among facilities	Set of suppliers with limited capacity provide raw materials to the different production plants
Determine the detailed batching, sequencing and timing of tasks in each production plant	Set of production plants and the distribution centers are located in specific geographical sites and provide final products to the markets
System limitations	System requirements
Mass balances have to be respected	Each production plant produces certain amounts of final products using equipment technologies which have defined installed capacity and minimum utilization rate
The final products are stored in one of the distribution centers before being sent to the markets	Maximum capacity limitations have to be considered for each treatment technology
Each market has a nominal demand of final products along a fixed time horizon	Available transportation links
A certain supply chain network structure	Production routes are defined in the product recipes, which contain mass balance coefficients and the consumption of production resources
	Unfulfilled demand cannot derive in back orders

As a result of the "Observe the system" phase, the case study has been captured using of the Enterprise Ontology Project, which allows to instantiate all the features of the system, from the production process level to the whole supply chain network. The following phase concerning "Problem definition" aims to capture four key issues: (i) goal description, (ii) key elements identification, (iii) system limitations, and (iv) system requirements (Table 2).

The "Model construction" has been derived from the "Problem definition", which provides the elements for the model creation. Specifically, links among the four different issues identified in the "Problem definition" are established by means of a relation matrix, and the sets, parameters, variables and groups of equations are accordingly derived. For example, the goal "Determine the amount of final products to be sold" is related to the continuous variable SLsff’t, which represents the sales of product s at time period t produced in facility f to market f’. Thus, the semantic modeling of the model is supported by the "Ontological Mathematical Representation", which is also related to the "Enterprise Ontology Project".

The "Data collection" stage concerns the relation of the different sets and parameters of the specific problem instantiation to their current value, which are usually stored in organized databases. Thus, the "Model solution" considers the algorithm selection according to the specific features of the mathematical model, and proceeds to the problem solution. Finally, the steps "Model validation" and "Implementation" consider the validation of the model according to historical data and decision-maker expertise, and the acceptance and application of the resulting solution for the real system.

5 Conclusions

The ontological framework provides a tool to build computational optimization models for the enterprise decision-making and to allow the comprehensive application of enterprise wide optimization throughout the process industry. This framework encompasses the steps of OR, and communicates with two previously existing semantic models related to enterprise wide and mathematical representation. This extended framework also allows the building of more accurate models for the chemical process industry and the full integration and solution of large-scale optimization models. The main contribution of this work is the systematic and rigorous representation of the whole decision-making process from the problem conception to implementation. As a result, creation and re-use of analytical models in the industry can be semantically supported, providing a higher flexibility and integration for model building of enterprise operations. Finally, Operations Research Ontology aims to support problem design and mathematical formulations (modeling & programming) for decision making in order to reach EWO by standardizing and facilitating the decision making process based on operations research, creating and classifying systems of mathematical equations according to the system's specification and decision type and reusing knowledge from other ontology models (Enterprise Project and Ontological Math Representation).

Acknowledgments Authors would like to acknowledge the Spanish Ministerio de Economía y Competitividad and the European Regional Development Fund for supporting the present research by projects EHMAN (DPI2009-09386) and SIGERA (DPI2012-37154-C02-01). Finally the financial support received from CIMAT México is also fully acknowledged.

References

1. Grossmann, I.E.: Enterprise-wide optimization: A new frontier in process systems engineering. AIChE J. **51**, 1846–1857 (2005)
2. Muñoz, E., Capón García, E., Laínez, J.M., Espuña, A., Puigjaner, L.: Integration of enterprise levels based on an ontological framework. Chem. Eng. Res. Des. **91**, 1542–1556 (2013)

3. Muñoz, E., Capón García, E., Laínez, J.M., Espuña, A., Puigjaner, L.: Mathematical knowledge management for enterprise decision making. In: Kraslawski, A., Turunen, I. (eds.) Computer Aided Design: Proceedings of the 23rd European Symposium on Computer Aided Chemical Engineering, Lappeeranta Finland, pp. 637–642 (2012)
4. Missikoff, M., Taglino, F.: Business and enterprise ontology management with symontox. In: Heidelberg, S.B. (ed.) The Semantic Web—ISWC 2002, 2342/2002, pp. 442–447. (2002)
5. Gruber, T.R.: A translation approach to portable ontology specifications. Knowl. Acquis. **5**(2), 199–220 (1993)
6. Muñoz, E.: Knowledge management tool for integrated decision-making in industries PhD Thesis, Universitat Politècnica de Catalunya Poner Barcelona, Spain (2011)
7. Ravindran, A.R.: Operations Research and Management Science Handbook. CRC Press Taylor and Francis Group, New York (2007)
8. Taha, H.A.: Operations Research. Prentice Hall, New Delhi (1995)
9. Horridge, M., Jupp, S., Moulton, G., Rector, A., Stevens, R., Wroe, C.: A practicalguide to building owl ontologies using protege 4 and CO-ODE tools. Technical report. The University of Manchester (2007)

Further Reading

10. Smith, T.F., Waterman, M.S.: Identification of common molecular subsequences. J. Mol. Biol. **147**, 195–197 (1981)