STUDIES IN FUZZINESS AND SOFT COMPUTING

Weldon A. Lodwick
Janusz Kacprzyk (Eds.)

Fuzzy Optimization

Springer

Fuzzy Optimization

Studies in Fuzziness and Soft Computing, Volume 254

Editor-in-Chief

Prof. Janusz Kacprzyk
Systems Research Institute
Polish Academy of Sciences
ul. Newelska 6
01-447 Warsaw
Poland
E-mail: kacprzyk@ibspan.waw.pl

Further volumes of this series can be found on our homepage: springer.com

Vol. 237. Kofi Kissi Dompere
Fuzziness and Approximate Reasoning, 2009
ISBN 978-3-540-88086-8

Vol. 238. Atanu Sengupta, Tapan Kumar Pal
Fuzzy Preference Ordering of Interval Numbers in Decision Problems, 2009
ISBN 978-3-540-89914-3

Vol. 239. Baoding Liu
Theory and Practice of Uncertain Programming, 2009
ISBN 978-3-540-89483-4

Vol. 240. Asli Celikyilmaz, I. Burhan Türksen
Modeling Uncertainty with Fuzzy Logic, 2009
ISBN 978-3-540-89923-5

Vol. 241. Jacek Kluska
Analytical Methods in Fuzzy Modeling and Control, 2009
ISBN 978-3-540-89926-6

Vol. 242. Yaochu Jin, Lipo Wang
Fuzzy Systems in Bioinformatics and Computational Biology, 2009
ISBN 978-3-540-89967-9

Vol. 243. Rudolf Seising (Ed.)
Views on Fuzzy Sets and Systems from Different Perspectives, 2009
ISBN 978-3-540-93801-9

Vol. 244. Xiaodong Liu and Witold Pedrycz
Axiomatic Fuzzy Set Theory and Its Applications, 2009
ISBN 978-3-642-00401-8

Vol. 245. Xuzhu Wang, Da Ruan, Etienne E. Kerre
Mathematics of Fuzziness – Basic Issues, 2009
ISBN 978-3-540-78310-7

Vol. 246. Piedad Brox, Iluminada Castillo, Santiago Sánchez Solano
Fuzzy Logic-Based Algorithms for Video De-Interlacing, 2010
ISBN 978-3-642-10694-1

Vol. 247. Michael Glykas
Fuzzy Cognitive Maps, 2010
ISBN 978-3-642-03219-6

Vol. 248. Bing-Yuan Cao
Optimal Models and Methods with Fuzzy Quantities, 2010
ISBN 978-3-642-10710-8

Vol. 249. Bernadette Bouchon-Meunier, Luis Magdalena, Manuel Ojeda-Aciego, José-Luis Verdegay, Ronald R. Yager (Eds.)
Foundations of Reasoning under Uncertainty, 2010
ISBN 978-3-642-10726-9

Vol. 250. Xiaoxia Huang
Portfolio Analysis, 2010
ISBN 978-3-642-11213-3

Vol. 251. George A. Anastassiou
Fuzzy Mathematics: Approximation Theory, 2010
ISBN 978-3-642-11219-5

Vol. 252. Cengiz Kahraman, Mesut Yavuz (Eds.)
Production Engineering and Management under Fuzziness, 2010
ISBN 978-3-642-12051-0

Vol. 253. Badredine Arfi
Linguistic Fuzzy Logic Methods in Social Sciences, 2010
ISBN 978-3-642-13342-8

Vol. 254. Weldon A. Lodwick, Janusz Kacprzyk (Eds.)
Fuzzy Optimization, 2010
ISBN 978-3-642-13934-5

Weldon A. Lodwick and Janusz Kacprzyk
(Eds.)

Fuzzy Optimization

Recent Advances and Applications

Springer

Editors

Prof. Weldon A. Lodwick
University of Colorado Denver
Department of Mathemaitical and Statistical Sciences
PO Box 173364 Denver
Colorado 80217-3364
USA
E-mail: Weldon.Lodwick@ucdenver.edu

Prof. Janusz Kacprzyk
Systems Research Institute
Polish Academy of Sciences
ul. Newelska 6
01-447 Warsaw
Poland
E-mail: kacprzyk@ibspqan.waw.pl

ISBN 978-3-642-13934-5 e-ISBN 978-3-642-13935-2

DOI 10.1007/978-3-642-13935-2

Studies in Fuzziness and Soft Computing ISSN 1434-9922

Library of Congress Control Number: 2010928592

Typeset & Cover Design: Scientific Publishing Services Pvt. Ltd., Chennai, India.

Printed on acid-free paper

9 8 7 6 5 4 3 2 1

springer.com

Foreword

Optimisation and decision making has always been one of the main focuses of operations research, decision theory, management science and economic theory. Until the beginning of the 1970's the mathematical tools used were predominantly "crisp" and the rationality of decisions in decision logic was defined on the basis of dual logic. If uncertainty was included at all in decision calculi or optimisation algorithms, it was done by using probability theory, for instance, in stochastic linear programming. The basis for any kind of analysis in decision theory and operations research was the asymmetric choice model in which decision alternatives (or feasible solutions) were modelled either by enumeration or by constraints as an unordered solution space and in which an order in this space was generated by the objective function, utility function, or any other order generating tool.

This had three major consequences:

1. Models were rather unstable because a minimal violation of a single constraint led to infeasibility,
2. All considerations of uncertainty had to be cast into probabilistic models, and
3. If more than one objective function was to be considered the problem became complicated due to the several orders that were generated in the solution space.

These limitations also often reduced the degree to which the model adequately modelled the real problem.

It should probably also be mentioned that the "knowledge-based" systems (expert systems), that were emerging as an alternative to mathematical optimisation at the beginning of the 1970's as a way to find not optimal but good heuristic solutions, were also based on dual logic and were, therefore, not really "knowledge based" but rather symbol (truth values) processing systems.

Many things have changed since then: The phenomenon of uncertainty has drawn much more attention and several "uncertainty theories" have been developed. (There exist more than twenty by now). Fuzzy sets theory has grown tremendously which led to more than 50 000 publications by now. Multi criteria decision making (MCDM) has become a very large area with many different approaches to handle the problem of several objective functions. In addition to the

asymmetric choice model of classical decision logic the symmetrical decision model of fuzzy set theory and MCDM has been suggested. The rationality of decision making has been reconsidered on the basis of fuzzy logic. Several new "bio-inspired" areas (such as artificial neural networks, evolutionary computation, swarm theory and others) have also been developed. Partly they have been pooled into "Computational Intelligence" or into "Soft Computing" and generated new hybrid models and theories. As one reaction to uncertainty in real applications the demand for "robust solutions" has appeared besides the request for optimal solutions. Also one major change in the relationship between these theories has occurred: to a large degree they are no longer considered as competitors but rather as complements which has even more strengthened the development of hybrid theories and techniques.

Publications in these areas are distributed over a very wide range of journals and books, making it very difficult, even for a scientist, to stay up-to-date. It is even more difficult for a practitioner to oversee which new tools, models, and techniques are available to solve his or her optimisation problems as adequately as possible. It is, therefore, extremely valuable that a book appears that surveys as many of these new possibilities as one book can cover. It not only surveys these theories but it also investigates the relationships between these theories and a very important question: "Which theory should be applied, under which circumstances, to real problems?"

The editors of this book have done an excellent job in inviting and getting the participation of top authors working in theory as well as in applications. In order to make this volume more readable it has been structured into a number of parts, which are not independent of each other but complement each other very well: Part 1 leads the reader from classical decision theory and optimisation into the area of fuzziness and fuzzy optimisation. Even though the main focus of this volume is fuzzy optimisation, other uncertainty theories, such as the Dempster-Shafer theory, possibility theory, etc. are also introduced and their interrelationships discussed. Part 2 is devoted to a very central problem of optimisation, the aggregation of the different parts of optimisation models, i.e. constraints and objective functions. This, from the very beginning of fuzzy decision making theory, has been one of the most interesting areas. Part 3 then turns to developments in different areas of fuzzy optimisation. It is not surprising that a major part is devoted to variations of mathematical programming. While classical linear programming, or even the different kinds of non-linear programming, are characterized by specific clear-cut models, this changes immediately when fuzzy set theory (or other similar uncertainty theories) are applied. Thus, while crisp linear programming is well described by one typical model, fuzzy linear programming is a family of very different models, depending on the membership functions used, the type of aggregation operators introduced and the degrees to which a crisp linear programming model is "fuzzified". Even the solvability of such a fuzzy linear programming model can vary from being easy to solve (as an equivalent crisp linear program) to unsolvable. In addition to those families of fuzzy mathematical programming models there are, of course, various other fuzzy optimisation models, which are also investigated in this part of the book.

One part of (linear) mathematical programming has gained particular importance during the last decade: combinatorial programming. It is, therefore, justified that the next part of the book is devoted to fuzzified versions of this type of linear programming. Three rather different models are presented, which indicate the wide and not yet well-developed area of these models.

The final Part 5 of the volume is devoted to applications of theories considered in the first four parts to models of real problems. The seven examples cover a wide scope of areas from feature selection via investment and portfolio selection problems to biological circuit design and a real problem of the Spanish football league. These examples can certainly serve as an inspiration to numerous other applications in the real world.

Altogether this book is an excellent piece of theoretical and applied literature. The editors and the authors have to be congratulated to their work. I am sure that this book will be of great benefit to scientists and practitioners and I can only hope that it will lead to many further developments in the area of fuzzy optimisation.

Aachen Hans-Jürgen Zimmermann
February 2010

Preface

Optimization is an extremely important area in science and technology which provides powerful and useful tools and techniques for the formulation and solution of a multitude of problems in which we wish, or need, to find a best possible option or solution. It has been an important area of research for more than half a century, and particularly since the advent of digital computers. Over those years great progress has been attained in the area with the development of powerful theoretical and algorithmic results. A multitude of academic and commercial software packages have been developed which have made it possible to solve virtually all kinds of optimization problems. Applications of optimization tools and techniques span practically the entire spectrum of science and technology.

Real applications of optimization often contain information and data that is imperfect. Thus, attempts have been made since the early days to develop optimization models for handling such cases. As the first, natural approaches in this respect one can mention value intervals and probability distributions as representations of uncertain data. They have led to the development of various interval and stochastic optimization models.

Fuzzy sets theory has provided conceptually powerful and constructive tools and techniques to handle another aspect of imperfect information related to vagueness and imprecision. This has resulted in the emergence – more or less in the mid-1970s—of a new field, called fuzzy optimization (and its related fuzzy mathematical programming), in which many powerful theoretical and algorithmic results have been proposed too. Many books and edited volumes, and a multitude of articles have been published. Moreover, numerous applications have been reported too.

Due to the importance and a constant growth of interest, both among theoreticians and practitioners, we have decided to prepare this edited volume on fuzzy optimization. A substantial number of the most active researchers and practitioners in the field have responded positively to our application, and therefore we have been able to present to the readers a comprehensive account of many new and relevant developments in fuzzy optimization, in its theoretical direction and also in real world applications.

The volume is divided into a coupe of parts which present various aspects of fuzzy optimization, some related more general issues, and applications.

Part I, "Introductory Sections", is concerned with a comprehensive survey of some basic issues related to broadly perceived choice, decision making, and optimization. Aspects of utility, preferences, uncertain and imprecise information are discussed. Moreover, an account of how these aspects can be incorporated in operational optimization models are surveyed.

Kofi K. Dompere ("Fuzziness, Rationality, Optimality and Equilibrium in Decision and Economic Theories") discusses some basic issues related to decision making and optimization, and puts them in a perspective of fuzziness. The paper is an essay which presents main categories of theories of optimization. It begins with the classical system leading to the establishment of a point of departure for fuzzy optimization from the point of view of classical optimization. The author distinguishes the following four categories of optimization problems: the first two, i.e. exact (non-fuzzy) and non-stochastic, and exact (non-fuzzy) and stochastic follow somehow classical laws of thought and mathematics. On the other hand, the other two, i.e. fuzzy and non-stochastic, and fuzzy-stochastic problems are associated with laws of thought characteristic for fuzzy logic and mathematics. From these structures, similarities and differences in the problem structures and corresponding solutions are abstracted and discussed. They are attributed to properties of exactness and completeness about information-knowledge structures underlying the optimization problems. The assumed degrees of exactness and completeness establish defective information-knowledge structure that generates uncertainties and produces inter-category differences in the optimization problem. The differences of intra-category algorithms are attributed to differences in the assumed functional relationships of the variables that occur in the objective and constraint sets. A taxonomy of solution structures is provided and a discussion on future research directions is given.

W.A. Lodwick and E. Untiedt ("Introduction to Fuzzy and Possibilistic Optimization") provide a comprehensive introduction to various aspects of broadly perceived fuzzy and possibilistic optimization. However, as opposed to previous survey articles of that type, they go much deeper, providing insight into decision making, optimization and mathematical programming in general. First, they give an overview of various perspectives, points of view, on uncertain, imprecise, and incomplete, information and summarize various mathematical modeling attempts and algorithms. Differences between representations based on interval mathematics, probability theory, the Dempster-Shafer theory, fuzzy sets theory, possibility theory, and related issues are considered. Problems related to the choice of optimal (best) options or courses of action are mentioned. Finally, a review of rationale, fundamental features and solution techniques for basic classes of fuzzy optimization and fuzzy mathematical programming are analyzed.

Part II, "Basic Issues", is concerned with some foundational issues that are relevant for fuzzy optimization, both in the theoretical and algorithmic sense. The main concern in this part is an omnipresent problem of aggregation of partial scores, pieces of evidence, preferences, etc.

Vicenç Torra ("Aggregation Operators for Evaluating Alternatives") reviews the use of aggregation functions and operators in the field of decision making, and hence in optimization as a consequence. The author first presents an overview of

main decision making problems, and then shows that aggregation operators are commonly employed for their solution. Then, a comprehensive review of various aggregation operators is provided, and their formal properties, features and differences are pointed out.

Gleb Beliakov ("Optimization and Aggregation Functions") looks at connections between aggregation functions and optimization from two main perspectives. The insight into the connections is that of aggregation functions which are used to transform a multiobjective optimization problem into a single objective problem by combining several criteria into one. The second insight into connections of aggregation functions and optimization is that the construction of aggregation functions often involves an optimization problem. Aggregation functions, or operators, are functions that combine several input values into one output value which can be used to evaluate or rank the alternatives. The author concentrates on aggregation functions that take the inputs from a closed interval, like [0,1], and produce the output in the same interval; they are widely used, in virtually all areas. Since the choice of an aggregation function is application specific, and is frequently performed in ad hoc manner, there are natural attempts to try to automate their choice, in particular when data are available from which information needed can be extracted. This can be exemplified by an analysis of customers' responses to recommendations which can provide suitable aggregation rules. It is possible to construct suitable application specific aggregation functions from the recorded data by solving a regression problem, which for the weighted mean operators boils down to a standard quadratic programming, though for other aggregation functions, the solution may be much more difficult. In this contribution the author presents various alternative methods suitable for the construction of aggregation functions.

Pingke Li and Shu-Cherng Fang ("Chebyshev Approximation of Inconsistent Fuzzy Relational Equations with Max-T Composition") consider an important problem associated with fuzzy relational equations which are a powerful tool for the formulation of many problems. The authors deal with resolving the inconsistency of a system of fuzzy relational equations with the max-T composition by simultaneously modifying the coefficient matrix and the right hand side vector. They show that resolving the inconsistency of fuzzy relational equations with the max-T composition by means of the Chebyshev approximation is closely related to the generalized solvability of interval-valued fuzzy relational equations with the max-T composition. An efficient procedure is proposed to obtain a consistent system with the smallest perturbation in the sense of the Chebyshev distance.

Part III, "Various Types of Fuzzy Optimization and Fuzzy Mathematical Programming Models", is devoted to a comprehensive presentation of some important classes of fuzzy optimization and fuzzy mathematical programming problems that are relevant both from the theoretical and practical points of view.

Ricardo C. Silva, Carlos Cruz, José L. Verdegay and Akebo Yamakami ("A Survey of Fuzzy Convex Programming Models") consider some basic issues related to convex optimization which is characterized by a convex objective function and convex constraint functions over a convex set of the decision variables. This can be viewed, on the one hand, as a particular case of nonlinear

programming and, on the other hand, as a general case of linear programming. Since in many cases when we use convex optimization, we deal with data that cannot be formulated precisely, then it makes sense to apply fuzzy set theory as a way to mathematically describe this imperfect information. In this paper the authors review the theory of fuzzy convex optimization and describe some flexible and possibilistic programming models to solve fuzzy convex programming problems. Flexible programming uses fuzzy sets to represent the imprecisely specified decision maker's aspirations and constraints, while possibilistic programming models handle imprecise or ambiguous data by possibility distributions.

Masahiro Inuiguchi ("Approaches to Linear Programming Problems with Interactive Fuzzy Numbers") considers the following crucial problem. Though most fuzzy mathematical programming models have been developed under the assumption of non-interaction among fuzzy coefficients, this is not always, maybe rarely, the case in real world problems. Therefore, several approaches have been proposed to deal with the interaction among fuzzy coefficients. The author provides a comprehensive and critical review of how the interaction among fuzzy coefficients in fuzzy linear programming problems can be dealt with. Using a necessity fractile model of a simple linear program with fuzzy coefficients, he shows differences between the non-interactive and interactive problems. Then, a review of five approaches to interactive fuzzy numbers, i.e., weak independent fuzzy numbers, fuzzy vector with a quadratic membership function, scenario decomposed fuzzy numbers, an oblique fuzzy vector, and a fuzzy polytope is provided.

Alexander Yazenin and Ilia Soldatenko ("Possibilistic Optimization Tasks with Mutually *T*-related Parameters: Solution Methods and Comparative Analysis") consider the problems of possibilistic linear programming. The T-norms are used to describe the interaction (relatedness) of fuzzy parameters. Solution methods are proposed, models of possibilistic optimization are compared for different T-norms. Basically, in traditional works the relatedness of fuzzy parameters in possibilistic optimization problems was based generally on the standard conjunction operation that is widely used in fuzzy logic which may often be not quite appropriate. For example, using the standard conjuction operator leads to the linear growth of result's fuzziness, which is not always reasonable. The methods based on T-norms provide more flexibility in controlling fuzziness in decision-making. The authors follow this line of investigation with regard to possibilistic linear programming tasks. For the case of TW-norm, they study two models of possibilistic linear programming problems, and propose methods which are combinations of the indirect method and genetic algorithms use for their solution. A comparison of models of possibilistic optimization for the TW-norm and TM-norm is given.

Elizabeth Untiedt ("A Parametrized Model for Optimization with Mixed Fuzzy and Possibilistic Uncertainty") considers the problem when fuzzy and possibilistic uncertainty, which very closely related, and sometimes coexist in optimization under uncertainty problems. Basically fuzzy uncertainty in mathematical programming problems typically represents flexibility on the part of the decision make while possibilistic uncertainty generally expresses a lack of information about the values the parameters will assume. First, the author briefly surveys several existing models for mixed fuzzy and possibilistic programming problems and indicates

that the semantic interpretation of these models may be of questionable value. Namely, the mixed models in the literature find solutions in which the fuzzy uncertainty (or flexibility) and the possibilistic uncertainty (or lack of confidence in the outcome) are held to the same levels. The author proposes a new mixed model which allows a trade-off between fuzzy and possibilistic uncertainty and this trade-off corresponds to a semantic interpretations consistent with human decision-making. The new model shares characteristics with multi-objective programming and the Markowitz models, and its structure, semantic justification, and solution approaches are articulated.

Włodzimierz Ogryczak and Tomasz Śliwiński ("On Solving Optimization Problems with Ordered Average Criteria and Constraints") discuss the problem of aggregating multiple numerical attributes to form an overall measure of broadly perceived performance or utility. The use of Yager's ordered weighted averaging (OWA) aggregation, which use the weights assigned to the ordered values rather than to the specific attributes, makes it possible to model various aggregation preferences, preserving simultaneously the impartiality (neutrality) with respect to the individual attributes. However, the more general importance weighted averaging is a central task in multiattribute decision problems of many kinds, and can be performed by the Weighted OWA (WOWA) aggregation though the importance weights make the WOWA concept much more complicated than the original OWA. The authors analyze some novel solution procedures for optimization problems with the ordered average objective functions or constraints, and show that the WOWA aggregation with monotonic preferential weights can be reformulated in a way that makes it possible to introduce linear programming models, similar to the optimization models developed earlier by the authors for the OWA aggregation. Numerical results justify the computational efficiency of the proposed models.

Gia Sirbiladze ("Fuzzy Dynamic Programming Problem for Extremal Fuzzy Dynamic System") deals some problem related to the so-called Extremal Fuzzy Continuous Dynamic System (EFCDS) optimization developed by the author. The basic properties of extended extremal fuzzy measure are considered and several variants of their representation are given. For extremal fuzzy measures several transformation theorems are represented for extended lower and upper Sugeno integrals. Values of extended extremal conditional fuzzy measures are defined as a levels of expert knowledge reflections of EFCDS states in the fuzzy time intervals. The notions of extremal fuzzy time moments and intervals are introduced and their monotone algebraic structures that form the most important part of the fuzzy instrument of modeling extremal fuzzy dynamic systems are discussed. Some new approaches in modeling of EFCDS are developed, and fuzzy processes with possibilistic uncertainty, the source of which is extremal fuzzy time intervals, are constructed. Dynamics of EFCDS's is described, and the ergodicity of EFCDS's is considered. Fuzzy-integral representations of controllable extremal fuzzy processes are given. Sufficient and necessary conditions are presented for the existence of an extremal fuzzy optimal control processes using Bellman's optimality principle and taking into account the gain-loss fuzzy process. A separate consideration is given to the case where an extremal fuzzy control process acting on the EFCDS does not depend on an EFCDS state. Applying Bellman's optimality

principle and assuming that the gain-loss process exists for the EFCDS, a variant of the fuzzy integral representation of an optimal control is given for the EFCDS. This variant employs the extended extremal fuzzy composition measures. An example of how to construct the EFCDS optimal control is presented.

Milan Mareš ("Vaguely Motivated Cooperation") considers the transferable utility cooperative games which are used as an effective mathematical representation of cooperation and coalitions forming. The author discusses a modified form of such games in which the expected pay-offs of coalitions are known only vaguely, where the vagueness is modeled by means of fuzzy quantities and some other fuzzy set theoretical concepts. Then, for such games the author discusses an extension of their cores and Shapley values, as well as some other properties, from the point of view of the motivation of players to cooperate in coalitions, as well as the relation between the willingness to cooperate and the ability to find the conditions under that the cooperation can be perceived as fair. The usefulness of some fuzzy and possibilistic optimization type tools is indicated.

Part IV, "Fuzzy Network and Combinatorial Optimization", is mainly concerned with broadly perceived fuzzy integer programming, or – more generally – broadly perceived fuzzy combinatorial optimization models, notably those related to network optimization.

Adam Kasperski and Paweł Zieliński ("Computing min-max Regret Solutions in Possibilistic Combinatorial Optimization Problems") discuss a wide class of combinatorial optimization problems with a linear sum and a bottleneck cost function. First, the authors consider the case when the weights in the problem are modeled as closed intervals, and show how the concept of optimality can be extended by using the concept of a deviation interval. For choosing a solution to the problem considered, the authors adopt a robust approach by seeking a solution that minimizes the maximal regret, that is, the maximal deviation from the optimum over all weight realizations, called scenarios, which may occur. Then, they explore the case in which the weights are specified as fuzzy intervals and show that under the fuzzy weights the problem has an interpretation which is consistent with possibility theory. Namely, the fuzzy weights induce a possibility distribution over the set of scenarios and the possibility and necessity measures can be used to extend the optimality evaluation and the min-max regret approach.

Yue Ge and Hiroaki Ishii ("Stochastic Bottleneck Spanning Tree Problem on a Fuzzy Network") consider a fuzzy network version of the stochastic bottleneck spanning tree problem. The existence of each edge is not necessary certain and it is given by a certain value between 0 and 1, with 1 standing for that it exists certainly and 0 for that it does not exist. For intermediate numbers, a higher value corresponds to a higher possibility of existence. Furthermore each edge has a random cost independent to other edges. The probability that the maximum burden among these selected edges is not greater than the capacity should be not less than the fixed probability. In this setting, the authors look for a spanning tree minimizing the capacity and maximizing the minimal existence possibility among these selected edges. Since usually there is no spanning tree optimizing simultaneously these two objectives, the authors develop an efficient solution procedure to obtain a set of some non-dominated spanning trees.

Dorota Kuchta ("The Use of Fuzzy Numbers in Practical Project Planning and Control") proposes how to use fuzzy numbers in project planning and control in such a way that it would meet requirements and expectations of practitioners. The method proposed is fairly general and is meant for all the projects, but especially for those where in the initial phase knowledge about the project is very incomplete and is made stepwise more precise during the project execution, and also for those in which initial assumptions about the project execution times are due to later changes. The method proposed requires the users to think while estimating project parameters in terms of trapezoidal fuzzy numbers, which in fact means only giving four parameters: an optimistic one, a pessimistic one and one or the two medium ones, which may also be equal to each other. The approach requires in each control moment, not an automatic generation of numbers which do not take into account the really important information about the project history and its future, but a deeper insight into the development of the project, the influence of its environment and the interdependencies between various project elements (activities, resources etc.).

Part V, "Applications", presents some examples of successful applications of broadly perceived fuzzy optimization and fuzzy mathematical programming in diverse areas, from economic and management, through technological to biological problems.

Susana M. Vieira, João M. C. Sousa and Uzay Kaymak ("Ant Feature Selection Using Fuzzy Decision Functions") consider feature selection, one of the most important stages in data preprocessing for data mining. Real-world data analysis, data mining, classification and modeling problems usually involve a large number of candidate inputs or features, and less relevant or highly correlated features decrease in general the classification accuracy, and enlarge the complexity of the classifier. Basically, feature selection is a multi-criteria optimization problem with contradictory objectives which are difficult to properly describe by conventional cost functions. The authors propose the use of fuzzy optimization to improve the performance of this type of system, since it allows for an easier and more transparent description of the criteria used in the feature selection process. This paper is an extension of the authors' previous work in which an ant colony optimization algorithm for feature selection was proposed which minimized two objectives: the number of features and classification error. Now, in this chapter, the authors propose a fuzzy objective function to cope with the difficulty of weighting the different criteria involved in the optimization algorithm. They show an application of fuzzy feature selection to two benchmark problems that justify the usefulness of the proposed approach.

Hiroshi Tsuda and Seiji Saito ("Application of Fuzzy Theory to the Investment Decision Process") propose a new approach to portfolio optimization that allows portfolio managers to construct portfolios that reflect their views about risk assets by applying fuzzy sets theory. The proposed approach to the investment decision process is based on the mean-variance approach proposed by Markowitz and uses the concept of asset market equilibrium proposed by Sharpe. For portfolio managers, it is very meaningful to use the equilibrium expected excess returns associated with the capital market as a reference. The proposed approach enables a new

method for incorporating the views of portfolio managers to aid in the investment decision process. Moreover, in order to estimate the distribution of an unknown true membership function of the views of portfolio managers concerning risk assets, the authors propose a fuzzy information criterion to evaluate the fitness of the distribution between an unknown true membership function and a hypothetical membership function. In particular, the proposed approach enables a group of portfolio managers of pension funds to obtain an important solution that realizes optimal expected excess returns of risky assets by specifying the vague views of portfolio managers as a fuzzy number.

Anna M. Gil-Lafuente, José M. Merigó ("Decision Making Techniques in Political Management") develop a new decision making model meant for selecting the best governmental policy of different types such as fiscal, monetary and commercial, and the authors employ a framework based on the use of ideals in the decision process and several similarity measures. For each similarity measure, different aggregation operators are applied exemplified by the simple and weighted average, the ordered weighted averaging (OWA) operator and its generalizations. Basically, the approach deals with multiple attributes and different scenarios for the selection of policies arising in various institutions. The authors develop different techniques using as a starting point a selection process based on attributes under the assumption that the requirements for each attribute is different depending on the environment of the economy.

Takashi Hasuike and Hiroaki Ishii ("Mathematical Approaches for Fuzzy Portfolio Selection Problems with Normal Mixture Distributions") consider some versatile portfolio selection models with general normal mixture distributions and fuzzy or interval numbers. They develop some fuzzy optimization models to obtain an optimal portfolio. Basically, they formulate the proposed portfolio selection problems minimizing the total variance and maximizing the total future return with normal mixture distributions, respectively. They introduce uncertainty sets for the mean values, weights and probabilities as fuzzy numbers. Taking into account several portfolio selection problems including randomness and fuzziness, the authors construct a novel solution method. The results obtained are compared on numerical examples with standard approaches, and some advantages of the approach proposed are pointed out.

Shuming Wang and Junzo Watada ("Fuzzy Random Redundancy Allocation Problems") consider some relevant problems in reliability related to the fuzzy random parallel systems. Namely, due to subjective judgment, imprecise human knowledge and perception in capturing statistical data, the real data of lifetimes in many systems are both random and fuzzy in nature. Based on the fuzzy random variables that are used to characterize the lifetimes, the authors study the redundancy allocation problems to a fuzzy random parallel-series system. Two fuzzy random redundancy allocation models (FR-RAM) are developed through reliability maximization and cost minimization, respectively. Some properties of the FR-RAM are obtained, where an analytical formula of reliability with convex lifetimes is derived and the sensitivity of the reliability is discussed. To solve the FR-RAMs, the authors first address the computation of reliability. A random simulation method based on the derived analytical formula is proposed to compute the

reliability with convex lifetimes. As for the reliability with non-convex lifetimes, the technique of fuzzy random simulation together with the discretization method of fuzzy random variable is employed to compute the reliability, and a convergence theorem of the fuzzy random simulation is proved. This fuzzy approach is then combined with the use of a genetic algorithm (GA) to search for the approximately optimal redundancy allocation of the models. Numerical examples provided illustrate the performance of the solution algorithm.

Eva Sciacca and Salvatore Spinella ("Reliable Biological Circuit Design Including Uncertain Kinetic Parameters") deal with biological design problems which should be particularly important in the near future when it will be possible to produce biological entities and synthetic organisms for pharmacological and medical usage. The biological systems are considered in terms of performance or key features of the system. The idea adopted is that the set of parameters involved in the model can be classified into two different typologies: the uncertain kinetic parameters and the control design parameters. In order to design a robust and reliable biological system with respect to a target performance, the design parameter values are set up to balance the uncertainty of the kinetic parameters. To take into account these uncertainties arising from the estimations of the kinetic parameters, the function representing feedback is fuzzified and a measure of failure of the designed biological circuit is minimized to reach the required performance. For illustration, a case study of an autonomously oscillatory system is provided, namely the Drosophila Period Protein which is a central component of the Drosophila circadian clocks. The results compared with a deterministic method and advantages are shown.

Zach Richards ("Fuzzy Optimal Algorithms for Multiple Target Convergence") proposes the use of fuzzy algorithms for a networked swarm of autonomous vehicles, such as those used in planet exploration, and to be used in target location determination and convergence. In particular, an algorithm of this type could be used in an Autonomous Stratospheric Aircraft (ASA), thus having the possibility of being used for the exploration of a planet as well as many other space, military and civil applications. Upon finding an unknown location of a specified target, the algorithm would then swarm and eventually converge upon the location. The author proposes two similar, but fundamentally different algorithms which are capable of locating and converging upon multiple targeted locations simultaneously. This project is inspired by the current thought of NASA in the search of life on Mars with the targeted location to be a water source. The algorithms proposed by the author make use of combining a modified Particle Swarm Optimization algorithm combined with fuzzy variables for added intelligence. An analysis of them is presented and efficiency is discussed.

J.M. Cadenas, V. Liern, R. Sala and J.L. Verdegay ("Fuzzy Linear Programming in Practice: An Application to the Spanish Football League") consider fuzzy linear programming problems as a hybridization of fuzzy sets theory and linear programming. In particular, they present a novel application of fuzzy linear programming to the formulation and solution of some problems arising in the Spanish football (or soccer) league. Basically, the main motivation is that uncertainty inherently associated with the parameters related to soccer teams in the Spanish

league may best be handled using fuzzy tools and techniques. Fuzzy linear programming models are developed which optimize the returns on investments made to maintain a high quality competition, which is finally given in an efficiency measure of the different teams that can be classified. Fuzzy data envelopment analysis models are used to provide team predictions as to their efficiency score. First, the author briefly present some basic elements of fuzzy sets theory and a brief review of the most typical problems and methods in fuzzy linear programming. Next, they develop an application of some selected fuzzy linear programming model to the problem considered. An example is solved using real data from the Spanish football in the season 2006/07.

We wish to thank all the contributors for their excellent work. We hope that the volume will be interesting and useful to the fuzzy optimization research community as well as other communities in which people may find fuzzy optimization tools and techniques useful to formulate and solve their specific problems.

We also wish to thank Dr. Tom Ditzinger and Ms. Heather King from Springer for their multifaceted support and encouragement.

February, 2010 Weldon A. Lodwick
Denver and Warsaw Janusz Kacprzyk

Contents

Part I: Introductory Sections

Fuzziness, Rationality, Optimality and Equilibrium in Decision and Economic Theories 3
Kofi Kissi Dompere

Introduction to Fuzzy and Possibilistic Optimization 33
Weldon A. Lodwick, Elizabeth Untiedt

Part II: Basic Issues

Aggregation Operators for Evaluating Alternatives 65
Vicenç Torra

Optimization and Aggregation Functions 77
Gleb Beliakov

Chebyshev Approximation of Inconsistent Fuzzy Relational Equations with Max-T Composition 109
Pingke Li, Shu-Cherng Fang

Part III: Various Types of Fuzzy Optimization and Fuzzy Mathematical Programming Models

A Survey of Fuzzy Convex Programming Models 127
Ricardo C. Silva, Carlos Cruz, José L. Verdegay, Akebo Yamakami

Approaches to Linear Programming Problems with Interactive Fuzzy Numbers 145
Masahiro Inuiguchi

Possibilistic Optimization Tasks with Mutually T-Related Parameters: Solution Methods and Comparative Analysis ... 163
Alexander Yazenin, Ilia Soldatenko

A Parametrized Model for Optimization with Mixed Fuzzy and Possibilistic Uncertainty 193
Elizabeth Untiedt

On Solving Optimization Problems with Ordered Average Criteria and Constraints 209
Włodzimierz Ogryczak, Tomasz Śliwiński

Fuzzy Dynamic Programming Problem for Extremal Fuzzy Dynamic System 231
Gia Sirbiladze

Vaguely Motivated Cooperation 271
Milan Mareš

Part IV: Fuzzy Network and Combinatorial Optimization

Computing Min-Max Regret Solutions in Possibilistic Combinatorial Optimization Problems 287
Adam Kasperski, Paweł Zieliński

Stochastic Bottleneck Spanning Tree Problem on a Fuzzy Network 313
Yue Ge, Hiroaki Ishii

The Use of Fuzzy Numbers in Practical Project Planning and Control 323
Dorota Kuchta

Part V: Applications

Ant Feature Selection Using Fuzzy Decision Functions 343
Susana M. Vieira, João M.C. Sousa, Uzay Kaymak

Application of Fuzzy Theory to the Investment Decision Process 365
Hiroshi Tsuda, Seiji Saito

Decision Making Techniques in Political Management 389
Anna M. Gil-Lafuente, José M. Merigó

Mathematical Approaches for Fuzzy Portfolio Selection Problems with Normal Mixture Distributions 407
Takashi Hasuike, Hiroaki Ishii

Fuzzy Random Redundancy Allocation Problems 425
Shuming Wang, Junzo Watada

Reliable Biological Circuit Design Including Uncertain Kinetic Parameters 457
Eva Sciacca, Salvatore Spinella

Fuzzy Optimal Algorithms for Multiple Target Convergence 479
Zach Richards

Fuzzy Linear Programming in Practice: An Application to the Spanish Football League 503
J.M. Cadenas, V. Liern, R. Sala, J.L. Verdegay

Author Index 529

Part I
Introductory Sections

Fuzziness, Rationality, Optimality and Equilibrium in Decision and Economic Theories

Kofi Kissi Dompere

Abstract. This essay presents structural categories of theories of optimization. It begins with the classical system leading to the establishment of an entry point of fuzzy optimization from the logic of the classical optimization. Four categories of optimization problems are identified. They are exact (non-fuzzy) and non-stochastic, exact (non-fuzzy) and stochastic categories that are associated with classical laws of thought and mathematics. The other categories are fuzzy and non-stochastic, and fuzzy-stochastic problems that are associated with fuzzy laws of thought and fuzzy mathematics. The canonical structures of the problems and their solutions are presented.

From these structures, similarities and differences in the problem structures and corresponding solutions are abstracted and discussed. The similarities and differences in the problem-solution structures of different categories are attributed to properties of exactness and completeness about information-knowledge structures in which the optimization problems are formulated and solved. The assumed degrees of exactness and completeness establish defective information-knowledge structure that generates uncertainties and produces inter-category differences in the optimization problem. The specific differences of intra-category algorithms are attributed to differences in the assumed functional relationships of the variables that establish the objective and constraint sets. The essay is concluded with taxonomy of solution structures and discussions on future research directions.

1 Introduction

In economic and decision theories about humanistic and non-humanistic systems, we speak of concepts of rationality, optimality, equilibrium and stability. These

Kofi Kissi Dompere
Department of Eeconomics
Howard Univeristy
Washington , D.C. 20059, USA,
Tel. #: 202-806-7682
e-mail: kdompere@howard.edu

W.A. Lodwick & J. Kacprzyk (Eds.): Fuzzy Optimization, STUDFUZZ 254, pp. 3–32.
springerlink.com

concepts allow us to derive testable and analytical conclusions at the level of explanatory theories. The same concepts allow us to derive prescriptive rules at the level of prescriptive theories. The manner in which these concepts are used and interpreted depends on the nature of decision-choice problem, its environment, conditions of analysis and synthesis between results and the assumed initial conditions. In the classical decision-choice systems, the explicit relationships among optimality, rationality and equilibrium processes are more or less required for scientific understanding. In addition, there is a troublesomely associated concept of stability of the analysis of the systems behavior in reference to these concepts in time and over time.

The concept of fuzziness as it relates to scientific reasoning has no meaning in relations to these basic four concepts in the classical decision-choice system if they are not clearly defined. The clear definitions require explications of the concepts to provide them with scientific meaning, conditions of measurability and framework of their uses. In order to attach scientific meanings to, as well as extend conditions of measurability and uses to the concepts of fuzzy optimality, fuzzy rationality, fuzzy equilibrium and fuzzy stability, scientific reasoning demands us to clearly explicate optimality, rationality, equilibrium and stability first within the established classical paradigm and then extend them to areas of fuzziness. This explication seeks to provide an entry and departing points for the uses of these concepts in research and teaching within the fuzzy paradigm.

At the entry point, we must have a clear understanding of the vision of the research and teaching programs on decision-choice behavior as these concepts are related to human action and knowledge construction. At the point of departure, we must make it explicit and clear the points of analytical difficulties that require cognitive change. This essay is more on explications of these concepts at the foundational level and how the fuzzy paradigm composed of its laws of thought and mathematics can assist in unraveling new frontiers of knowledge construction as well as designing a framework for constructing efficient rules of managing dynamics of social and natural forces through human action in penumbral regions of uncertainty and judgments thus showing points of analytical entry and cognitive departure.

The initial postulate in this essay is that all human actions on social, natural and mechanical systems are induced and controlled by decision-choice processes that constitute complete action systems [16] [17][55]. The evaluations of the efficiency of decision-choice systems are cost-benefit based. Any of the decision-choice actions comes from a decision-choice space, D. The decision-choice action $d \in D$ has an objective w that belongs to a space of objectives Ω. The decision-choice action involves, given $w \in \Omega$, assessments of alternatives which constitute a set, A with a generic element, $a \in A$ in order to implement the decision-choice action in A. The elements in the set, A, relate to each other in terms of their degree of relative effectiveness in accomplishing the selected objective, $w \in \Omega$. We say that the elements in A constitute a relation, $\mathcal{R}$ or $\mathcal{R}$ is a relation in A. The degree of effectiveness presents itself as a net usefulness (benefit/ utility) or a net suffering (cost/ disutility) to the decision-choice agent relative to the objective in a particular decision-choice subsystem.

Whether a decision-choice action on alternatives is seen from utility or disutility viewpoint, an order relation $(\succcurlyeq)$ is assumed over the alternatives in A in either ascending or descending order of subjective or objective magnitude producing an ordered set, $(\mathsf{A},\succcurlyeq)$. The symbol $(\succcurlyeq)$ may be interpreted in a general way as *better than or indifferent to* where the particular interpretation will depend on the specific decision-choice problem. The ordered set, $(\mathsf{A},\succcurlyeq)$, is conceptually consistent with our linguistic value system as well as our number system, $\mathbb{R}$ that is ordered by a relation "greater than or equal to" $(\geq)$ with an ordered set, $(\mathbb{R},\geq)$ in one-dimensional framework. In decision-choice systems, the partial ordering is induced by cost-benefit balances in such a way that the desired element may not be the largest. The concept and measure of cost-benefit balances present an important information set, $\mathbb{I}$ that restricts the objective that may be pursued and the domain of the set of alternatives that may be opened for decision-choice action. In other words, the relation $(\succcurlyeq)$ in A is not necessarily isomorphic to the order relation, $(\geq)$ in $\mathbb{R}$ even though they project the same idea to human thought.

2 Preconditions to Concepts of Rationality, Optimality, Equilibrium and Stability

To operationalize the order relation $(\succcurlyeq)$ through the concept of cost-benefit balances, two important concepts are introduced to close each decision-choice subspace. They are the set of objectives $\mathbb{A}$ that either presents benefit or cost characteristics but not both; and there is a set of constraints $\mathbb{B}$ that presents cost or benefit characteristics but not both in such a way that if $\mathbb{A}$ is benefit characteristic set then $\mathbb{B}$ is cost characteristic set. The constructions of $\mathbb{A}$ and $\mathbb{B}$ are such that there are functions that link their elements to the set of alternatives in A with order-preserving properties of $(\succcurlyeq)$.The functional relations are defined with the information set $\mathbb{I}$ that is processed to abstract a parameter space Θ that represents the knowledge space, $\mathbb{K}$. On the basis of the knowledge space, decision-choice actions are undertaken. Operating the interactions between the set of alternatives A and the information set $\mathbb{I}$, the criterion space, Λ and constraint space, Γ on which $\mathbb{A}$ and $\mathbb{B}$ acquire decision-choice meanings are established. The criterion space, Λ and constraint space, Γ are linked to the elements of the set of alternatives that act as the space of variables, $\mathbb{V}$ that is partitioned into decision-choice variables and the state variables. The decision-choice agent has a set of personality characteristics, Φ

that allows judgment to be rendered in the decision-choice space according to or in support of the criterion used.

Every decision-choice space is a complex and complicated interactions of all these subspaces. For the purpose of examining the meanings of rationality, optimality, equilibrium and stability in decision-choice processes and how they interrelate, we may represent the general decision-choice space D in a canonical form as:

$$D = \{\mathbb{V}, \Lambda, \Gamma, \Omega, \Theta, \succcurlyeq | \Phi\} \tag{1}$$

Where,

$\mathbb{V}$ = the space of state and decision-choice variables, Λ= criterion space

Θ = parameter space, Γ= the constaint space, Ω = the objective space

Φ= a set of personality characteristics, $\succcurlyeq$ represents order relation

From the decision-choice space the benefit space $\mathbb{B}$ and cost space $\mathbb{A}$ may be specified as

$$\left\{\begin{array}{l} \mathbb{B} = \{(b, \succcurlyeq) \in (\mathbb{V}, \Lambda, \Theta, \succcurlyeq) | \Phi, w \in \Omega\} \\ \\ \mathbb{A} = \{(c, \succcurlyeq) \in (\mathbb{V}, \Gamma, \Theta, \succcurlyeq) | \Phi, w \in \Omega\} \end{array}\right\} \tag{2}$$

Both $\mathbb{A}$ and $\mathbb{B}$ may be viewed individually as unconstrained decision-choice problems and together they constitute a constrained decision-choice problem. The problem is to select an element that yields the *best* benefit or cost value, and notice that the elements in both Λ and Γ that produce $\mathbb{A}$ and $\mathbb{B}$ are functions defined, under parametric restriction of Θ, in $(\mathbb{V}, \Lambda)$ and $(\mathbb{V}, \Gamma)$ respectively. The action of selecting the best as part of the characteristics of decision-choice behavior of decision-choice agents is an assumption that is abstracted as an element in the personality characteristic set, Φ.

The task of the decision-choice agent is to develop cognitive algorithms that will allow the best to be located in the decision-choice space whether it is constrained or unconstrained problem. The concept of *best* fits into what has been characterized in [15], [16] as the Euler's max-min principle that; *Nothing happens in the universe that does not have a sense of either certain maximum or minimum* [103]. The search for cognitive algorithms for finding the best has given rise to the concept and theory of optimization with related concepts and theories of rationality, equilibrium and stability. To abstract how these concepts and theories are related let us visit some aspects of the classical paradigm composed of its logic and mathematics for classical optimization theory.

3 The Concepts of Optimization, Rationality, Equilibrium and Stability in Decision-Choice Theories

Every theory about decision-choice behavior in both humanistic and non-humanistic behavior encounters the concept of the best act, path, design and many others. The concept of best is interpreted in terms of optimality, rationality and equilibrium and then examined relative to their stabilities. What are these concepts and how do they relate to each other. Let us provide explications to these concepts within the theories of decision-choice behavior which include control systems. Given these explications, we shall define the operating zones of the classical system that will provide a useful entry and point into the fuzzy logical system and departing point from classical logical system for optimization, rationality, equilibrium and stability. Since the analytical systems that carry these concepts are many, we shall provide general linguistic definitions for their explications that may have a universal coverage as well as lend themselves to symbolic mathematical representation and selected laws of thought.

A this cognitive juncture, it may be useful to remind ourselves of the classical laws of thought and exact symbolic representation of propositions, criticisms levied against it and responses that may be abstracted from fuzzy paradigm. The classical laws of thought is that: '*All propositions are either true or false'* These laws of thought have been criticized by Russell and suggested that *some substitute must be found, or all general accounts of deduction become impossible* [94, 63-64]. Furthermore, the exact symbols representing propositions and ideas within the classical laws of thought are also criticized by both Russell [93] and Max Black as they view classical logic and its mathematics. A statement by Black is useful ."*For*(classical) *mathematics is the study of all structures whose form can be expressed in symbols, it is the grammar of all symbolic systems and, as such, its methods are peculiarly appropriate to the investigation of its own internal structure*" [1, p.4].

These criticisms about the failure of the classical logic to account for vagueness in representation, ambiguities in interpretations and non-acceptance of contradiction in the classical laws of thought is met with the fuzzy paradigm where vagueness, ambiguities, and linguistic numbers are allowed in symbolic representation of propositions and ideas. This is supported by fuzzy laws of thought that has been stated in words as *"every statement is a set of true-false characteristics in varying proportion and that the acceptance of a proposition to be true or false is done on the basis of subjective decision-choice action in reconciling the conflict in the true-false proportion"* [15][16][17]. These fuzzy laws of thought replace the classical laws of thought. The importance of these classical-fuzzy analytics is seen in ordering and reasoning with items of sets. The process of replacing classical *exact symbols* with *fuzzy symbols* is basically referred to as *fuzzification*. The process of replacing the *classical laws of thought* with *fuzzy laws of thought* is basically refer to as *defuzzification*. This fuzzification-defuzzification process points to a propositional claim.

Proposition 3.1

Every problem in the classical paradigm has a fuzzy a counterpart, however, not every problem in the fuzzy paradigm has a classical counterpart.

Proposition 3.1 can be stated in terms of symbolic logic and mathematics and proven. The question of interest, at the moment, is: how does this fuzzification-defuzzification process carry on to analytical processes that involve optimization, rationality, equilibrium and stability? Few definitions are needed at this point.

Definition 3.1: Concept of Optimization

Optimization is a process about ordering the decision-choice elements in the decision choice space given the information concerning the behavior of the decision-choice space in such a way that the ordering is consistent with our number system or system of our linguistic magnitudes.

Definition 3.1.2: Concept of Rationality

Decision-choice rationality is a behavioral action that allows a selection to be made on the basis of a criterion where such a selection must satisfy the information constraint as defined by resources in the decision-choice space given the ordered set of alternatives.

Definition 3.1.3: Concept of Equilibrium

Equilibrium is a state in static or dynamic decision-choice system where the forces of expectations, broadly defined, in the environment do not cause a change in the knowledge support of the ordering of the decision-choice elements by optimization and hence do not change the action of rationality.

Definition 3.1.4: Concept of Stability

Stability in decision-choice system is a set of conditions describing the behavioral responses of the system's equilibrium such that expectation disturbances of the information set do not alter the ordered elements by optimization and the decision-choice action by rationality for the given knowledge support.

These four concepts are basically central organizing ideas that provide logical coherence for the understanding of decision-choice and control processes with given information-knowledge structures. The logical coherence imposes analytical conditions on the system such that decision-choice agents will not maintain their actions in states in which either preferred or better ones are available to them given the information-knowledge support. Because of various uses of the notion of rationality, it is analytically useful in the theories about decision-choice systems to use the term *optimal rationality* that is consistent with the conditions of logical coherence of decision-choice actions. The optimal rational actions are said to be in equilibrium states when the expectations relative to information signals flowing

from these states are such that the intentions of decision-choice agents are mutually consistent and can be implemented without a change of the intended action relative to the information-knowledge structure.

The concepts of optimality, rationality and equilibrium are not the same in meaning and content. They coincide and, therefore, interchangeable under specific conditions in the behavior of the decision-choice system. Optimization involves finding the best way of doing things given the information-knowledge support of the environment. The best way of doing things involves every aspect of human action in knowledge production and its uses to bring about, economic production, city planning, path of negotiation and other decisions involving mechanical systems, medical treatment, automation and many others. The practice of optimization involves more than locating the best relative to the index of rationality. One important thing in the practice of optimization is the nurturing of subjective appreciation of the value of information needed to construct the knowledge support for decision-choice action.

The knowledge abstracted from the information about the decision-choice system is essential in understanding the qualitative and quantitative nature of the decision-choice process or the control process. It allows a symbolic representation of the optimization problem in order to apply the available algorithms or to develop new ones to abstract the best in terms of solution. It, further, provides us with an appreciation of the role that subjective judgment interacts with the final result of the optimization leading to implementation. When the information-knowledge support about the decision-choice system, the measure of the rational criterion and the system's constraints are known, their real costs and benefits are defined [18][19]. In this respect, optimization, rationality, equilibrium and stability are computationally decisive within the classical paradigm in that possible choices are basically reduced to the best one. Optimization, rationality, equilibrium and stability are useful after the decision-choice problem is well posed. The concept of stability is intimately connected with the concept of sensitivity analysis which has come to be known in decision analysis, particularly, in economics as comparative statics and dynamics. The analytical process encompasses the general sensitivity theory about equilibrium and stability in the system's behavior in the parameter space Θ relative to the behavior of information-knowledge space $(\mathbb{I}\text{-}\mathbb{K})$.

4 Information-Knowledge Structure for the Development of Optimization

The areas of development of classical optimization and fuzzy optimization are conceptually, logically and mathematically different. They, however, interact in a useful way. These areas are defined not by the nature of mathematical reasoning and the set of algorithms but by the nature of the information-knowledge structure that imposes conditions on the structures of the decision-choice problem, the logic of reasoning and the needed mathematical algorithms as well as the entry and departing points of analyses and syntheses. The information-knowledge space is

induced by two sets of certainty and uncertainty characteristics. The space of certainty characteristics is then partitioned into exact full-information space and fuzzy (vague) full-information space. The uncertainty space is divided into exact and incomplete information and fuzzy and incomplete information subspaces. The nature of the assumed information set and the logic of the information processing will determine the structure of the knowledge space that will be represented by the parametric space. The information-knowledge structure will determine the problem structure and the needed algorithm for optimization, rationality, equilibrium and stability. These information-knowledge subspaces partition the problem space into four cohorts that are diagrammatically illustrated in Figure 1.

<table>
<tr><td>COMPLETENESS
EXAACTNESS</td><td>COMPLETE INFORMATION</td><td colspan="2">INCOMPLETE INFORMATION</td></tr>
<tr><td>EXACTNESS (NON-FUZZINESS)</td><td>EXACT AND COMPLETE INFORMATION SPACE (UNCERTAINTY-FREE)
<u>COHORT I</u></td><td colspan="2">EXACT AND INCOMPLETE INFORMATION SPACE (STOCHASTIC UNCERTAINTY)
<u>COHORT II</u></td></tr>
<tr><td rowspan="2">INEXACT/ FUZZINESS</td><td rowspan="2">COMPLETE AND FUZZY INFOEMATION SPACE (FUZZY NON-STOCHASTIC UNCERTAINTY)
<u>COHORT III</u></td><td colspan="2">INCOMPLETE AND FUZZY INFORMATION SPACE
<u>COHORT IV</u></td></tr>
<tr><td>FUZZY-STOCHASTIC UNCERTAINTY
<u>COHORT IVA</u></td><td>STOCHASTIC-FUZZY UNCERTAINTY
<u>COHORT IVB</u></td></tr>
</table>

Fig. 1 Exactness-Completeness Partition of the Development of Optimization Problems and Required Algorithms

5 Optimization, Rationality and Equilibrium in the Classical Paradigm

There are two aspects of the classical paradigm that are of concern to us and for discussions about optimization, rationality and equilibrium. The first one is the classical *laws of thought* in developing cognitive algorithms. The second one is the assumption of *exactness* in the information-knowledge spaces leading to exact symbolic representation, exact reasoning, analysis, synthesis and results in deriving cognitive algorithms that relate to optimality, rationality, equilibrium and stability. The development of the classical theories of optimization with the corresponding rationality, equilibrium and stability are restricted to uncertainty-free and exact stochastic uncertainty decision-choice sub-spaces. Corresponding to these two spaces are the classical exact topology that is free from stochastic

uncertainty, and classical exact topology with stochastic uncertainty. They are classical in the sense that the logic and algorithmic developments follow the classical laws of thought and the corresponding mathematics as well as satisfying the postulate of symbolic exactness.

At the level of ordinal and cardinal analyses, the classical mathematical spaces for the framing and analyses of the decision-choice problems for optimization are shown in Figure 2 where $\mathsf{T} = \langle \mathfrak{X}, \mathfrak{I} \rangle$ is a topological space that correspond to a world of exactness and certainty with $\mathfrak{X}$ defined as nonempty set of classical exact points and $\mathfrak{I}$ is a family of subsets of $\mathfrak{X}$ such that $\mathfrak{X} \in \mathfrak{I}$ named as Cohort I. Similarly, $\mathscr{P} = (\Omega, \mathfrak{B}, \mathrm{P})$ is a topological space that correspond to the world of exactness and stochastic uncertainty where, Ω is a sample space; $\mathfrak{B}$ is Borel field and P is a set of probability density functions defined over $\mathfrak{B}$ and named as Cohort II.

Cohort I may be referred to as classical exact and non-stochastic mathematical space where substantial portion of theories on decision-choice behavior have been formulated, solved and analyzed with the toolbox of classical optimization, mathematical programming and Aristotelian logic with principle of non-contradiction.

	Non-Stochastic	**Stochastic**
Non-Fuzzy	**COHORT I** **Non-Stochastic and Non-fuzzy Topological Space** $\mathsf{T} = \langle \mathfrak{X}, \mathfrak{I} \rangle$	**COHORT II** **Non-Fuzzy and Stochastic Topological Space** $\mathscr{P} = (\Omega, \mathfrak{B}, \mathrm{P})$

Fig. 2 Various Topological Spaces and Categories of Classical Mathematical Environment for Classical Optimization and Decision-Choice Theories for Decision-Choice Rationality

The classical optimization problems that correspond to Cohort I follow a canonical representation of the form:

$$\left.\begin{aligned} \underset{\psi \in \mathbb{A} \cap \mathbb{B}}{\text{opt}} \ & \lambda(\psi \mid \theta) \in \Lambda(\Psi, \Theta) \\ s.t. \ & \mathbb{A} \cap \mathbb{B} \neq \varnothing \\ & \mathbb{A} \subset \{\Psi \mid \Theta\} \\ & \mathbb{B} \subset \{\Gamma \mid \Theta\} \\ & \Theta \subseteq \{\mathbb{I} \mid \Phi\} \end{aligned}\right\} \tag{3}$$

where Λ is the criterion space, Ψ , is the objective space, Γ , the constraint space. The elements of these spaces are defined over the control (decision-choice) and state variables $\mathbb{V}$. The $\mathbb{I}$, is a full and exact information space from which an a parameter space Θ is abstracted to represent an exact and full knowledge structure, $\mathbb{K}$ and Φ is the space of personality characteristics of the decision-choice agent. This is the *exact and non-stochastic classical optimization problem* whose solution appears in the form:

$$\lambda^* = \lambda\left(\psi^* \in \left(\mathbb{A} \bigcap \mathbb{B}\right) | \theta \in \Theta \subset \mathbb{I}\right) \tag{4}$$

The term $\left(\mathbb{A} \bigcap \mathbb{B}\right)$ defines the *critical region* of optimal decision-choice activity. It is the stationary value relative to Θ as a parameter space. The structure of equations (3) and (4) includes all classes of deterministic optimization problems whose optimal solutions may be abstracted by techniques and methods of linear and nonlinear computational algorithms in addition to others that the classical logic and mathematics offer [20][44],[45][70]. This may also be called the *perfect information-knowledge optimization problem*. The information-knowledge structure is said to be perfect for a class of optimization problems if it is devoid of all uncertainties (that is vagueness and incompleteness). The optimization problem is said to be perfect if, additionally, its algorithms and analytics are derived from classical mathematics and the interpretation of the results follow the classical laws of thought.

The optimization problems in the exact and incomplete information-knowledge structure of Cohort II introduce extra complication for the development of optimizing algorithms in Cohort I. It is classical in that the information-knowledge structure is exact but defective in terms of completeness. The defective information-knowledge structure introduces *stochastic uncertainty* that requires the inclusion of the degree of belief of knowledge certainty available in formulating the optimization problem and constructing the algorithms for the ranking of the alternative in either ascending or descending order of magnitude for the application of optimal rationality. The measure of degree of belief on knowledge certainty is specified around each of the optimizing variables that affect the ordering process.

The process may be called stochastic ordering which is associated with the probability distribution. Given that the degrees of stochastic uncertainty are preordered by a probability distribution, the expected consequences are then ordered by preferences or some index of ranking. The preference ranking, or a criterion function, is induced by expected-value function on the basis of which optimization algorithms, that satisfy the goal-constraint configuration, are abstracted. An information-knowledge structure that is exact but incomplete is said to be defective with stochastic uncertainty and the variable for reasoning is called a random variable.

The preordering of the degree of stochastic uncertainty by the probability distribution is conditional on the absence of vagueness and imprecision of concepts, measurement and reasoning. In general, the optimization problem is a logical

extension of the classical optimization problem with a perfect information-knowledge structure of Cohort I. The general canonical representation may be written as:

$$\left.\begin{array}{l} \underset{\psi\in\mathbb{A}\bigcap\mathbb{B}}{\text{opt}}\ \lambda\left(\psi\,|\,\theta,\mathbb{P}\right)\in\Lambda\left(\left\{\Psi,\mathbb{B}_0,\mathbb{P}\right\},\Theta\right) \\ \mathbb{A}\bigcap\mathbb{B}\neq\varnothing \\ \mathbb{A}\subset\left(\left\{\Psi,\mathbb{B},\mathbb{P}\right\}|\,\Theta\subset\hat{\mathbb{I}}\right) \\ \mathbb{B}\subset\left(\left\{\Psi,\mathbb{B},\mathbb{P}\right\}|\,\Theta\subset\hat{\mathbb{I}}\right) \\ \Theta\subset\left(\hat{\mathbb{I}},\mathbb{P}|\Phi\right) \end{array}\right\} \tag{5}$$

The optimization process that allows the search for the best element $\psi=\psi^*\in\mathbb{A}\bigcap\mathbb{B}$ conditional on a given set of parameters and probability distributions that are defined over both the sets of alternatives and constraints is called stochastic optimization which includes expected-value optimization. Again, notice that $\mathbb{A}\bigcap\mathbb{B}$ equipped with probability distribution function is the critical region of decision-choice activity. Generally, the optimal elements appear as:

$$\lambda\left(\left(\psi^*,\mathrm{p}^*\right)|\,\theta\subset\Theta,\mathrm{p}^*\in\mathbb{P}\right)=\underset{\psi\in\mathbb{A}\bigcap\mathbb{B}}{\text{opt}}\ \lambda\left(\psi\,|\,\theta,\mathbb{P}\right) \tag{6}$$

The solution given in eqn. (6) defines a family of optimal decisions over the parameter and probability spaces. The parameter variations specify the paths of informationally optimal values as a family of sensitivity functions that depend on the environment of optimization characterized by the relevant parameter space and the corresponding probability distribution functions. The symbol $\hat{\mathbb{I}}$ represents incomplete information with defective knowledge structure, $\hat{\mathbb{K}}$ with stochastic uncertainty and $\mathbb{P}$ represents a probability space whose specific distribution is known. A typical class from the above canonical form of eqn. (5) may be written as:

$$\left.\begin{array}{l} \underset{\psi\in\mathbb{A}}{\text{opt}}\ \underset{\mathbb{A}\subset\Psi}{\int}\lambda\left(\psi\,|\,\Phi\right)f_0\left(\psi\right)\mathrm{d}\psi\left(\left\{\Psi,\mathbb{B},\mathbb{P}\right\}|\,\Theta\subset\hat{\mathbb{H}}\right) \\ \qquad\text{s.t.}\ \underset{\psi\in\mathbb{B}\subset\Gamma}{\int}\gamma\left(\psi\,|\,\Phi\right)f_1\left(\psi\right)\mathrm{d}\psi\leq\overline{\pi} \end{array}\right\} \tag{7}$$

where $\overline{\pi}$ is a predetermined value, and $f_0\left(\bullet\right)$ and $f_1\left(\bullet\right)$ are probability density functions defined over spaces of alternatives goals and constraints respectively. This is the class of classical exact stochastic optimization problems.

6 Optimization, Rationality, Equilibrium and Stability in the Fuzzy Paradigm

Let us now relate our understanding of the structure of the classical paradigm as it relates to optimization, rationality and equilibrium in theories about decision-choice behavior to those of the fuzzy paradigm. Like the classical paradigm, there are two aspects of the fuzzy paradigm that are of concern to us and for discussions about optimization, rationality and equilibrium. The first one is the *fuzzy laws of thought* in developing cognitive algorithms. The second one is the assumption of *inexactness* or *vagueness* in the information-knowledge spaces leading to fuzzy or approximate reasoning, analysis, synthesis and results in deriving cognitive algorithms that relate to optimality, rationality, equilibrium and stability in the decision-choice spaces. The development of the fuzzy theories of optimization with the corresponding fuzzy rationality, fuzzy equilibrium and fuzzy stability are restricted to *fuzzy-non-stochastic uncertainty* and *fuzzy-stochastic uncertainty* decision-choice sub-spaces. Corresponding to these two spaces are the *non-stochastic-fuzzy topology* that is free from stochastic uncertainty, *fuzzy-stochastic* or *stochastic-fuzzy topology* with stochastic uncertainty. The uncertainties and topologies are said to be fuzzy in the sense that the logic and algorithmic developments follow the fuzzy laws of thought and the corresponding mathematics and symbolic representation of ideas that incorporate subjectivity, vagueness and ambiguities as part of reasoning. The relevant information-knowledge subspaces are Cohorts III and IV in Figure 1.

6.1 The Mathematical Spaces for the Development of Fuzzy Optimization

The mathematical subspaces that correspond to Cohort III and IV may be obtained by fuzzification process of classical exact and non-stochastic topological space $\mathsf{T} = \langle \mathfrak{X}, \mathfrak{I} \rangle$ to obtain fuzzy non-stochastic topological space, while the fuzzification process allows the cognitive transformation of classical exact stochastic topological space, $\mathcal{P} = (\Omega, \mathcal{B}, \mathrm{P})$ to either fuzzy-stochastic or stochastic-fuzzy topological space respectively . These will form the mathematical spaces for stating the problems and the development of algorithms for fuzzy optimization, and analysis of fuzzy rationality, equilibrium and stability.

The fuzzification process is such that the corresponding fuzzy non-stochastic topological space will appear as a triplet of the form $\mathcal{A} = \langle \mathcal{X}, \mathcal{T}, \mathfrak{M} \rangle$. Interestingly, it may be noted that $\mathcal{X}$ is nonempty set of fuzzy points where each fuzzy point is a set of the form $\mathcal{D} \in \mathcal{X}$ and equipped with a membership function $\mu_{\mathcal{D}}(\bullet) \in \mathfrak{M}$ where $\mathfrak{M}$, is the set of membership functions that is a fuzzifier on the classical exact topological space. We note that $\mathcal{X}$ is a family of

sets and $\mu_{\mathfrak{D}}(\bullet) \in \mathfrak{M}$ specifies the degree to which each element of $\mathfrak{D}$ belongs to $\mathfrak{D}$. The family is specified as $\boldsymbol{\mathcal{X}} = \left\{\mathfrak{D} \mid \mathfrak{D} = \left(x, \mu_{\mathfrak{D}}(x)\right) : x \in \mathfrak{D}, \mu(x) \in \mathfrak{M}_{\mathfrak{D}}\right\}$. The structure $\boldsymbol{\mathcal{T}}$ defines a family of families, $\boldsymbol{\mathcal{X}}$ of fuzzy sets as specified and $\mathfrak{M}_{\mathfrak{D}}$ is a set of membership characteristic function relevant to $\mathfrak{D}$ where the fuzzy topology is constructed around fuzzy points through the fuzzification of classical exact points. The classical points will correspond to classical exact numbers while the fuzzy points will correspond to fuzzy numbers (see Figure 3).

	Non-Stochastic	**Stochastic**	
FUZZY	**COHORT III** **Fuzzy and Non-Stochastic Topological Space** $\tilde{\mathcal{A}} = \langle \mathcal{X}, \mathcal{T}, \mathfrak{M}_{\mathcal{X}} \rangle$ **OR** $\tilde{\mathcal{B}} = \langle \mathcal{X}, \mathcal{T}, \mathfrak{M}_{\mathcal{T}} \rangle$ **(Fuzzy-non-stochastic variable)**	**COHORT IV**	
		Non-Separable	**Separable**
		A)Fuzzy-Stochastic Topological Space (fuzzy probability with Fuzzy-random variable $\tilde{\mathcal{K}} = (\Theta, \mathfrak{B}_{\Theta}, \mathfrak{M}_{\mathrm{P}}, \mathrm{P})$	**A.** $\tilde{\mathcal{K}} = \tilde{\mathcal{A}} \otimes \mathcal{P}$ **Reduced into fuzzy variable units for summability in the fuzzy and non-stochastic topological space**
		B) Stochastic-Fuzzy Topological Space (probability of fuzzy Variable defined on random-fuzzy variable $\tilde{\mathcal{K}} = \langle \mathcal{V}, \mathcal{T}, \mathfrak{M}_{\mathcal{T}}, \mathrm{P}_{\mathfrak{M}_{\mathcal{T}}} \rangle$	**B**. $\tilde{\mathcal{C}} = \tilde{\mathcal{A}} \otimes \mathcal{P}$ **Reduced into random variable units for summability in the non-fuzzy and stochastic topological space.**

Fig. 3 Fuzzy Topological Spaces and Categories of Mathematical Environment for Fuzzy Optimization Theories for Decision-Choice Rationality

The fuzzification of the classical exact stochastic topological space corresponding to Cohort II is complex as well as logically challenging. The complexity arises from the emergence of two different but interrelated variables of fuzzy-random variable and random-fuzzy variable. Corresponding to the fuzzy-random variable we have fuzzy-stochastic topology where the stochastic topology

of Cohort II must be fuzzified. Similarly, corresponding to random-fuzzy variable we have stochastic-fuzzy topology where the fuzzy topology must be randomized. The logical challenge involves not only finding the right topologies but examining the nature of fuzzy-random interactions and determining the conditions of separability and non-separability of fuzziness and randomness associated with each of the two complex variables that will lead to a clear problem statement. The mathematical spaces for logical reasoning and development of system of optimizing algorithms are shown in Figure 3.

The symbolic representations in Figure 3 have the following conceptual meanings:

1. $\mathsf{T} = \langle \mathfrak{X}, \mathfrak{I} \rangle$ is the classical exact non-stochastic topological space and the symbols are as have been explained in Figure 1 with $\mathfrak{X}$ is defined as nonempty set of classical exact points and $\mathfrak{I}$ is a family of subsets of $\mathfrak{X}$ such that $\mathfrak{X} \in \mathfrak{I}$.
2. $\mathscr{P} = (\Omega, \mathcal{B}, \mathrm{P})$ is a topological space that corresponds to the world of exactness and stochastic uncertainty where, Ω is a sample space; $\mathcal{B}$ is Borel field and P is probability density function defined over $\mathcal{B}$
3. $\mathcal{A} = \langle \mathcal{X}, \mathcal{T}, \mathfrak{M}_{\mathcal{X}} \rangle$ is fuzzy non-stochastic topological space constructed around a set of fuzzy points $\mathcal{D} \in \mathcal{X}$ with $\mathcal{T}$ as a family of fuzzy subsets of $\mathcal{X}$ where $\mathcal{X} \in \mathcal{T}$ and $\mathfrak{M}_{\mathcal{X}}$ is a set of membership characteristic functions that define the degrees to which fuzzy points $\mathcal{D}$ belong to $\mathcal{X}$.
4. $\mathcal{B} = \langle \mathcal{X}, \mathcal{T}, \mathfrak{M}_{\mathcal{T}} \rangle$ is an alternative construct of fuzzy non-stochastic topological space that is different from $\mathcal{A} = \langle \mathcal{X}, \mathcal{T}, \mathfrak{M}_{\mathcal{X}} \rangle$ in that the membership characteristic functions $\mathfrak{M}_{\mathcal{T}}$ define the degree to which a subset of $\mathcal{X}$ belongs to $\mathcal{T}$ instead of an element of $\mathcal{X}$ belonging to $\mathcal{X}$.
5. $\mathcal{K} = (\tilde{\Theta}, \mathcal{B}_{\Theta}, \mathfrak{M}_{\mathrm{P}}, \mathrm{P})$ is fuzzy-stochastic topological space where $\tilde{\Theta}$ **is a fuzzy sample space,** $\mathcal{B}_{\Theta}$ is a fuzzy Borel set, P is a fuzzy set of probability values whose degrees of belonging are defines by a set of membership functions $\mathfrak{M}_{\mathrm{P}}$ where any $\mathrm{p} \in \mathrm{P}$ is fuzzy (inexact) probability [This is a fuzzification of Cohort II].
6. $\mathcal{R} = \langle \mathcal{V}, \mathcal{T}, \mathfrak{M}_{\mathcal{V}}, \mathrm{P}_{\mathfrak{M}_{\mathcal{V}}} \rangle$ is the construct of stochastic-fuzzy topological space where $\mathcal{V}$ is a set of fuzzy points and $\mathcal{T}$ is a family of subsets of $\mathcal{V}$ such that $\mathcal{V} \in \mathcal{T}$, $\mathfrak{M}_{\mathcal{V}}$ is a set of membership functions defining the degrees to which the elements of $\mathcal{V}$ belonging to it. The $\mathrm{P}_{\mathfrak{M}_{\mathcal{V}}}$ defines a set of probability values that are

attached to the elements in $\mathfrak{M}_{\nu}$. [This is a randomization of Cohort III]. The two fuzzy topological spaces of $\mathcal{K} = \left(\Theta, \mathcal{B}_{\Theta}, \mathfrak{M}_{\mathrm{P}}, \mathrm{P}\right)$ and $\mathcal{R} = \left\langle \mathcal{V}, \mathcal{T}, \mathfrak{M}_{\nu}, \mathrm{P}_{\mathfrak{M}_{\nu}} \right\rangle$ are constructed under the assumption that fuzziness and randomness are insepararable. When they are separable, we then combine exact stochastic topological space of Cohort II with fuzzy non-stochastic topological space of Cohort III to obtain either $\mathcal{R} = \mathcal{A} \otimes \mathscr{P}$ or $\mathcal{Q} = \mathcal{A} \otimes \mathscr{P}$.

6.2 Different Structures of Fuzzy Optimization Problems for Rationality and Equilibrium: The Non-stochastic Case

The optimization problem under full information and fuzzy environment where the symbol $(\sim)$ indicates fuzzy sets may be canonically represented as:

$$\left.\begin{array}{l} \underset{\psi \in \tilde{\mathbb{A}} \cap \tilde{\mathbb{B}}}{\text{opt.}}\ \lambda\left(\psi \mid \theta, \mu\right) \in \Lambda\left(\tilde{\Psi}, \tilde{\Theta}\right) \\ \quad s.t.\quad \tilde{\mathbb{A}} \cap \tilde{\mathbb{B}} \neq \varnothing \\ \qquad \tilde{\mathbb{A}} \subset \left\{\tilde{\Psi} \mid \tilde{\Theta}, \mathbb{M}\right\} \\ \qquad \tilde{\mathbb{B}} \subset \left\{\tilde{\Gamma} \mid \tilde{\Theta}, \mathbb{M}\right\} \\ \qquad \tilde{\Theta} \subseteq \left\{\tilde{\mathbb{I}} \mid \Phi, \mathbb{M}\right\} \end{array}\right\} \tag{8}$$

Under such a fuzzy specification of optimization problem, the algorithms are developed to abstract the optimal fuzzy element that appears as:

$$\sigma^{*} = \delta\left(\psi^{*}\mu^{*} \mid \theta, \Phi\right),\ \ \sigma^{*} \in \Delta = \left(\tilde{\mathbb{A}} \cap \tilde{\mathbb{B}}\right),\ \psi^{*} \in \Psi,\ \mu^{*} \in \mathbb{M} \tag{9}$$

The distinguishing characteristics of the fuzzy non-stochastic class of optimization are:

1) All the relevant information about the optimization problem is full.
2) However, the relevant full information is fuzzy in the sense that one or more of the components is fuzzy.
3) The fuzzy information is processed into knowledge base and summarized by fuzzy structural parameters where errors and approximations are allowed.

A typical case is to optimize a fuzzy goal $\left(\tilde{\mathbb{A}}, \mu_{\tilde{\mathbb{A}}}(\psi)\right)$ subject to a fuzzy constraint $\left(\tilde{\mathbb{B}}, \mu_{\tilde{\mathbb{B}}}(\psi)\right)$ which is simply

$$\left.\begin{array}{l} \underset{\psi\in \tilde{\mathbb{A}}\cap\mathbb{B}=\Delta}{\text{opt}} \ \mu_{\tilde{\mathbb{A}}}(\psi) \\ \text{s.t}\left[\mu_{\tilde{\mathbb{A}}}(\psi)-\mu_{\tilde{\mathbb{B}}}(\psi)\right]\le 0 \end{array}\right\} \tag{10}$$

As a fuzzy program where $\mu_{\tilde{\Delta}}(\psi)=\mu_{\tilde{\mathbb{A}}}(\psi)\wedge\mu_{\tilde{\mathbb{B}}}(\psi)$ **(see [18] [72] [19]).**

6.3 Different Structures of Fuzzy Optimization Problems for Rationality and Equilibrium: The Stochastic Case

The optimization problems and the developments of relevant algorithms may proceed from a judicious extension of conditions of non-stochastic fuzzy optimization into the domain of classical stochastic optimization to obtain conditions of optimality that take into account the presence of both fuzzy and stochastic characteristics in the optimizing space. The variables in this space are of two types of fuzzy-random variable that is associated with the fuzzification of Cohort II, and random-fuzzy variable that is associated with randomization of Cohort III where the fuzziness and randomness may be separable or inseparable. The presence of these two variables in optimizing activities leads to fuzzy-stochastic optimization and stochastic-fuzzy optimization. In the discussions that follow , fuzziness and randomness may be viewed in order of prior or posterior.

6.3.1 Fuzzy-Stochastic Optimization

An approach to fuzzy-stochastic optimization may proceed where stochastic uncertainty is taken to precede the fuzzy uncertainty. A problem example would be a case where the variable is random and the outcome acquires a linguistic variable such as large or small. The question, then, is how large or how small is small after the outcome? In this case, the uncertainty inherent in the optimizing variable is dichotomized. As such, a randomization process is first introduced on the relevant optimizing variables where the problem is formulated to allow for stochastic optimization in order to obtain optimal stochastic certainty equivalences and optimal stochastic decision rules in a fuzzy environment. In the example given above the fuzzy environment is in relation to an outcome that is to be determined as large or small. The stochastic-certainty-equivalent values are then considered as fuzzy variables to allow the establishment of effective fuzzification-defuzzification process. The process of randomization and stochastic optimization is to account for information-knowledge deficiency due to stochastic uncertainty. The fuzzification process is to account for vagueness, ambiguities, inexactness and approximations that constrain concept formations, thought representations, interpretations of optimal stochastic values and subjective formation of probability distributions. In other words, to introduce fuzzy outcome and fuzzy logical reasoning into the optimization process where fuzzy optimization allows

defuzzification through the use of fuzzy laws of thought (logic of fuzzy mathematics).

In general fuzzy-stochastic optimization leads to three optimal quantitative elements of exact certainty equivalent value of the optimizing variable ψ^*, associated probability value, $p\left(\psi^*\right)$ as measure of stochastic uncertainty (probabilistic belief), and optimal degree of exactness $\mu\left(\psi^*\right)$ as a measure of fuzzy uncertainty (possibilistic belief) associated with exactness of outcomes. The relative similarities and differences in probabiliatic and possibilistic beliefs in knowledge production and mathematics of uncertainties are discussed in [15] [16] [57] [61] [63] [64].

Essentially, the analytical process to formulate the optimization problem is such that we optimize the expected value of an objective function, $\lambda(\bullet)\in\Lambda(\bullet)$, among a set of objective functions, given the relevant information-knowledge structure $\Theta\subset\mathbb{I}$ and probability structure. The probability structure is a measure of degree of knowledge completeness that may be viewed as defining the state of knowledge in support of the optimizing system subject to an expected value of the constraint. A representation of the fuzzy-stochastic optimization process may be represented in two steps as:

$$\left.\begin{array}{l}\underset{\psi\in\mathbb{A}\cap\mathbb{B}}{\text{opt}} \int\limits_{\mathbb{A}\subset\Psi}\lambda(\psi\,|\,\theta)\,f(\psi\,|\,\theta)\,d\psi \\ \text{s.t.}\int\limits_{\mathbb{B}\subset\Psi}\gamma(\psi\,|\,\theta)g(\psi\,|\,\theta)\,d\psi\le\bar{u}\end{array}\right\}=\sigma^* \tag{11}$$

The stochastic optimal value is σ^* with corresponding probability $p=P\left(\sigma^*\right)$ that value σ^* is now viewed in dual fuzzy sets as an objective set, $\tilde{\mathbb{A}}$ with $\mu_{\tilde{\mathbb{A}}}\left(\sigma^*\right)$ and constraint set, $\tilde{\mathbb{B}}$ with $\mu_{\tilde{\mathbb{B}}}\left(\sigma^*\right)$ in determining the needed optimal crisp value by fuzzy optimization to obtain $\mu_{\tilde{\mathbb{A}}}\left(\sigma^*\right)$ subject to $\mu_{\tilde{\mathbb{B}}}\left(\sigma^*\right)$. The problem is the same as finding $\sigma^*\in\tilde{\mathbb{A}}\bigcap\tilde{\mathbb{B}}$ which reduces to:

$$\left.\begin{array}{l}\underset{\sigma^*\in\Delta}{\text{opt}}\,\mu_{\tilde{\mathbb{A}}}\in[0,1] \\ \text{s.t.}\quad\left[\mu_{\tilde{\mathbb{A}}}\left(\sigma^*\right)-\mu_{\tilde{\mathbb{B}}}\left(\sigma^*\right)\right]\le 1\end{array}\right\}=\sigma^{**} \tag{12}$$

where $\bar{u}$ is a predetermined value, $f(\psi\,|\,\theta)$ and $g(\psi\,|\,\theta)$ are probability density functions defined over the objective and constraint spaces conditional on

parameter space respectively. The optimal solution σ^{**} in terms of ψ is of the form

$$\sigma^{**} = \left(\psi^{**}, p\left(\psi^{**}\right), \mu\left(\psi^{**}\right) \mid \theta, \mathbb{P}, \mathbb{M}, \Phi\right) \tag{13}$$

Alternatively, we may specify the present value of the goal and the present value of the constraint. We then induce a fuzzification process over them to obtain fuzzy present values of goals and constraints (We may note that the fuzzy present value is not the same as present fuzzy value). The problem is then formulated for fuzzy present value optimization to obtain exact present value equivalences and optimizing algorithms in support of fuzzy- stochastic optimality. Here, the optimal conditions are derived under further conditions that the probability distributions are given or known without vagueness. The vagueness is found in the stochastic optimal outcomes bringing into the optimization process another dimension of uncertainty that is to be modeled through the fuzzy process with membership functions or possibility distributions. The needed membership function or the possibility distribution is either selected from a set of membership functions or subjectively constructed to reflect expressed possibilistic belief on the expected values.

The fuzzy optimization is then undertaken subject to the stochastic uncertainty through present-value equivalences of goals and constraints. The general structure of the optimizing problem leading to fuzzy-stochastic optimality may be represented in a canonical form as:

$$\left.\begin{aligned} &\underset{\psi \in \tilde{\mathbb{A}} \cap \tilde{\mathbb{B}}}{\text{opt}}\ \lambda\left(\psi \mid \theta, \mu(\bullet), p(\bullet)\right) \in \Lambda\left(\Psi \mid \mathfrak{B}_{\psi}, \mathbb{M}_{\psi}, p, \theta, \Phi\right) \\ &\text{s.t.}\quad \tilde{\mathbb{A}} \cap \tilde{\mathbb{B}} \neq \varnothing \\ &\qquad \tilde{\mathbb{A}} \subset \left(\Psi \mid \mathfrak{B}_{\psi}, \mathbb{M}_{\psi}, p, \tilde{\Theta}\right) \\ &\qquad \tilde{\mathbb{B}} \subset \left(\Gamma \mid \mathfrak{B}_{\gamma}, \mathbb{M}_{\gamma}, p, \tilde{\Theta}\right) \\ &\qquad \tilde{\Theta} \subset \tilde{\mathbb{I}} \end{aligned}\right\} \tag{14}$$

where $\tilde{\mathbb{A}}$ is a fuzzy present value goal with probability density function $p = f\left(\psi \mid \bullet\right),\ \psi \in \Psi,$. Similarly $\tilde{\mathbb{B}}$ is a fuzzy present value constraint also with probability density function, $p = g\left(\psi \mid \bullet\right),\ \gamma \in \Gamma$, with $\mathfrak{B}$ as a general Borel set on decision space where $\mathbb{M}$ is a general set of membership functions for fuzziness and $\Psi \bigcap \Gamma \neq \varnothing$. If $\mu_{\tilde{\mathbb{A}}}(\psi)$ and $\mu_{\tilde{\mathbb{B}}}(\psi)$ are the membership functions for the fuzzy present-value goal and fuzzy present-value constraint then the fuzzy logical process leads to the fuzzy decision that may be written as

$\tilde{\tilde{\Delta}} = \tilde{\tilde{\mathbb{A}}} \bigcap \tilde{\tilde{\mathbb{B}}}$ with membership function $\mu_{\tilde{\tilde{\Delta}}}(\psi)$, and $\psi = \gamma \in \Psi \bigcap \Gamma$ whose fuzzy optimization problem may be written as:

$$\left.\begin{aligned} &\underset{\psi \in \tilde{\tilde{\Delta}}}{\text{opt}}\, \mu_{\tilde{\tilde{\mathbb{A}}}}(\hat{\psi}) \geq 0 \\ &\text{s.t.} \quad \left[\mu_{\tilde{\tilde{\mathbb{A}}}}(\hat{\psi}) - \mu_{\tilde{\tilde{\mathbb{B}}}}(\hat{\psi})\right] \leq 0 \end{aligned}\right\} \sigma^* \tag{15}$$

where $\hat{\psi}$ is fuzzy present value variable. The optimal values in support of fuzzy stochastic rationality may then be written as:

$$\sigma^* = \sigma\left(\psi^*, p\left(\psi^*\right), \mu\left(\psi^*\right) \mid \theta, \mathbb{P}, \mathbb{M}, \Phi\right) \in \Delta \tag{16}$$

There is another case of fuzzy-stochastic optimization where the conditions of fuzziness are found in the probability distribution of the random variable. This gives rise to fuzzy or inexact probabilities that are associated with the values of the random variable. In this case, probability values are fuzzy sets and the optimization structure is such that the membership function is expressed over the probabilities. The problem representation of eqn. (14) becomes

$$\left.\begin{aligned} &\underset{\psi \in \tilde{\tilde{\mathbb{A}}} \cap \tilde{\tilde{\mathbb{B}}}}{\text{opt}}\, \lambda\left(\psi \mid \theta, \mu(\bullet), p(\bullet)\right) \in \Lambda\left(\Psi \mid \mathfrak{B}_{\psi}, \mathbb{M}_p, p, \theta, \Phi\right) \\ &\text{s.t.} \quad \tilde{\tilde{\mathbb{A}}} \bigcap \tilde{\tilde{\mathbb{B}}} \neq \varnothing \\ &\qquad \tilde{\tilde{\mathbb{A}}} \subset \left(\Psi \mid \mathfrak{B}_{\psi}, \mathbb{M}_p, p, \tilde{\Theta}\right) \\ &\qquad \tilde{\tilde{\mathbb{B}}} \subset \left(\Gamma \mid \mathfrak{B}_{\gamma}, \mathbb{M}_p, p, \tilde{\Theta}\right) \\ &\qquad \tilde{\Theta} \subset \tilde{\mathbb{I}} \\ &\qquad \psi, \lambda \in \left(\Psi \bigcap \Gamma\right) \end{aligned}\right\} \tag{17}$$

The optimization problem expressed by eqn. (17) reflects in part the problem of fuzzy (inexact) probability that has lead to the Ellsberg's paradox [16]. The problem of inexact probabilities arises from the structure that the possibility space from which the probability space is constructed is unknown and hence generates vagueness or ambiguities when the two spaces are connected [13] [15] [16] [50] [62].

6.3.2 Stochastic-Fuzzy Optimization

There is an alternative conceptual approach to the order of separability or non-separability of the uncertainty space. Instead of viewing outcomes as prior stochastic and posterior fuzzy, outcomes may be viewed as prior fuzzy and posterior stochastic. Given prior fuzzy uncertainty, the probability distribution is established over estimates of degree of fuzziness as expressed by membership

function or possibility values. Here the best outcome is a weighted average degree of fuzziness $\mu^{*}(\bullet)$ where the weights are the corresponding probabilities and corresponding to the average is a probability value $p(\mu^{*})$. Here, we obtain the Professor Zadeh's specification of probability of random fuzzy set, $\tilde{\tilde{\mathbb{F}}}$ as:

$$p\left(\tilde{\tilde{\mathbb{F}}}\right)=\int_{\Psi}\mu_{\tilde{\tilde{\mathbb{F}}}}(\psi)dp(\psi),\quad \tilde{\tilde{\mathbb{F}}}\subset\Psi \tag{18}$$

The probability of random fuzzy variable is defined in terms of the expected value of its membership function, $\mu_{\tilde{\tilde{\mathbb{F}}}}(\bullet)$ [57] [61] [119]. In this case, the fuzzy topological space is the subject of randomization.

To obtain the conditions of stochastic-fuzzy optimality and the supporting set of optimal rules, let us consider a random fuzzy goal, $\tilde{\tilde{\mathbb{A}}}$ whose membership function is $\mu_{\tilde{\tilde{\mathbb{A}}}}(\psi)$ and probability density is $f(\psi)$. The probability of the random fuzzy goal may be specified as either continuous or discrete process respectively as:

$$\left.\begin{aligned} p\left(\tilde{\tilde{\mathbb{A}}}|\psi\right)&=\int_{\Psi}\mu_{\tilde{\tilde{\mathbb{A}}}}(\psi)f(\psi)d\psi=\mathrm{E}\left[\mu_{\tilde{\tilde{\mathbb{A}}}}(\psi)\right]\\ &\text{or}\\ p\left(\tilde{\tilde{\mathbb{A}}}\,|\,\psi\right)&=\sum_{\psi\in\Psi}\mu_{\tilde{\tilde{\mathbb{A}}}}(\psi)\,f(\psi)=\mathrm{E}\left[\mu_{\tilde{\tilde{\mathbb{A}}}}(\psi)\right]\end{aligned}\right\} \tag{19}$$

Furthermore, let a random fuzzy constraint be $\tilde{\tilde{\mathbb{B}}}$ with a membership function, $\mu_{\tilde{\tilde{\mathbb{B}}}}(\psi)$ and probability density function $g(\psi)$. Here the optimizing variable is fuzzy, such as large or small, however, the outcome is random. For example, find an output level that will optimize expected large profit subject to a reasonable cost given the conditions of the factor markets. This example is different from raising the question whether the optimized profit outcome is large and the cost is reasonable. In the first case, the fuzziness is prior to randomness while in the second case the randomness is prior to fuzziness. The probability of the random-fuzzy constraint may be written continuously and discretely as:

$$\left.\begin{aligned} p\left(\tilde{\tilde{\mathbb{B}}}\,|\,\psi\right)&=\int_{\Gamma}\mu_{\tilde{\tilde{\mathbb{B}}}}(\psi)g(\psi)d\psi=\mathrm{E}\left[\mu_{\tilde{\tilde{\mathbb{B}}}}(\psi)\right]\\ &\text{or}\\ p\left(\tilde{\tilde{\mathbb{B}}}\,|\,\psi\right)&=\sum_{\psi\in\Gamma}\mu_{\tilde{\tilde{\mathbb{B}}}}(\psi)\,g(\psi)=\mathrm{E}\left[\mu_{\tilde{\tilde{\mathbb{B}}}}(\psi)\right]\end{aligned}\right\} \tag{20}$$

Eqns. (19) and (20) may then be combined to specify the expected random-fuzzy optimizing problem, $\tilde{\tilde{\Delta}} = \tilde{\tilde{\mathbb{A}}} \bigcap \tilde{\tilde{\mathbb{B}}}$ with membership function, $\mu_{\tilde{\tilde{\Delta}}}(\psi) = \mu_{\tilde{\tilde{\mathbb{A}}}}(\psi) \wedge \mu_{\tilde{\tilde{\mathbb{B}}}}(\psi)$ as:

$$\left.\begin{aligned} \mathrm{E}\left(\tilde{\tilde{\Delta}}\right) &= \int_{\tilde{\Psi} \cap \tilde{\Gamma}} \left\{ \left[\mu_{\tilde{\tilde{\mathbb{A}}}}(\psi) f(\psi) \right] \wedge \left[\mu_{\tilde{\tilde{\mathbb{B}}}}(\psi) g(\psi) \right] \right\} d\psi \\ &= \mathrm{E}\left[\mu_{\tilde{\tilde{\Delta}}}(\psi) \right] = \mathrm{E}\left[\mu_{\tilde{\tilde{\mathbb{A}}}}(\psi) \wedge \mu_{\tilde{\tilde{\mathbb{B}}}}(\psi) \right] \end{aligned} \right\} \tag{21}$$

The stochastic-fuzzy problem to be optimized is then abstracted from the probability of random-fuzzy variable as optimization of fuzzy integral objective function subject to a fuzzy integral constraint. This may be presented as:

$$\left.\begin{aligned} &\underset{\psi \in \Psi}{\mathrm{opt}} \int_{\Psi} \varphi_0(\mathbb{A}) \mathrm{E}\left[\mu_{\mathbb{A}}(\psi) \right] d\psi \\ &\text{s.t.} \quad \int_{\Psi} \varphi_1(\mathbb{B}) \mathrm{E}\left[\mu_{\mathbb{B}}(\psi) \right] d\psi \le \bar{k}, \\ &\qquad\qquad \psi \ge 0 \end{aligned} \right\} \tag{22}$$

where $\bar{k}$ is a known constant $\varphi_1(\bullet)$ is monotonically increasing concave function of constraint and $\varphi_0(\bullet)$ is a monotonically increasing convex function of goal with , $\varphi_0(0) = \varphi_1(0) = 0$, $\varphi_0(\mathsf{A}) = \mathbb{A}$, $\varphi_1(\mathsf{A}) = \mathbb{B}$ and A is an action set. The optimal decision under prior fuzzy and posterior stochastic may be abstracted from eqn. (22) as

$$\sigma^* = \sigma\left(\psi^* \mid \theta, \mu\left(\psi^*\right), p\left(\psi^*\right)\right) = \underset{\psi \in \tilde{\Psi}}{\mathrm{opt}}\, \sigma\left(\psi \mid \theta, \mu(\bullet), p(\bullet)\right) \tag{23}$$

and corresponding to ψ^* are $\mu_{\tilde{\Delta}}\left(\psi^*\right)$ and $p\left(\mu_{\tilde{\Delta}}\left(\psi^*\right)\right)$ which together constitute the set of optimal decision rules under non-separable uncertainty with prior fuzzy and stochastic posterior.

The supporting set of optimal decision-choice rules of stochastic-fuzzy optimality reveals itself in terms of three basic set of optimal values of decision-choice element, ψ^*, expected confidence index, $\mu_{\tilde{\Delta}}\left(\psi^*\right)$ measuring the degree of exactness of the optimal values with $\left[1 - \mu_{\tilde{\Delta}}\left(\psi^*\right)\right]$ measuring the degree of fuzzy uncertainty attached to the optimality conditions. The degree of known knowledge attached to the set of optimality conditions is measured as the

probability $p\left(\mu_{\tilde{\Delta}}(\bullet)\right)$ and the random uncertainty due to ignorance attached to the stochastic-fuzzy optimality is $\left[1-p\left(\mu_{\tilde{\Delta}}(\bullet)\right)\right]$. Here the probabilistic characterization of random uncertainty is known and the fuzzy membership function is subjectively selected or constructed. The structure presents us with a concept of fuzzy-random variable for the decision-choice process. The behavior of the fuzzy-random variable and the introduction to the mathematics to deal with it is also provided in [57] [61][84] [87] [105] [108].

7 Similarities and Differences in the Optimizing Structures

Four different structures of optimization problems have been presented. The objective is to show the evolutionary process of optimality from the classical system with exact symbolism and classical laws of thought to the fuzzy system with vague symbolism and fuzzy laws of thought and their relationships to equilibrium, rationality and stability. The objective in this essay is a focus on the nature of fuzzy optimization, areas of success and areas of challenges for further scientific research on fuzzy phenomenon. The four sets of optimization problems are represented in table 1.

The essential differences and similarities are better revealed by both problem structure with the needed algorithms and the solution structures. To illustrate with the case of the solution structure, let $\mathbb{X}$ be a general set of variables with generic element $x \in \mathbb{X}$, $\mathbb{P}$, a set of probabilities that is taken to be a set of measures of degrees of knowledge completeness, with generic element $p(x) \in \mathbb{P}$ and $p(x) \in [0,1]$ and $\mathbb{M}$, a set of membership functions that measure the degrees of exactness in the system's behavior with $\mu(x) \in \mathbb{M}$ and $\mu(x) \in [0,1]$. Finally, let S be a set of optimal solutions with generic element $\sigma^* \in \mathsf{S}$. With these values, we can provide a table of taxonomy of the optimal solution sets to all the four areas of optimization problems to illustrate their differences and similarities.

Figures 4 and 5 provide us with some reflective conditions for comparative analyses of the differences and similarities between optimization in the classical system and optimization in the fuzzy system. These similarities and differences point to opportunities and challenges of research on development of problem definitions for fuzzy optimization and construction of algorithms and theorems for their solutions and analysis. There are two important conceptual elements and measurements in the whole development of optimization and its relationship to rationality, equilibrium and stability. The first one is *information-knowledge incompleteness* that generates *stochastic uncertainty* which then creates *probabilistic belief* that is measured by a *probability value,* $p(x) \in [0,1]$ irrespective of how this probability value is constructed. The second one is the

CLASSICAL NON-FUZZY AND NON-STOCHASTIC OPTIMIZATION **Variable:** ***classical exact non-stochastic Variable.*** **Logic:** ***Classical laws of thought*** **Mathematics:** ***Classical Mathematics*** **Algorithms:** ***Classical optimizing Algorithms*** **Information-knowledge structure:** ***Free of all uncertainties.***	**CLASSICAL NON-FUZZY STOCHASTIC OPTIMIZATION** **Variable:** ***Classical exact random Variable*** **Logic:** ***Classical laws of thought*** **Mathematics:** ***Classical Mathematics*** **Algorithms: Classical Optimizing Algorithms** **Information-knowledge structure:** ***Stochastic Uncertainty and free of Fuzzy Uncertainty***
FUZZY NON-STOCHASTIC OPTIMIZATION **Variable:** ***Fuzzy non-Stochastic Variable*** **Logic:** ***Fuzzy Laws of Thought*** **Mathematics:** ***Fuzzy Mathematics*** **Algorithms:** ***Fuzzy-optimizing Algorithms*** **Information-knowledge Structure:** ***Fuzzy Uncertainty and Free of Stochastic Uncertainty***	**FUZZY-STOCHASTIC / STOCHASTIC-FUZZY OPTIMIZATION** **Variable:** ***Fuzzy-Stochastic/ Stochastic-fuzzy Variable*** **Logic:** ***Fuzzy Laws of Thought*** **Mathematics:** ***Fuzzy Mathematics*** **Algorithms:** ***Fuzzy-optimizing Algorithms*** **Information-knowledge Structure:** ***Both Stochastic and Fuzzy Uncertainties***

Fig. 4 Four Categories of Optimization Problems and Conditions of their Similarities and Differences

information-knowledge inexactness that generates *fuzzy uncertainty* which then creates *possibilistic belief* and measured by *membership value* $\mu(x) \in [0,1]$ irrespective of how this value is constructed. For the discussions on the differences in probabilistic and possibilistic belief systems see[13] [16] [17][50][57] [61] [62].

The probability value is viewed as a measure of degree of knowledge associated with the optimization, while the membership value is the degree of exactness associated with knowledge. The probability value $p(x) = 1$ and membership value $\mu(x) = 1$ imply perfect information-knowledge structure with the belief that the optimal value is accurate in that it is *exact and true*. The conditions, probability value $p(x) \in (0,1)$ and membership value $\mu(x) = 1$ imply a deficient information-knowledge structure due to incompleteness and the belief that the value that has been obtained may be an inaccurate optimal value in

	NON-STOCHASTIC	STOCHASTIC
EXACT	**The CLASSICAL EXACT NON-STOCHASTIC OPTIMAL SOLUTION** $\sigma^* = \{x^*, p(x^*) = 1, \mu(x^*) = 1\}$	**The CLASSICAL EXACT STOCHASTIC OPTIMAL SOLUTION** $\sigma^* = \{x^*, p(x^*) \in [0,1], \mu(x^*) = 1\}$
INEXACT	**The FUZZY NON-STOCHASTIC OPTIMAL SOLUTION** $\sigma^* = \{x^*, p(x^*) = 1, \mu(x^*) \in [0,1]\}$	**The FUZZY-STOCHASTIC / STOCHASTIC-FUZZY OPTIMAL SOLUTION** **a)** $\sigma^* = \{x^*, p(x^*) \in [0,1], \mu(x^*) \in [0,1]\}$ **b)** $\sigma^* = \{x^*, p(x^*) \in [0,1], \mu_P(p(x^*)) \in [0,1]\}$ **c)** $\sigma^* = \{x^*, p(\mu_X(x^*)) \in [0,1], \mu_X(x^*) \in [0,1]\}$

Fig. 5 The Structures of Solution of Different Optimization Problems given their Information-knowledge Supports

the sense that, even though it is exact, it may not be true value due to limited knowledge. The situation where $p(x) = 1$ and $\mu(x) \in (0,1)$, implies a defective information-knowledge structure due to inexactness and the belief in the value that has been obtained as optimal may be inaccurate in sense that even though it is information-knowledge complete may not be true due to information-knowledge ambiguities (inexactness). The conditions where $p(x) \in (0,1)$ and $\mu(x) \in (0,1)$ imply a completely defective information-knowledge structure due to both information-knowledge incompleteness and inexactness (ambiguities) and hence the belief in the value that has been obtained may be inaccurate in that it is inexact and may not be true.

The discussions that we have developed bring into focus the developments of optimization theories that have intra-category differences and inter-category differences. The differences in optimization algorithms within categories are category specific. Such intra-category differences of algorithms will depend on the assumptions imposed on the functional structures associated with the objective and constraint sets within the problem category. The inter-categorial differences in the optimization algorithms are due to the nature of information-knowledge structure imposed on the development of the optimization problem. It is the imposed assumption of information-knowledge structure that generates either exact-non-stochastic optimization process with ordinary algebraic variable or exact-stochastic optimization process with algebraic random variable or fuzzy-non-stochastic

optimization process with algebraic fuzzy variable and fuzzy-stochastic optimization process with algebraic fuzzy-random and random-fuzzy variables for the development of algorithms for computable optimization systems.

Substantial portion of research on fuzzy optimization has taken place with non-stochastic fuzzy variable where the information deficiency is due to fuzzy uncertainty. Some of the algorithms developed in this environment have been applied to optimization problems in classical environment of exact and non-stochastic information-knowledge structure. The research challenges facing the fuzzy research program on optimization is in Cohort IV where there are simultaneous existence of fuzziness and randomness giving rise to fuzzy and stochastic uncertainties with corresponding fuzzy-random and random-fuzzy variables for modeling optimal computable systems under conditions of fuzziness and randomness. The development may take two forms where a system of mathematics is created for problem formulation reasoning and a system of algorithms is created to abstract solutions. In other words, we are to create an effective fuzzification-defuzzification system for the general optimization under fuzzy and stochastic uncertainties. There are some works that are helping in these directions of research on mathematical and decision-choice frameworks under the phenomena of fuzziness, randomness and their simultaneities [31] [37] [50], [57][61],[64] [84] [88] [89] [90] [91] [92], [108] [110].

By the way of emphasis, let us note that the essential core of fuzzy paradigm is composed of changes in the two important principles: the principle of symbolic representation, and the principle of reasoning in the classical paradigm for information-knowledge production.

1) There is the replacement of classical exact symbolic representation of ideas and propositions that show themselves as exact classical algebraic variables with fuzzy symbolic representation of ideas and propositions that show themselves as fuzzy variables.

2) There is the replacement of classical laws of thought that: *'All propositions are either true or false'* with its principle of excluded middle, by fuzzy laws of thought that *"every statement is a set of true-false characteristics in varying proportion and that the acceptance of a proposition to be true or false is done on the basis of subjective decision-choice action in reconciling the conflict in the true-false proportion"*. with the acceptance principle of contradiction in true-false claims thus rejecting the principle of the excluded middle.

The importance of these classical-fuzzy analytics as has been stated above is seen in ordering and reasoning with items of sets The process of replacing the classical *exact symbols* with *fuzzy symbols* is *fuzzification*. The process of replacing the *classical laws of thought* with *fuzzy laws of thought* is *defuzzification*. This fuzzification-defuzzification technique must be revealed in the fuzzy optimization process in terms of problem formulation and solution abstracted for the selection of the best. The optimal solution, in terms of rationality and equilibrium, is such that the fuzzy optimal rationality is a classical suboptimal rationality and provides a best mathematical approach for models of satisficing (bounded rationality) or reasonability or levels of aspirations that are common in psychological approaches to decision theory.

References

[1] Black, Max, The Nature of Mathematics, Totowa, Littlefield, Adams and Co. (1965)
[2] Boolos, G.S., Jeffrey, R.C.: Computability and Logic. Cambridge University Press, New York (1989)
[3] Bose, R.K., Sahani, D.: Fuzzy Mappings and Fixed Point Theorems. Fuzzy Sets and Systems 21, 53–58 (1987)
[4] Buckley, J.J.: Fuzzy Programming And the Pareto Optimal Set. Fuzzy Set and Systems 10, 57–63 (1983)
[5] Butnariu, D.: Fixed Points for Fuzzy Mappings. Fuzzy Sets and Systems 7, 191–207 (1982)
[6] Carlsson, G.: Solving Ill-Structured Problems Through Well Structured Fuzzy Programming. In: Brans, J.P. (ed.) Operation Research, vol. 81, pp. 467–477. North-Holland, Amsterdam (1981)
[7] Carlsson, C.: Tackling an AMDM - Problem with the Help of Some Results From Fuzzy Set Theory. European Journal of Operational Research 10(3), 270–281 (1982)
[8] Cerny, M.: Fuzzy Approach to Vector Optimization. Intern. Jour. of General Systems 20(1), 23–29
[9] Chang, S.S.: Fixed Point Theorems for Fuzzy Mappings. Fuzzy Sets and Systems 17, 181–187 (1985)
[10] Chang, S.S.L.: Fuzzy Dynamic Programming and the Decision Making Process. In: Proc. 3rd Princeton Conference on Information Science and Systems, Princeton, pp. 200–203 (1969)
[11] Churchman, C.W.: Prediction and Optimal Decision. Pretice-Hall, Englewood Cliffs (1961)
[12] Cohn, D.L.: Measure Theory. Birkhauser, Boston (1980)
[13] Delgado, M., et al.: On the Concept of Possibility-Probability Consistency. Fuzzy Sets and Systems 21(3), 311–318 (1987)
[14] DiNola, A., et al. (eds.): The Mathematics of Fuzzy Systems. Verlag TUV Rheinland, Koln (1986)
[15] Dompere, K.K.: Epistemic Foundations of Fuzziness: Unified Theories on Decision-Choice Processes. Studies in Fuzziness and Soft Computing, vol. 236. Springer, New York (2009)
[16] Dompere, K.K.: Fuzziness and Approximate Reasoning: Epistemics on Uncertainty, Expectations and Risk in Rational Behavior. Studies in Fuzziness and Soft Computing, vol. 237. Springer, New York (2009)
[17] Dompere, K.K.: Fuzzy Rationality: Methodological Critique and Unity of Classical, Bounded and Other Rationalities. Studies in Fuzziness and Soft Computing, vol. 235. Springer, New York (2009)
[18] Dompere, K.K.: Cost-Benefit Analysis and the Theory of Fuzzy Decision: Identification and Measurement Theory. Springer, New York (2004)
[19] Dompere, K.K.: Cost-Benefit Analysis and the Theory of Fuzzy Decision: Fuzzy Value Theory. Springer, New York (2004)
[20] Delbaen, F.: Stochastic Preferences and General Equilibrium Theory. In: Dreze, J.H. (ed.) Allocation Under Uncertainty: Equilibrium and Optimality. Wiley, New York (1974)
[21] Dubois, D., et al.: Systems of Linear Fuzzy Constraints. Fuzzy Sets and Systems 3(1), 37–48 (1980)
[22] Dubois, D.: An Application of Fuzzy Arithmetic to the Optimization of Industrial Machining Processes. Mathematical Modelling 9(6), 461–475 (1987)

[23] Dumitrescu, D.: Entropy of a Fuzzy Process. Fuzzy Sets and Systems 55(2), 169–177 (1993)
[24] Eaves, B.C.: Computing Kakutani Fixed Points. Journal of Applied Mathematics 21, 236–244 (1971)
[25] El Rayes, A.B., et al.: Generalized Possibility Measures. Information Sciences 79, 201–222 (1994)
[26] Ernst, G.C., et al.: Principles of Structural Equilibrium, a Study of Equilibrium Conditions by Graphic, Force Moment and Virtual Displacement (virtual work). University of Nebraska Press, Lincoln Na. (1962)
[27] Feng, Y.J.: A Method Using Fuzzy Mathematics to Solve the Vector Maxim Problem. Fuzzy Set and Systems 9(2), 129–136 (1983)
[28] Fischer, R.B., Peters, G.: Chemical Equilibrium. Saunders Pub., Philadelphia (1970)
[29] Fisher, F.M.: Disequilibrium Foundations of Equilibrium Economics. Cambridge University Press, New York (1983)
[30] Foster, D.H.: Fuzzy Topological Groups. Journal of Math. Analysis and Applications 67, 549–564 (1979)
[31] Goetschel Jr., R., et al.: Topological Properties of Fuzzy Number. Fuzzy Sets and Systems 10(1), 87–99 (1983)
[32] Fourgeaud, C., Gouriéroux, C.: Learning Procedures and Convergence to Rationality. Econometrica 54, 845–868 (1986)
[33] Freeman, A., Carchedi, G. (eds.): Marx and Non-Equilibrium Economics, Cheltenham, UK, Edward Elgar (1996)
[34] Gaines, B.R.: Fuzzy and Probability Uncertainty logics. Information and Control 38(2), 154–169 (1978)
[35] Newton, G., Hare, P.H. (eds.): Naturalism and Rationality. Prometheus Books, Buffalo (1986)
[36] Ginsburgh, V.: Activity Analysis and General Equilibrium Modelling. North-Holland, New York (1981)
[37] Grabish, M., et al.: Fundamentals of Uncertainty Calculi with Application to Fuzzy Systems. Kluwer, Boston (1994)
[38] Hahn, F.: Equilibrium and Macroeconomics. MIT Press, Cambridge (1984)
[39] Hamacher, H., et al.: Sensitivity Analysis in Fuzzy Linear Programming. Fuzzy Sets and Systems 1, 269–281 (1978)
[40] Hannan, E.L.: On the Efficiency of the Product Operator in Fuzzy Programming with Multiple Objectives. Fuzzy Sets and Systems 2(3), 259–262 (1979)
[41] Hannan, E.L.: Linear Programming with Multiple Fuzzy Goals. Fuzzy Sets and Systems 6(3), 235–248 (1981)
[42] Hansen, B.: A Survey of General Equilibrium Systems. McGraw-Hill, New York (1970)
[43] Heilpern, S.: Fuzzy Mappings and Fixed Point Theorem. Journal of Mathematical Analysis and Applications 83, 566–569 (1981)
[44] Hestenes, M.R.: Optimization Theory: The Finite Dimensional Case, New York (1975)
[45] Heyman, D.P., Sobel, M.J.: Stochastic Models in Operations Research. Stochastic Optimization, vol. II. Mc Graw-Hill, New York (1984)
[46] Ulrich, H.: A Mathematical Theory of Uncertainty. In: Yager, R.R. (ed.) Fuzzy Set and Possibility Theory: Recent Developments, pp. 344–355. Pergamon, New York (1982)
[47] Ignizio, J.P., et al.: Fuzzy Multicriteria Integer Programming via Fuzzy Generalized Networks. Fuzzy Sets and Systems 10(3), 261–270 (1983)
[48] Jarvis, R.A.: Optimization Strategies in Adaptive Control: A Selective Survey. IEEE Trans. Syst. Man. Cybernetics SMC-5, 83–94 (1975)

[49] Kabbara, G.: New Utilization of Fuzzy Optimization Method. In: Gupta, M.M., et al. (eds.) Approximate Reasoning In Decision Analysis, pp. 239–246. North Holland, New York (1982)
[50] Kacprzyk, J., Fedrizzi, M. (eds.): Combining Fuzzy Imprecision with Probabilistic Uncertainty in Decision Making. Plenum Press, New York (1992)
[51] Kacprzyk, J., et al. (eds.): Optimization Models Using Fuzzy Sets and Possibility Theory. D. Reidel, Boston (1987)
[52] Kakutani, S.: A Generalization of Brouwer's Fixed Point Theorem. Duke Mathematical Journal 8, 416–427 (1941)
[53] Kaleva, O.: A Note on Fixed Points for Fuzzy Mappings. Fuzzy Sets and Systems 15, 99–100 (1985)
[54] Kandel, A.: On Minimization of Fuzzy Functions. IEEE Trans. Comp. C-22, 826–832 (1973)
[55] Kandel, A.: Comments on Minimization of Fuzzy Functions. IEEE Trans. Comp. C-22, 217 (1973)
[56] Kandel, A.: On the Minimization of Incompletely Specified Fuzzy Functions. Information, and Control 26, 141–153 (1974)
[57] Kaufmann, A., Gupta, M.M.: Introduction to fuzzy arithmetic: Theory and applications. Van Nostrand Rheinhold, New York (1991)
[58] Kaufmann, A.: Introduction to the Theory of Fuzzy Subsets, vol. 1. Academic Press, New York (1975)
[59] Kaufmann, A.: The Science of Decision-Making: Introduction to Praxeology. McGraw-Hill, New York (1968)
[60] Kickert, W.J.M.: Organization of Decision-Making: A Systems-Theoretical Approach. North-Holland, Amsterdam (1980)
[61] Klir, G.J., et al.: Fuzzy Sets, Uncertainty and Information. Prentice Hall, Englewood Cliffs (1988)
[62] Klir, G.J., et al.: Probability-Possibility Transformations: A Comparison. Intern. Jour. of General Systems 21(3), 291–310 (1992)
[63] Klir, G.J.: Where Do we Stand on Measures of Uncertainty. Ambignity, Fuzziness and the like? Fuzzy Sets and Systems 24(2), 141–160 (1987)
[64] Kosko, B.: Fuzziness vs Probability. Intern. Jour. of General Systems 17(1-3), 211–240 (1990)
[65] Lai, Y., et al.: Fuzzy Mathematical Programming. Springer, New York (1992)
[66] Lee, E.S., et al.: Fuzzy Multiple Objective Programming and Compromise Programming with Pareto Optimum. Fuzzy Sets and Systems 53(3), 275–288 (1993)
[67] Lowen, R.: Connex Fuzzy Sets. Fuzzy Sets and Systems 3, 291–310 (1980)
[68] Luhandjula, M.K.: Compensatory Operators in Fuzzy Linear Programming with Multiple Objectives. Fuzzy Sets and Systems 8(3), 245–252 (1982)
[69] Luhandjula, M.K.: Linear Programming Under Randomness and Fuzziness. Fuzzy Sets and Systems 10(1), 45–54 (1983)
[70] McKenzie, L.W.: Classical General Equilibrium Theory. MIT Press, Cambridge (2002)
[71] Negoita, C.V., et al.: Fuzzy Linear Programming and Tolerances in Planning. Econ. Group Cybernetic Studies 1, 3–15 (1976)
[72] Negoita, C.V., Stefanescu, A.C.: On Fuzzy Optimization. In: Gupta, M.M., et al. (eds.) Approximate Reasoning In Decision Analysis, pp. 247–250. North Holland, New York (1982)
[73] Negoita, C.V.: The Current Interest in Fuzzy Optimization. Fuzzy Sets and Systems 6(3), 261–270 (1981)

[74] Negoita, C.V., et al.: On Fuzzy Environment in Optimization Problems. In: Rose, J., et al. (eds.) Modern Trends in Cybernetics and Systems, pp. 13–24. Springer, Berlin (1977)
[75] Nguyen, H.T.: Random Sets and Belief Functions. Jour. of Math. Analysis and Applications 65(3), 531–542 (1978)
[76] Orlovsky, S.A.: On Programming with Fuzzy Constraint Sets. Kybernetes 6, 197–201 (1977)
[77] Orlovsky, S.A.: On Formulation of General Fuzzy Mathematical Problem. Fuzzy Sets and Systems 3, 311–321 (1980)
[78] Ostasiewicz, W.: A New Approach to Fuzzy Programming. Fuzzy Sets and Systems 7(2), 139–152 (1982)
[79] Patty Wayne, C.: Foundations of Topology. PWS Pub. Co., Boston (1993)
[80] Parthasarath, K.R.: Probability Measure on Metric Spaces. Academic Press, New York (1967)
[81] Ponsard, G.: Partial Spatial Equilibra With Fuzzy Constraints. Journal of Regional Science 22(2), 159–175 (1982)
[82] Prade, H., et al.: Representation and Combination of Uncertainty with belief Functions and Possibility Measures. Comput. Intell. 4, 244–264 (1988)
[83] Prade, M.: Operations Research with Fuzzy Data. In: Want, P.P., et al. (eds.) Fuzzy Sets, pp. 155–170. Plenum, New York (1980)
[84] Puri, M.L., et al.: Fuzzy Random Variables. Jour. of Mathematical Analysis and Applications 114(2), 409–422 (1986)
[85] Ralescu, D.: Optimization in a Fuzzy Environment. In: Gupta, M.M., et al. (eds.) Advances in Fuzzy Set Theory and Applications, pp. 77–91. North-Holland, New York (1979)
[86] Ralescu, D.A.: Orderings, Preferences and Fuzzy Optimization. In: Rose, J. (ed.) Current Topics in Cybernetics and Systems. Springer, Berlin (1978)
[87] Rao, M.B., et al.: Some Comments on Fuzzy Variables. Fuzzy Sets and Systems 6(2), 285–292 (1981)
[88] Rao, N.B., Rashed, A.: Some Comments on Fuzzy Random Variables. Fuzzy Sets and Systems 6(3), 285–292 (1981)
[89] Ralescu, D.: Toward a General Theory of Fuzzy Variables. Jour. of Math. Analysis and Applications 86(1), 176–193 (1982)
[90] Stein, N.E., Talaki, K.: Convex Fuzzy Random Variables. Fuzzy Sets and Systems 6(3), 271–284 (1981)
[91] Rodabaugh, S.E.: Fuzzy Arithmetic and Fuzzy Topology. In: Lasker, G.E. (ed.) Applied Systems and Cybernetics. Fuzzy Sets and Systems, vol. VI, pp. 2803–2807. Pergamon Press, New York (1981)
[92] Rodabaugh, S., et al. (eds.): Application of Category Theory to Fuzzy Subsets. Kluwer, Boston (1992)
[93] Russell, B.: Vagueness. Australian Journal of Philosophy 1, 84–92 (1923)
[94] Russell, B.: Logic and Knowledge: Essays 1901-1950. Capricorn Books, New York (1971)
[95] Sakawa, M.: Fuzzy Sets and Interactive Multiobjective Optimization. Plenum Press, New York (1993)
[96] Sakawa, M., et al.: Feasibility and Pareto Optimality for Multi-objective Nonlinear Programming Problems with Fuzzy Parameters. Fuzzy Sets and Systems 43(1), 1–15
[97] Tanaka, K., et al.: Fuzzy Programs and Their Execution. In: Zadeh, L.A., et al. (eds.) Fuzzy Sets and Their Applications to Cognitive and Decision Processes, pp. 41–76 (1974)
[98] Tanaka, K., et al.: Fuzzy Mathematical Programming. Transactions of SICE, 109–115 (1973)

[99] Tanaka, H., et al.: On Fuzzy-Mathematical Programming. Journal of Cybernetics 3(4), 37–46 (1974)
[100] Tanaka, H., et al.: Fuzzy Linear Programming, Based on Fuzzy Functions. Bulletin of Univ. of Osaka Prefecture, Series A 29(2), 113–125 (1980)
[101] Tomovic, R., Vukobratovic, M.: General Sensitivity Theory. American Elsevier, New York (1972)
[102] Torr, C.: Equilibrium, Expectations, and Information: A Study of General Theory and Modern Classical Economics. Westview Press, Boulder Colo (1988)
[103] Tsypkin, Y.Z.: Foundations of the Theory of Learning Systems. Academic Press, New York (1973)
[104] Valentinuzzi, M.: The Organs of Equilibrium and Orientation as a Control System. Hardwood Academic Pub., New York (1980)
[105] Vira, J.: Fuzzy Expectation Values in Multistage Optimization Problems. Fuzzy Sets and Systems 6(2), 161–168 (1981)
[106] Verdegay, J.L.: Fuzzy Mathematical Programming. In: Gupta, M.M., et al. (eds.) Fuzzy Information and Decision Processes, pp. 231–238. North-Holland, New York (1982)
[107] Walsh, V.C., Gram, H.: Classical and Neoclassical Theories of General Equilibrium: Historical Origins and Mathematical Structure. Oxford University Press, New York (1980)
[108] Wang, G.Y., et al.: The Theory of Fuzzy Stochastic Processes. Fuzzy Sets and Systems 51(2), 161–178 (1992)
[109] Wang, X., et al.: Fuzzy Linear Regression Analysis of Fuzzy Valued Variable. Fuzzy Sets and Systems 36(1), 19
[110] Warren, R.H.: Optimality in Fuzzy Topological Polysystems. Jour. Math. Anal. 54, 309–315 (1976)
[111] Roy, W.E.: General Equilibrium Analysis: Studies in Appraisal. Cambridge University Press, Cambridge (1985)
[112] Roy, W.E.: Microfoundations: The Compatibility of Microeconomics and Macroeconomics. Cambridge University Press, Cambridge (1980)
[113] Weiss, M.D.: Fixed Points, Separation and Induced Topologies for Fuzzy Sets. Jour. Math. Anal. and Appln. 50, 142–150 (1975)
[114] Whittle, P.: Systems in Stochastic Equilibrium. Wiley, New York (1986)
[115] Wiedey, G., Zimmermann, H.J.: Media Selection and Fuzzy Linear Programming. Journal Oper. Res. Society 29, 1071–1084 (1978)
[116] Wong, C.K.: Fuzzy Points and Local Properties of Fuzzy Topology. Jour. Math. Anal. and Appln. 46, 316–328 (1987)
[117] Yager, R.R.: Mathematical Programming with Fuzzy Constraints and Preference on the Objective. Kybernetes 8, 285–291 (1979)
[118] Zadeh, L.A.: The Role of Fuzzy Logic in the Management of Uncertainty in expert Systems. Fuzzy Sets and Systems 11, 199–227 (1983)
[119] Zadeh, L.A.: Probability Measure of Fuzzy Event. Jour. of Math. Analysis and Applications 23, 421–427 (1968)
[120] Zimmerman, H.-J.: Description and Optimization of Fuzzy Systems. Intern. Jour. Gen. Syst. 2(4), 209–215 (1975)
[121] Zimmerman, H.-J.: Fuzzy Programming and Linear Programming with Several Objective Functions. Fuzzy Sets and Systems 1(1), 45–56 (1978)
[122] Zimmerman, H.J.: Applications of Fuzzy Set Theory to Mathematical Programming. Information Science 36(1), 29–58 (1985)

Introduction to Fuzzy and Possibilistic Optimization

Weldon A. Lodwick and Elizabeth Untiedt

1 Introduction

Deterministic optimization is a normative process which extracts the best from a set of options, usually under constraints. It is arguably true that optimization is one of the most used areas of mathematical applications. It is the thesis of this book that applied mathematical programming problems should be solved predominantly by using a fuzzy and possibilistic approaches. Rommelfanger ([42], p. 295), states that the only operations research methods that is widely applied is linear programming. He goes on to state that even though this is true, of the 167 production (linear) programming systems investigated and surveyed by Fandel [18], only 13 of these were "purely" (my interpretation) linear programming systems. Thus, Rommelfanger concludes that even with this most highly used and applied operations research method, there is a discrepancy between classical linear programming and what is applied. Deterministic and stochastic optimization models require well-defined input parameters (coefficients, right-hand side values), relationships (inequalities, equalities), and preferences (real valued functions to maximize, minimize) either as real numbers or real valued distribution functions. Any large scale model requires significant data gathering efforts. If the model has projections of future values, it is clear that real numbers and real valued distributions are inadequate representations of parameters, even assuming that the model correctly captures the underlying system. It is also known from mathematical programming theory that only a few of the variables and constraints are necessary to describe an optimal solution (basic variables and active constraints), assuming a correct deterministic normative criterion (objective function). The ratio of variables that are basic and constraints that are active compared to the total becomes smaller, in general, as the model increases in size since in general large-scale models tend to become more sparse. Thus, only a few parameters need to be obtained precisely. Of course the problem is that it is not known *a priori* which variables will be basic and which constraints will be active.

W.A. Lodwick & J. Kacprzyk (Eds.): Fuzzy Optimization, STUDFUZZ 254, pp. 33–62.
springerlink.com

Moreover, a model of an actual problem is always an abbreviated view of the underlying actual system. If a problem is able to be manipulated *in situ* to obtain a solution without a symbolic representation, then there would be no need for modeling the problem mathematically. Inherently, a mathematical model is a symbolic representation of the problem. Bertrand Russell ([44], p. 85, 86) states,

> "The law of excluded middle is true when precise symbols are employed, but it is not true when symbols are vague, as, in fact, all symbols are."

At the heart of (deterministic) pure mathematics are theorems whose proofs are mathematical statements that are either true or false, but not both (law of the excluded middle). This concept is also stated as a property of classical sets - either an element belongs to a set or does not, but not both. Fuzzy set theory adds gradation to this Boolean notion of set belonging. Thus it is the rare exception that a mathematical (optimization) model is a precise representation of the underlying system even if the symbols represent real numbers or real valued distributions. It is rare that an optimal solution from an optimization model is in reality best. Herbert Simon in many places (in particular see [46], p. 35, 36) states,

> "Of course the decision that is optimal in the simplified model will seldom be optimal in the real world. The decision maker has a choice between optimal decisions for an imaginary simplified world or decisions that are 'good enough,' that *satisfice,* for a world approximating the complex real one more closely. ... What a person *cannot* do he *will not* do, no matter how much he wants to do it. Normative economics has shown that exact solutions to the larger optimization problems of the real world are simply not within reach or sight. In the face of this complexity the real-world business firm turns to procedures that find good enough answer to questions whose best answers are unknowable. Thus normative microeconomics, by showing real-world optimization to be impossible, demonstrates that economic man is in fact a satisficer, a person who accepts 'good enough' alternatives, not because he prefers less to more but because he has no choice."

From an email discussion, Rommelfanger [43] relates the following.

> "In fact Herbert Simon develops a decision making approach which he calls the *Concept of Bounded Rationality.* He formulated the following two theses. **Thesis 1**: In general a human being does not strive for optimal decisions, but he tends to choose a course of action that meets minimum standards for satisfaction. The reason for this is that truly rational research can never be completed. **Thesis 2**: Courses of alternative actions and consequences are in general not known *a priori*, but they must be found by means of a search procedure."

The central tenet of this book is that most optimization is and should be a *satisficing* process. To this end, fuzzy and possibilistic optimization play a key, if not the most important, role. Why is fuzzy/possibilistic optimization not an explicit part of every Operations Research and Optimization curriculum? Why is it not in wide-spread practice yet? Is it that the practitioners do not know these new theoretical instruments? Is it that too many university professors do not understand the new concepts? Is it that clear and compelling "industrial strength" models have yet to appear in a way to make an impact? There are some "industrial strength" models (see the application section of this volume and [31], [38]). This volume presents the compelling reasons for including fuzzy and possibilistic optimization at the heart of normative decision theory.

We first define two terms that are frequently used in conjunction with fuzzy set theory and possibility theory, "uncertainty" and "incomplete information." From [14], we have:

Definition 1. *A piece of information is said to be* ***incomplete*** *in a given context if it is not sufficient to allow the agent to answer a relevant question in this context. A piece of information is* ***uncertain*** *for an agent when the latter does not know whether this piece of information is true or false.*

2 Fuzzy Set Theory and Possibility Theory in Optimization

This section begins with an example "industrial strength" application which shows the applicability and relevance of fuzzy and possibilistic optimization in an actual application. This problem was reported in [31], [38].

Example 2. *(Radiation Therapy of Tumors - [31]) The determination of how to use particle beams to treat tumors is called the radiation therapy problem (RTP). Beams of particles, usually photons or electrons, are oriented at various angles and with varying intensities to deposit dose (energy/unit mass) to the tumor. The idea is to deposit just enough radiation to the tumor to kill all the tumor cells while sparing normal tissue. The process begins with the patient's computed tomography (CT) scan. Each CT image is examined to identify and contour the tumor and normal structures. The image subsequently is vectorized. Likewise, candidate beams are discretized into beamlets, where each beamlet is the width of a CT pixel. A pixel is the mathematical entity or structure (a square in the two-dimensional case and a cube in three dimensions) that is used to represent a unit area or volume of the body at a particular location. For two-dimensional problems, about seventeen CT scans (slices) are sequentially "stacked" (to form a three-dimensional image that covers the tumor), and a variety of resolutions might be considered,* $64 \times 64, 128 \times 128, 256 \times 256, 512 \times 512$. *One set of beams each at ten or more equally spaced angles is not unusual. Since we constrain the dosage at*

each pixel, for ten equally spaced angles, the complexity of the problem ranges from an order of $17 \cdot 10 \cdot 64^2$ *to* $17 \cdot 10 \cdot 512^2$ *potential constraints. However, since all pixels are not in the paths of the radiation beams that hit the tumor, and some are outside the body, we a priori set the delivered dosages at these pixels to zero and remove them from our analysis. This corresponds to blocking the beam, which is always done in practice. The identification of a set of beam angles and weights that provide a lethal dose to the tumor cells, while sparing healthy tissue with a resulting dose distribution acceptable and approved by the attendant radiation oncologist, is called a treatment plan. The largest actual problem solved is on the order of* $500,000$ *constraints in a fuzzy/possibilistic optimization problem. A dose transfer matrix* A^T *(representing how one unit of radiation intensity in each beamlet is deposited in pixels - for historical reasons, we use a transpose to emphasize its origin as the discrete version of the inverse Radon transform), called here the attenuation matrix, specific to the patient's geometry, is formed where columns of* A^T *correspond to the beamlets and rows represent pixels. A component of a column of the matrix* A^T *is non-zero if the corresponding beamlet intersects a pixel, in which case the value is the positive fraction of the area of the intersection of the pixels with the beamlet otherwise it is zero. The beams then are attenuated according to a factor dependent on the distance from where the beam enters the body to a pixel within the body and the type of tissue at that pixel. The variables are vectors* x *that represent the beamlet intensities. The mathematical programming problem is*

$$\begin{aligned} z &= \min f(c, x) \\ A^T x &\leq b. \end{aligned}$$

For this problem, there are four places in which the data turn this problem into a fuzzy/possibilistic optimization problem. First, the objective function can be a probability function (upper/lower bounding), the probability that the radiation intensity vector x *will turn a health pixel into a cancerous one. Each row of the left side, that is, each row of the constraint matrix* A^T*, represents each pixel in the path of the beam, the "beam's eye view" of the tumor. Thus, each row accumulates pixel by pixel total radiation deposited by the radiation intensity vector* x *at that pixel. This will occur mathematically since the* i^{th} *row vector (the* i^{th} *pixel)* A_i^T *when dotted into the vector* x *is the sum of radiation at the pixel,* $A_i^T x$*. Since a pixel can be cancerous, or cancerous to a degree (the boundaries between cancerous and non-cancerous are gradual, transitional and thus fuzzy), the left side matrix* A^T *matrix is composed of fuzzy intervals. The right-hand side value is the maximum allowable dosage (for critical organs in the path of the beam and maximal value representing pre-burning for the tumor cells). Separately, for tumor cells, there is also an associated minimal value, the smallest value a radiation oncologist does not allow the radiation to go below (it is the minimal acceptable radiation that will kill a cancer cell). These values may be considered to be possibilistic*

since these are values derived from research, expert knowledge (epistemic values), and experience with preferred values within the range. Therefore, the right-hand side value may be considered either as a target (with preferred ranges) or as purely as a possibilistic value.

A radiation oncologist is a *satisficer*. Any good treatment regime (identified by the mathematical programming problem) will suffice. Of course the optimal would be better, but given the variations in the data and model, "best" is an illusive value and arguably unattainable. The above example problem illustrates fuzzy, probability, and possibility optimization in an applied problem which has been solved (see [31], [38]).

Possibility theory as a mathematical theory is based on a set of axioms or properties found below. It is used to model systems of entities or variables that are uncertain due to *knowledge deficit* (inherent or acquired) or *incomplete information*. Fuzzy set theory deals with sets which are a generalization of the classical notion of set. It is used to model systems of entities or variables whose belonging (to a set) is *gradual* or *transitional*. Underlying possibility theory is the principle of minimal specificity (see [16]). That is, when the values of parameters or variables are not completely specified (for whatever reason - by choice, by finances, by lack of the ability to obtain the precise value, by the fact that inherently a precise value cannot be obtained), is there still sufficient structure to the information that is available for mathematical analysis? One of these mathematical structures is possibility theory.

The following simple example shows that *probability alone* is insufficient to describe uncertainty of every type. Suppose all that is known is that

$$x \in [1, 4]. \tag{1}$$

Clearly, $x \in [1, 4]$ implies that the real value that x represents is not certain (albeit bounded). If the uncertainty that $x \in [1, 4]$ represents were probabilistic (x is a random variable that lies in this interval), then every distribution having support contained in $[1, 4]$ would be equally valid given (1). Thus, if one chooses the uniform probability density distribution on $[1, 4]$,

$$p(x) = \begin{cases} \frac{1}{3} & 1 \leq x \leq 4 \\ 0 & otherwise, \end{cases}$$

which is the usual choice given no other indication other than the support, one gives up information. The approach that keeps the entire uncertainty of (1) considers it as all distributions whose support is $[1, 4]$. The pair of cumulative distributions that bound **all** cumulative distributions with this given support is depicted in Figure 1. The statement $x \in [1, 4]$ not only represents a random variable whose support is $[1, 4]$, but it can be a mathematical entity, an interval. When $[x] = [1, 4]$ is a mathematical entity, an interval, the statement $x \in [1, 4]$ has no associated uncertainty. It is complete, precise, and coherent in contradistinction with $[1, 4]$ containing all probability density

functions whose support is contained in this interval. It is true that the same object, $[1, 4]$, interpreted probabilistically and as an interval, has two semantically distinct and analytically distinct meanings, they have a different calculus, metric, convergence structure, and so on.

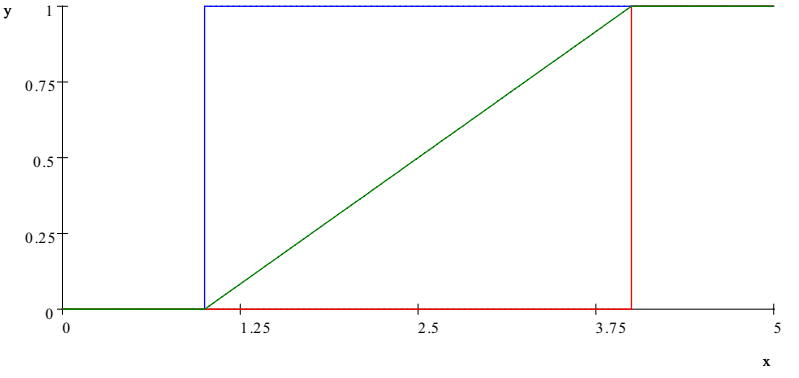


Fig. 1 BOUNDING CUMULATIVE DISTRIBUTIONS - Possibility (blue), Necessity (red), Uniform (green)

The upper cumulative distribution depicted in Figure 1 is a *possibility distribution* (blue), and the lower cumulative distribution is a *necessity distribution* (red). When the statement $x \in [1, 4]$ represents an unknown random variable whose support is the interval, to keep all the information about the uncertainty not only requires a pair of bounding functions, but a different arithmetic and mathematical analysis than "simply" functional analysis on probability distributions. The uniform distribution is precisely the intuitive solution to lack of information, "Choose the midpoint of the distribution pair as the solution *if one has to choose*." Of course, the case is made here that $x \in [1, 4]$ can also be tied to uncertainty which is purely non-probabilistic information deficiency in addition to an uncountably infinite set of random variables (whose support is contained in this interval).

2.1 Fuzzy Set Theory

Fuzzy sets are sets in which the Boolean property of belonging that characterizes classical sets is generalized to allow degrees of belonging continuously from zero (indicating not belonging for sure) to one (indicating belonging for sure). A classical set is one in which every element in the universal set has a degree of belonging that is described by the characteristic function (zero/one), whereas a fuzzy set is one in which every element in the universal set is described by a function, called a *membership function*, whose range is

between zero and one. A fuzzy set $A \in \Omega$ (universal set) is uniquely described by its membership function

$$\mu_A(x) : \Omega \to [0,1], x \in \Omega. \tag{2}$$

A classical set is one whose membership function is,

$$\mu_A(x) = \begin{cases} 1 & x \in A \\ 0 & x \notin A. \end{cases} \tag{3}$$

The general definition of a fuzzy set (2) does not impose any condition except that the fuzzy membership be a function. In the context of optimization, we will restrict ourselves to fuzzy numbers and fuzzy intervals, and the membership function must be upper/lower semi-continuous (or more practically continuous), where at least one value of the domain has a membership value of one. In contrast to probability for finite spaces (see discussion in [16]), for an event $x \in A$, "... $prob(A)$ is the probability that an ill-known single-valued variable x ranging in Ω hits the fixed well-defined set A." For example, suppose we are rolling two die, and $A = \{4, 5\}$. The probability $prob(A) = \frac{7}{36}$. Here x, which ranges in $\Omega = \{2, 3, ..., 12\}$, is unknown (it is the outcome of a throw of the dice), but the set A is precisely known. When we consider $\mu_A(x)$, then x is "fixed", known precisely, whereas the set A is ill-defined (transitional).

> [8] A fuzzy set (2) is an abstract mathematical notion. It presupposes nothing about what it could be applied to the notion. It corresponds to the ideal of non-Boolean predicates. The idea of gradualness or grade (non-Boolean) is opposed to "all or nothing" (Boolean). Thus, fuzzy means *gradual* and not vague or uncertain. So, consider a set. Some sets are real entities, for example, the set of all salaried tenured professors who are employed by the Mathematical and Statistical Sciences Department at the University Colorado Denver during the 2008-2009 academic year or the set of *older* mathematicians. Some of these real sets are fuzzy, for example, the set of older mathematicians, since the concept of "older" is gradual, graded. These (fuzzy) sets are taken as "lumped" entities (*older* is a "lumped" entity). They are a *conjunctive or linked* set of elements having more or less weight.

Let us look again at our initial example. The interval $[1, 4]$ considered as a fuzzy set has membership function

$$\mu_{[1,4]}(x) = \begin{cases} 1 & x \in [1,4] \\ 0 & otherwise. \end{cases}$$

Considered as a fuzzy set, the elements of $[1, 4]$ are "lumped" or linked into one entity, an interval. This intertwined or conjunctive property characterizing fuzzy sets is contrasted with the "mutually exclusive", unlinked, or

disjunctive property that characterizes possibility theory and probability theory, as we shall discuss further below.

Mathematical analysis on fuzzy sets (here we are thinking of fuzzy sets as abstract mathematical entities, that is, sets of membership functions ([2])) require a measure, a fuzzy measure, which is used to define order, extent, convergence, integrals (transformation of a function into a number), and so on. Fuzzy measures and integrals are used in fuzzy logic to compare/order and to "defuzzify" (turn a fuzzy set into a number and thus an action - "shift into overdrive" for example). This is akin to determining the expected value for continuous distributions in probability theory that transforms a (probability distribution) function into a single real number (the mean).

Definition 3. *[30] Given a universal set Ω and a nonempty family $\mathcal{F}$ of subsets of Ω, a* ***fuzzy measure*** *on $\langle \Omega, \mathcal{F} \rangle$ is a function*

$$g : \mathcal{F} \to [0, 1]$$

that satisfies the following properties:
1) $g(\varnothing) = 0$, and $g(\Omega) = 1$ (boundary conditions)
2) $\forall A, B \in \mathcal{F}$, if $A \subset B$, then $g(A) \leq g(B)$ (monotonicity)
3) $\forall$ increasing sequence $A_1 \subset A_2 \subset ...$ in $\mathcal{F}$, if $\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$, then

$$\lim_{i \to \infty} g(A_i) = g\left(\bigcup_{i=1}^{\infty} A_i \right) \text{ (continuity from below)}$$

4) $\forall$ decreasing sequence $A_1 \supset A_2 \supset ...$ in $\mathcal{F}$, if $\bigcap_{i=1}^{\infty} A_i \in \mathcal{F}$, then

$$\lim_{i \to \infty} g(A_i) = g\left(\bigcap_{i=1}^{\infty} A_i \right) \text{ (continuity from above).}$$

In the above, 3) and 4) can be viewed as endowing consistency to fuzzy measures. Usually, the fuzzy measures are defined over a structured set of sets such as a ring or a σ-algebra or the full power set $P(\Omega)$. Fuzzy measures may be considered as generalizations of probability measures and, in the broader sense, classical measures. The difference between the fuzzy measure and a probability measure is that the additivity condition is replaced by a weaker one of monotonicity and continuity.

Property 2) of Definition 3 implies that $g(A \cup B) \geq \max\{g(A), g(B)\}$ and $g(A \cap B) \leq \min\{g(A), g(B)\}$. A fuzzy measure is a real valued measure that is used for mathematical analysis of fuzzy sets. The semantic associated with fuzzy measure is that the assigned number $g(A)$ to the set $A \in \mathcal{F} \subset \Omega$ indicates the degree to which the given element, set A of $\mathcal{F}$, whose characterization is transitional, belongs to $\mathcal{F}$.

2.2 Possibility Theory

The mathematical models of systems which are described by the knowledge we as humans have (epistemic) about the system are often not probabilistic. Deterministic models are (supposed to be) precise statements about properties of the real system, the thing itself. They refer to the real world. Possibilistic models often refer to the knowledge humans have about the system. The model exists or refers to what is in someone's (or a group of people's) head, not the actual system. Possibility is a mathematical structure with a set of operations that allow us to model some uncertainties that are not probabilistic, uncertainties that are not random, frequencies, or chance. However, as in the interval example, they also provide a structure for probabilistic bounds which themselves are derived from (cumulative) probabilities. Even when the statement of the problem is clearly probabilistic (the probability that radiation intensity x will turn a health cell into a cancerous cell), a single probability distribution function is impossible to obtain for every cell, and for every human body, and for every bodily condition (obesity, anorexia, lean, and so on), and for every type of cancer, and every type of radiation type (electron, proton, neutron), and must be approximated. To obtain one single probability distribution function $p(x)$ is an approximation at best. It is perhaps more useful to find bounding functions and to do our analysis on these rather than committing to a single probability at the start.

Some set representations are epistemic entities, the information possessed by people about real entities. These set representations do not refer to the real world, but to what people know about the real world. Dubois uses the following example. I may "know" epistemically that the Prime Minister of France is between 40 and 55 years of age. This is the extent of my knowledge of the age of the prime minister of France, which may be wrong, and it may differ from your knowledge. The interval [40,55] refers to an interval of real ages. However, the age of Prime Minister Sarkozy, which exists, is a *point,* not a set. The set $[40, 55]$ is the set of possible ages of Prime Minister Sarkozy according to my knowledge, where the actual age of the Prime minister is a point. My interval of possible ages in this case is not a conjunction of elements, it is a disjunction of elements (the set of distinct numbers in the interval). A generalized characteristic function of this set is a possibility distribution, despite the fact that there is only one age of Prime Minister Sarkozy. If some values of ages in the interval $[40, 55]$ are more plausible than others, then this plausibility defines a preferential function on $[40, 55]$ which is a possibility distribution.

Possibility is also an abstract mathematical structure which is independent of applications.

Definition 4. *[55] Let* $\mathbf{P}(\mathbf{\Omega})$ *denote the power set of the universal set* Ω *(although a* σ*-algebra would also work). A* ***possibility measure*** *Pos:*$\mathbf{P}(\mathbf{\Omega}) \rightarrow$ *[0,1] has the following properties:*

1) Pos(∅)=0, Pos(Ω)=1
2) $A \subset B \Rightarrow Pos(A) \leq Pos(B)$, *for* $A, B \in \Omega$.
3) $Pos(\bigcup_{i \in I} A_i) = \sup_{i \in I} Pos(A_i)$.

Possibility is a non-additive measure like fuzzy measures, but it differs from a fuzzy measure since the two continuity conditions are replaced by the supremum condition on the union. Moreover, property 3) of Definition 4 implies that

$$Pos(A \cup B) = \max\{Pos(A), Pos(B)\}.$$

As an abstract structure, it also differs from a probability measure because the additivity condition is replaced by the supremum condition. Besides defining possibility via the three properties given above, possibility measures also can be constructed beginning from probability in four different ways one of which as was given by our first example [1, 4]. The other three ways are presented in the sequel.

Any possibility measure is determined uniquely by a given possibility (distribution) function

$$\pi : \Omega \to [0, 1]$$

via the formula

$$Pos(A) = \sup_{x \in A} \pi(x), A \subset \Omega. \tag{4}$$

Also, given a possibility measure, we can define a possibility (distribution) function $\pi : \Omega \to [0, 1]$

$$\pi(x) = Pos(\{x\}). \tag{5}$$

That is, we can go from a measure to a possibility distribution function in a natural way. Moreover, possibility measures are distinct from fuzzy measures (see [41]).

It turns out that Dempster-Shafer's plausibility and belief functions (see [6] and [45]) using an auxiliary probability assignment function with the requirement that the focal elements (elements whose probability assignment functions are non-zero) be nested. When this occurs, the resultant is a possibility measure. Thus, possibility theory can be considered as a particular case of Dempster-Shafer theory. To see this, consider a finite universal set Ω, for the sake of simplicity, and where $P(\Omega)$ denotes the power set of Ω. We restate possibility and necessity in this context (finite universe) and call a function $Nec : P(\Omega) \to [0, 1]$ a *necessity measure* if and only if

$$\begin{aligned} Nec(\emptyset) &= 0 \\ Nec(\Omega) &= 1 \\ Nec(A \cap B) &= \min\{Nec(A), Nec(B)\}, \end{aligned} \tag{6}$$

for all $A, B \in P(\Omega)$. A function $Pos : P(\Omega) \rightarrow [0,1]$ is called a *possibility measure* if and only if,

$$\begin{aligned} Pos(\emptyset) &= 0 \\ Pos(\Omega) &= 1 \\ Pos(A \cup B) &= \max\{Pos(A), Pos(B)\}, \end{aligned} \tag{7}$$

for all $A, B \in P(\Omega)$. Note that (7) implies that if $A \subseteq B, Pos(A) \leq Pos(B)$. Moreover,

$$Nec(A) = 1 - Pos(A^c).$$

The relationship between necessity and Dempster-Shafer theory is as follows. Evidence theory is based on two non-additive measures on a universal set Ω, a belief measure, Bel, and a plausibility measure, Pl, defined as (see [30])

$$\begin{aligned} Bel \ &: \ P(\Omega) \rightarrow [0,1] \\ Bel(\emptyset) &= 0, Bel(\Omega) = 1, \\ Bel(\bigcup_{i=1}^{n} A_i) &\geq \sum_i Bel(A_i) - \sum_{i<j} Bel(A_i \cap A_j) + ... \\ &\quad +(-1)^{n+1} Bel(\bigcap_{i=1}^{n} A_i) \end{aligned} \tag{8}$$

for all possible families of subsets of Ω. Belief is superadditive. The dual of belief is plausibility which is defined

$$Pl(A) = 1 - Bel(A^C). \tag{9}$$

A function $m : P(\Omega) \rightarrow [0,1]$ is called a *probability assignment function* if and only if

$$\begin{aligned} m(\emptyset) &= 0 \\ \sum_{A \in P(\Omega)} m(A) &= 1. \end{aligned} \tag{10}$$

Lemma 5. *[45] Given a probability assignment function, a Belief and Plausibility measure on Ω can be defined*

$$Bel(A) = \sum_{B \subseteq A} m(B) \tag{11}$$

for all $A \in P(\Omega)$, and

$$Pl(A) = \sum_{A \cap B \neq \emptyset} m(B). \tag{12}$$

Conversely, given a belief measure, a probability assignment function can be defined by

$$m(A) = \sum_{B \subseteq A} (-1)^{|A-B|} Bel(B), \tag{13}$$

where $|A - B|$ *denotes the cardinality of* $A - B$*, for all* $A, B \in P(\Omega)$.

Essentially, starting from probability, we start from (10) where the probabilities are known on *sets* of the space, **not** elements of the space, with no particular assumed structure (they may be overlapping, for example) other than that the sums of their probabilities add to one. From this partial information, we can derive a mathematical structure which does not, in general, have the additive property of probabilities which is called Belief and Plausibility. On the other hand, we can start from the point of view that we have a sub/super additive set functions called Belief and Plausibility (whose properties are is defined by (8) and (9)) and derive a probability assignment function from (13). If the assignment function is known on *elements* of the space, then belief is equal to plausibility, and the resulting structure is probability (with additivity).

The subsets of Ω with positive values of a given probability assignment function are called *focal elements*. Suppose all focal elements of the basic assignment functions (13) can be ordered so that they are nested. That is, suppose we have an order to the focal elements determined by (13) so that the finite sequence

$$\{A_1, A_2, ..., A_K\}$$

and $m(A_i) > 0$ has the property that $A_i \subseteq A_j$ for $i \leq j$. Then

$$Nec(A) = \sum_{B \subseteq A} m(B) \tag{14}$$

is a necessity measure. Moreover,

$$Pos(A) = \sum_{A \cap B \neq \emptyset} m(B) \tag{15}$$

is a possibility measure. In other words, (10) can be given and Bel and Pl constructed via (11) and (12). Or Bel (and Pl) can be given and the probability assignment function be constructed via (13). Moreover, when we have nested sets, we take the focal elements of the probability assignment function on these nested sets and construct the possibility and necessity which are precisely the belief and plausibility measures.

We can obtain a probability assignment function from a given necessity measure by using (13) so that

$$m(A) = \sum_{B \subseteq A} (-1)^{|A-B|} Nec(B).$$

Thus, necessity measures are belief measures when the focal elements are nested. Fuzzy intervals are the collection of nested focal elements constructed from the non-zero $\alpha - cuts$. That is, for $\alpha > 0$, in the discretized case, let

$$\begin{aligned} 0 = \alpha_0 < \alpha_1 < ... < \alpha_j < \alpha_{j+1} < ... < \alpha_N = 1, \\ m(A_{\alpha_j}) = \alpha_j - \alpha_{j-1}, j = 1, ..., N \\ A_\alpha = \{x | \mu_{fuzint}(x) \geq \alpha\}, \end{aligned}$$

where μ is the membership function of the fuzzy interval. Hence, possibility theory may be viewed as a special type of evidence theory. This allows possibility theory to be considered from a generalized probabilistic point of reference. Zadeh [55] defined only possibility measures/distributions. Dubois and Prade [11] were the first to develop the dual to possibility, necessity.

The previous section outlined several ways that lead to possibility distributions. There is a development (see [4]) which takes the possibility measure given above and develops the full mathematical structure such that the mathematical analysis is well-defined. We are thinking of possibility theory applied to mathematical analysis (optimization in particular). To this end, we construct possibility and necessity distributions in one of the following ways:

1. [37] Given a set of probabilities $\Omega = \{p_\alpha(x), x \in \mathbb{R}, \alpha \in I$, where I is an index set$\}$,

$$\begin{aligned} Pos(x) = \sup_{\alpha \in I} p_\alpha(x) \\ Nec(x) = \inf_{\alpha \in I} p_\alpha(x). \end{aligned}$$

2. [28] Given an unknown probability $p(x)$ which is known to exist inside a bounding pair of functions $p(x) \in [\underline{f}(x), \overline{f}(x)]$, construct necessity/possibility distributions such that $p(x) \in [Nec(x), Pos(x)]$.
3. [6], [45] Given a probability assignment function m whose focal elements are nested, construct necessity/possibility distributions according (15) and (14).
4. A fuzzy interval, defined below, generates a possibility and necessity pair. Fuzzy intervals, as used here, are what is called in most of the literature fuzzy numbers. The possibility and necessity functions are constructed as was done from our initial example $[1, 4]$ (also see Figure 3 below).

The most prevalent approach is to define the entities of interest in optimization (the coefficients and/or the right-hand side values, for example) to be fuzzy intervals in which case they will be able to capture both gradualness or transition and lack of specificity/information as we specify below. Thus, possibility distributions used in possibilistic optimization typically are associated with the membership function of fuzzy numbers, whose generalization is called a *fuzzy interval*. If the coefficients arise from probability-based

possibility (as in items 1-3 listed above), then this generates upper and lower possibilistic optimization (see [37]). A fuzzy number is a fuzzy set with upper/lower semi-continuous membership function with one and only one value, x^*, such that $\mu(x^*) = 1$ (where x^* is the "fuzzified" number). The set of numbers for which the membership value are one is called the *core*. So for example, a fuzzy number 2 would have $\mu(2) = 1$, which is depicted in Figure 2.

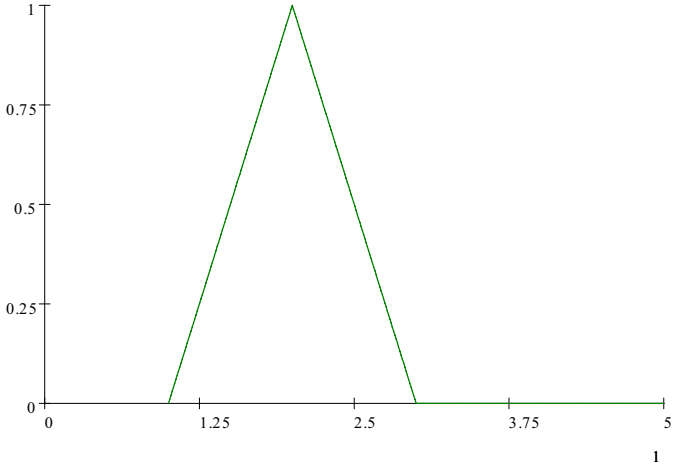


Fig. 2 A Fuzzy number two

A fuzzy interval M, depicted as a trapezoid in Figure 3, is a fuzzy number except the core (which must also exist) does not have to be a singleton. There are various views (applications) of fuzzy intervals. A fuzzy interval can be used to enclose a set of probability distributions where the bounds are constructed from the fuzzy interval (blue line being the possibility and green line being the necessity in Figure 3). The core of the fuzzy interval is the top, the horizontal line segment between 2 and 3 at height 1, indicated in Figure 3. The possibility and necessity as indicated below enclose all probabilities whose upper limit is the possibility (blue line) and lower limit is the necessity (green line). Thus, according to [15],

> "A fuzzy interval M can thus be viewed as encoding a family of probabilities, a set of probability measures P_M defined by
>
> $$P_M = \{P | \Pi_M(A) = \sup_{a \in A} M(a) \geq prob(A), A \text{ measurable}\}.$$
>
> It is important to notice that there are actually three probabilistic views of a fuzzy interval:

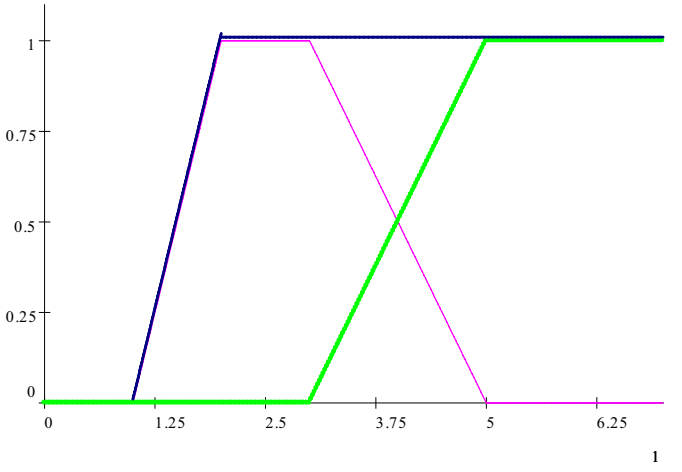


Fig. 3 Fuzzy interval - Megenta

a) The *imprecise probability* view whereby M encodes a set of probability measures shown in Figure 3.
b) The *pair of PDFs* view whereby M is defined by two random variables x^- and x^+ with cumulative distributions in blue and green of Figure 3 and M stands for the random interval $[x^-, x^+]$.
c) The *random set* view whereby M encodes the one point coverage function of a random interval, defined by the probability measure on the unit interval (for instance the uniformly distributed one) and a family of nested intervals (the $\alpha - cuts$), via a multivalued mapping from (0,1] to $\mathbb{R}$, following Dempster [6]."

A reason that one might want to use probability-based possibility (interpretations a), b) or c)) rather than probability is precisely in situations for which real values or complete probability density functions for data are not available. For example: (1) we don't know which of a given set of probabilities to use, (2) all we know is that the probability is bounded by two functions, or (3) we do not have the probability distribution on singletons, but on sets. Whether an entity of interest inherently lacks specificity (the minimal radiation that will kill a particular patient's prostate tumor cell located at (x, y, z) is $Pos(x, y, z)$), lacks sufficient research to determine its precise value or its precise probability density function, its deterministic functional representation is not required, in the sense that one can get by with a more general form than its deterministic equivalent - perfect information, for the use to which it is put (the light wave reflection measured by a satellite sensor to impute the depth of the ocean, low/medium/high might suffice), or complexity reduction (low, medium, high speed for the automatic gear shifting mechanism on a car), lack of information/specificity is a part of many if not most

problems. Moreover, when we have models that are built from epistemic knowledge (human ideas about the system rather than the system itself), many of these linguistically derived models are possibilistic either in their entirety or partially.

2.3 Fuzzy Set Theory and Possibility Theory - Distinctions

The semantics of fuzzy sets and possibility are different. The difference between fuzzy set theory and possibility theory is that the semantics of fuzzy are tied to gradualness whereas the semantics of possibility refers to incomplete (deficient) information about an existent entity. As we have mentioned above, fuzzy is a non-Boolean set whose elements transitionally belong to a given set. Possibility is tied to incomplete information. A fuzzy set membership function uniquely describes the set and defines the degree to which an element belongs to the set. The possibility distribution assigns the degree to which the evidence supports the element's belonging to the set in question. In particular, consider the two statements

$$\mu_A(x) = 1,$$

and

$$\pi_A(x) = 1.$$

In the first instance, the membership function value of x is 1. In the second, the possibility distribution at x is 1. In the first case, it is certain that x belongs to the set A. In the second case, all that can be said is that the best information at hand indicates that it is most plausible (or possible) that x belongs to A. $\mu_A(x) = 0$ means that $x \notin A$ for sure. $\pi_A(x) = 0$ means that all the given evidence at hand indicates that x does not belong to A (or x is not A).

[16] "Limited specificity can be modelled in a natural way by possibility theory. The mathematical structure of possibility theory equips fuzzy sets with set functions, conditioning tools, notions of independence/dependence, decision-making capabilities (lattices). Lack of information or lack of specificity means we do not have 'the negation of a proposition is improbable if and only if the proposition is probable.' In the setting of lack of specificity, 'the negation of a proposition is impossible if and only if the proposition is necessarily true.' Hence, in possibility theory pairs of possibility and necessity are used to capture the notions of plausibility and certainty. When pairs of functions are used we may be able to capture or model lack of information. *A membership function is a possibility only when the domain of a fuzzy set is decomposable into mutually exclusive elements.* A second difference (between probability and possibility) lies in the underlying assumption

regarding a probability distribution. Namely all values of positive probability are mutually exclusive. A fuzzy set is a conjunction of elements. For instance, in image processing, imprecise regions are often modelled by fuzzy sets. However, the pixels in the region are not mutually exclusive (although they do not overlap). Namely the region contains several pixels, not a single unknown one. *When the assumption of mutual exclusion of elements of a fuzzy set is explicitly made, then, and only then, the membership function is interpreted as a possibility distribution; this is the case of fuzzy numbers describing the ill-located unique value of a parameter.*" (my *italicized* emphases)

More recently, [14] state

"A set used for representing a piece of incomplete information is called a *disjunctive set.* It contrasts with a conjunctive view of a set considered as a collection of elements. A conjunctive set represents a precise piece of information. For instance, consider the quantity $v = sisters(Pierre)$ whose range is the set of subsets of possible names for Pierre's sisters. The piece of information $\{Marie, Sylvie\}$ is precise and means that Pierre's sisters are Marie and Sylvie. Indeed, the frame is then $S = 2^{NAMES}$, where $NAMES$ is the set of all female first names. In this setting, a piece of incomplete information would be encoded as a disjunction of conjunctive subsets of $NAMES$."

For example, an image could be segmented/classified into two sets, stomach lining and stomach muscle. Every pixel in the image is given a value v where $0 \leq v \leq 1$ with respect to being stomach lining or stomach muscle. This is conjunctive and thus a fuzzy set. That is, each pixel is stomach lining to specified degree (between 0 and 1), and (conjunction) each pixel is also stomach muscle to a specified degree (between 0 and 1). On the other hand, suppose we use the fuzzy trapezoid interval 59/59.9/60.1/61 to model the possibilistic notion of a tumorcidal dose to each tumor pixel. This is an incomplete set of information about each tumor pixel. That is, a tumor pixel has an associated *distribution*, a fuzzy interval, a function. A tumor pixel is a fuzzy interval 59/59.9/60.1/61. The pixel is not represented by a number (as in the case of stomach lining or muscle - there are two numbers one for each class), the pixel is represented by an entire distribution. Moreover, possibility is always normalized since the semantics of possibility is tied to an existential entity. Thus, *not all fuzzy set membership functions can be transformed into possibility distributions.*

2.4 Fuzzy Set Theory and Possibility Theory in Optimization

Gradualness characterizes many linguistic descriptions of what we know about properties of entities and systems of these entities. Indeed, classification of

reality, which often is the first step in mathematical modelling, may reflect the inherent transitional nature of the class as well as objects in that class. Thus, "tumor" is a classification which is "intertwined." Each "tumor" pixel within a CT scan may consist of fully cancerous cells in a pixel and (conjunctive) precancerous cells coexisting in the same pixel.

Mathematical modelling has two directions - one toward more specificity (a real value being one "end point" of specificity) or more generality (the universe being at the other end of specificity). Generality is often useful to simplify analysis and reduce the complexity. The greatest use and application of fuzzy set theory is fuzzy logic controllers. In mathematical analysis, it can be argued that fuzzy optimization has been the most successful application. Fuzzy optimization is an extension to flexible programming and allows for a broader and more ample approach to flexible programming.

The appropriate classification of *possibilistic optimization* is *optimization under uncertainty,* where some (or all) the input data (parameters) to the optimization model lack specificity, and/or the information is insufficient to yield a real valued number or a probability distribution. This is distinguished from *fuzzy optimization* that appropriately belongs in the class of *flexible programming* problems. The uncertain parameters that are possibilistic must adhere to (7), which means that more information about the parameter can never yield less certainty (if $A \subseteq B, P(A) \leq P(B)$), and these sets must be nested. This property (more information leading to greater certainty) is not axiomatically present in fuzzy measures. For example, the fuzzy set "older" remains transitionally "older," regardless of how much more information about older is obtained. However, the interval of the age of the prime minister of France will narrow with more (correct) information.

Not all systems possess this property of more information not increasing uncertainty. In competitive markets, (cold) wars, adversarial relationships (such as political campaigns, trials, or propaganda), more information might be worse given the propensity to disinformation, deceptive information, or lying all mixed together with legitimate information. When the parameters are defined as fuzzy intervals, they are also possibility distributions and automatically adhere to the axioms (definition) of possibility even if they are fuzzy entities. When fuzzy intervals are used for the parameters, what must be checked in terms of appropriate use is the semantics. The semantics associated with possibilistic uncertainty must be tied to information deficiency (lack of specificity).

The next set of paragraphs are taken from [35] both directly and in modified form. By fuzzy and possibilistic optimization, as used here, we mean optimization when at least one element of the input data is a real valued interval, a real valued random variable, a real valued fuzzy number, or a real valued number described by a possibility/necessity distribution. The use of necessity distributions are done similarly to possibility except that the necessity sematic is a pessimistic one, while the possibility semantic is optimistic.

We consider the following general programming problem:

$$\begin{aligned} z &= \min f(x,c) \\ \text{subject to } \; g_i(x,a) &\leq b \;\; i = 1,...,M_1 \\ h_j(x,d) &= e \;\; j = 1,...,M_2. \end{aligned} \tag{16}$$

The constraint set is denoted $\Omega = \{x \mid g_i(x,a) \leq b \;\; i = 1,...,M_1, h_j(x,d) = e \;\; j = 1,...,M_2\}$. It is assumed that $\Omega \neq \emptyset$. The values of a, b, c, d, and e are inputs (data, coefficients, parameters) of the programming problem. These values are subject to uncertainty for a variety of reasons. Depending on the nature of the uncertainty, they may be probability distributions, intervals, fuzzy sets or possibilistic distributions. Moreover, the operator *min* and relationships $=$ and $\leq$ can take on a flexible or fuzzy meaning becoming a soft relationship or constraint. For example, the equality and inequality relationships may be aspirations, that is, they may take on the meaning, "Come as close as possible to satisfying the relationships with some degree of violation being permissible." On the other hand, the value of $a,\ b, c, d$, or e may be described by a probability, interval, fuzzy or possibilistic distribution. In either case, the meaning of the relationships must be specified. When the objective function and/or constraints are defined by functions other than real valued convex functions, the optimization problem may not be (undoubted is not) convex so that typical solution methods are local. In very simple cases where the constraint is of the form $Ax \leq b$, and the coefficients of the matrix and right-hand side values are intervals, the solution set can be a star-shaped region (see [20]). Recall that an interval is a fuzzy number. Moreover, when the components of the matrix A are other than real valued, this means that the underlying model as specified by linear relationships is not known exactly or that the model is precise, but knowledge of what the value of the data are incomplete. We use a tilde, $\tilde{}$, to denote a fuzzy set, and a "hat", $\hat{}$, to denote a possibility distribution.

2.4.1 Fuzzy and Possibilistic Optimization Semantics

Next what is meant by decision-making in the presence of fuzzy and possibilistic entities is defined. These definitions are central to the semantics and methods. In their book (Chapter 5) Dubois and Prade [9] give clear definitions and distinctions of fuzzy measures, possibility and probabilities often forgotten and ignored by researchers (see also Chapters 1 and 7 of [13]).

1. *Fuzzy Decision Making:* Given the set of real valued (crisp) decisions, Ω, and fuzzy sets, $\{\tilde{F}_i \mid i = 1 \text{ to } n\}$, find the optimal decision in the set Ω. That is,

$$\sup_{x \in \Omega} h\left(\tilde{F}_1(x),...,\tilde{F}_n(x)\right), \tag{17}$$

where $h : [0,1]^n \to [0,1]$ is an aggregation operator [30], often taken to be the *min* operator, and $\tilde{F}_i(x) \in [0,1]$ is the fuzzy membership of x in

fuzzy set $\tilde{F}_i$. The decision space Ω is a set of real numbers (*crisp set*), and the optimal decision satisfies a mutual membership condition defined by the aggregation operator h. This is the method of Bellman and Zadeh [1], Tanaka, Okuda and Asai [48], [49], and Zimmermann [56], who were the first to develop fuzzy mathematical programming. While the aggregation operator h historically has been the *min* operator, it can be, for example, any $t-norm$ that is consistent with the context of the problem and/or decision methods (see [29]).

2. *Possibilistic Decision Making:* Given the set of real valued (crisp) decisions, Ω, and the set of possibility distributions representing the uncertain outcomes from selecting decision $\mathbf{x} = (x_1, ..., x_n)^T$ denoted $\Psi_x = \{\hat{F}_x^i, i = 1, \ldots, n\}$, find the optimal decision that produces the best set of possible outcomes with respect to an ordering U of the outcomes. That is,

$$\sup_{\Psi_x \in \Psi} U(\Psi_x), \tag{18}$$

where $U(\Psi_x)$ represents a "utility" of the set of distributions of possible outcomes $\Psi = \{\Psi_x | x \in \Omega\}$. The decision space Ψ is a **set of possibility distributions** $\Psi_x : \Omega \to [0, 1]$ resulting from taking decision $x \in \Omega$. This is the semantic taken in the possibilistic optimization of Inuiguchi [22], [23], [24] and Jamison and Lodwick [27]. If $\hat{F}_x = \hat{2}x_1 + \hat{3}x_2$, then each $\mathbf{x} = (x_1, x_2)^T$ generates the possibility distribution $\hat{F}_x = \hat{2}x_1 + \hat{3}x_2$.

Remark 6. *Let us summarize what we have just presented. For fuzzy sets* $\tilde{F}_i$, $i = 1$, ..., n, *given* x, $[\tilde{F}_1(x), ..., \tilde{F}_n(x)]^T$ *is a real valued* ***vector****. Thus, we need a way to aggregate the components of the vectors into a single real value. This is done by a t-norm, min for example. For possibility, given* x, $\Psi_x = \{\hat{F}_x^i, i = 1, \ldots, n\}$ *is a set of* ***distributions****, so we need a way to turn this set of distributions into a single real value. This is done by the utility function, a generalized expectation, for example.*

Very simply, fuzzy decision-making selects from a set of real valued, crisp, elements ordered by an aggregation operator on corresponding membership functions, while possibilistic decision making selects from a set of distributions measured by a utility operator that orders the corresponding distributions. These two different approaches have two different ordering operators (an aggregation operation for fuzzy sets such as *min* and a utility function in the case of possibility such as a generalized expectation) and lead to two different optimization methods (see [35]). The underlying sets associated with fuzzy decision-making are fuzzy, where one forms the decision space of real valued elements from operations ("*min*" and "*and*", for example, in the case of optimization of [1], [49] and [56]) on these fuzzy sets. The underlying sets associated with possibilistic decision making are real value sets, where one forms the decision space of (possibility) distributions from operations on real valued sets.

The construction of an appropriate utility function is a challenge. The axioms of utility theory as developed by Von Neumann and Morgenstern [52] are usually required. The type of utility function that is used is a challenge and decision maker dependent. For example, if one is radiating a tumor that is quite aggressive, one's utility might have higher risk (the first derivative is large and positive over the domain) than if one were radiating a tumor that was growing very slowly (the first derivative is small and positive over the domain). For this presentation, we put aside the question of how to obtain an appropriate utility function noting that it is a key to the successful implementation of the methods contained herein. The key point is that in possibilistic optimization, one is using a utility such as a generalization of the expectation to transform distributions into one real valued function (which is then optimized), whereas in fuzzy optimization, one is using an aggregation operator such as a *min* or *t-norm* to transform vectors into one real valued function (which is then optimized).

The idea of the use of utility for decision making under uncertainty problems is discussed in [17] who show how to use two qualitative counterparts to the expected utility criterion, one type of utility, U, that can be used in (18), to express uncertainty and preferences in decision making under uncertainty. Thus, what is called here *possibilistic decision making,* (18), is related to what [17] develop. However, optimization as articulated here are quantitative methods (the mapping $U : \Psi \rightarrow \mathbb{R}$), whereas the focus of [17] is more qualitative.

2.4.2 Fuzzy Decision Making Using Fuzzy Optimization

Fuzzy decision making using fuzzy optimization was first operationalized by Tanaka, Okuda, and Asai (see [48], [49]) and Zimmermann (see [56]). This approach, based on the landmark theoretical paper by Bellman and Zadeh [1], relaxes systems of inequalities $Ax \leq b$ to denote aspirations. The results are *soft constraints,* where the number b to the right of the soft inequality is a target such that, if the constraint is less than or equal to b, the membership value is one (the constraint is satisfied with certainty), and, if the constraint is greater than $b + d$, (for an *a priori* given $d > 0$), the membership is zero (the constraint is not satisfied with certainty). In between, the membership function is interpolated so that it is consistent with the definition of a fuzzy number membership function in the context of the problem. Linear interpolation was the original form (see [56]). This models a fuzzy meaning of inequality that is translated into a fuzzy membership function and is the source of our use of the designation of **flexible programming** for these classes of problems. The $\alpha - level$ represents the degree of feasibility of the constraints, consistent with the aspiration that the inequality be less than b but definitely not more than $b + d$. Thus, the objective (according to [56]) is to simultaneously satisfy all constraints at the highest possible level

of feasibility as measured by the $\alpha-levels$ of the membership functions (that is "and" all membership functions).

The approach of Tanaka, Okuda, and Asai (see [48], [49]) and Zimmermann (see [56]) deals with one way to minimize constraint violations. However, their operationalization is not always Pareto optimal [8]. Their approach must be iterated - fix the constraints at bounds and re-optimize. Their method falls within a goal satisfaction approach in optimization in which the highest degree of goal attainment is sought. They do this by minimizing the violation of the most stringent constraint. Thus, for example, this approach may guarantee that every constraint is satisfied to a 0.65 degree or more, and it may be the case that every constraint is satisfied to the 0.65 level. However, it may also be that if one of the constraints were relaxed to a 0.6 constraint violation level, all others may be satisfied at a 0.95 level. That is, this approach does not look at the aggregate satisfaction, only the most constraining one. It is minimizing the maximum constraint violation.

An aggregate goal attainment tries to maximize an overall *measure* of aggregate goal satisfaction. The *aggregate* sum of goal attainment focuses on maximizing the cumulative satisfaction of the goals. The surprise function (see [38], [40]) is one such measure for an aggregate set of (fuzzy) goals. In particular, when the right-hand side values are interpreted as goals rather than rigid constraints, the problem may be translated into one of optimizing the aggregate goal satisfaction. Thus, for soft constraints derived as,

$$\text{hard} \quad y_i = (A\mathbf{x})_i \leq b_i \Rightarrow \text{soft} \quad y_i = (A\mathbf{x})_i \leq \tilde{b}_i, \tag{19}$$

where the right-hand side values of the soft constraint are *fuzzy numbers,* the transformation into a set of aggregate goal satisfaction problem using the surprise function as the measure for the cumulative goal satisfaction is attained as follows. A (soft) fuzzy inequality (19) is translated into a fuzzy membership function, $\mu_i(x)$, which is the possibility $pos(\tilde{b}_i \geq x)$. Each membership function is translated into a surprise by

$$s_i(x) = \left(\frac{1}{\mu_i(x)} - 1\right)^2. \tag{20}$$

These functions are added to obtain a total surprise

$$S(\mathbf{x}) = \sum_i s_i(y_i) = \sum_i s_i(y_i) = \sum_i s_i((A\mathbf{x})_i). \tag{21}$$

Note that (21) is an aggregation operator. A best compromise solution based on the surprise function is given by the nonlinear optimization problem

$$\min \ z = \sum_i s_i((A\mathbf{x})_i)$$
$$\text{subject to } x \in \Omega \text{ (possible hard constraints).}$$

That is, a real valued inequality constraint whose right-hand side value is a fuzzy number is translated into a fuzzy set. This fuzzy set then is transformed into a real valued set, one for each $\alpha - level$, via a surprise function whose domain, of course, is the $\alpha - level$ of the fuzzy sets. The objective is to minimize the **sum** of all surprise function values. Unlike Tanaka and Zimmermann, the constraints are not restricted such that *all* satisfy a minimal level. The surprise function approach effectively sums each of the $\alpha - levels$ for each of the constraints, then maximizes this sum with respect to α. Since the optimization is over sets of crisp values coming from fuzzy sets, the surprise approach is a fuzzy optimization method. The salient feature is that surprise uses a dynamic penalty for falling outside distribution/membership values of one. The advantage is that the individual penalties are convex functions, which become infinite as the values approach the endpoints of the support. Moreover, this approach is computationally tractable.

Again, the surprise approach may be used to handle soft constraints of Tanaka, Okuda, and Asai (see [49]) and Zimmermann (see [56]), since these soft constraints can be considered to be fuzzy numbers. However, if soft constraints are handled using surprise functions, the sum of the failure to meet the constraints is minimized rather than forcing each constraint to meet a minimal (fuzzy) feasibility level.

Another historically significant interpretation of fuzzy optimization comes from Verdegay [51], who proposes a method for obtaining a fuzzy solution to a fuzzy problem. This is a deviation from the solutions examined so far, which have been real valued, crisp solutions. Verdegay considers a problem with fuzzy constraints,

$$\begin{aligned} z &= f(x) \\ x &\subseteq \tilde{C}, \end{aligned} \tag{22}$$

where the set of constraints have a membership function μ_C, with $\alpha - cut$ $\tilde{C}_\alpha$.

Verdegay defines x_α as the set of solutions that satisfy constraints $\tilde{C}_\alpha$. Then a fuzzy solution to the fuzzy mathematical programming problem is

$$\begin{aligned} &\max_{x \in \tilde{C}_\alpha} z = f(x) \\ &\quad \forall\, \alpha \in [0,1]. \end{aligned} \tag{23}$$

Verdegay proposes solving (23) parametrically for $\alpha \in [0,1]$ to obtain a fuzzy solution $\tilde{\chi}$, with $\alpha - cut\ \chi_\alpha$, which yields fuzzy objective value $\tilde{z}$, with α-cut z_α.

2.4.3 Possibilistic Decision Making Using Possibilistic Optimization

One approach to possibility distributions of parameters ([27] and [32]) allows all constraint violations at an established cost or penalty and minimizes the

expected average, a generalization of expected value ([10], [26], [27], [53], and [54]). This approach considers all possible outcomes as a weighted expected average penalty. The expected average is a type of utility. Another utility is minimizing regret, from Kasperski's article in this volume. This particular utility takes violations as penalties on all outcomes of the constraints. It optimizes over sets of possibility distributions, so it is possibilistic optimization.

Another approach that optimizes over possibility distributions [23] and [24]also optimize over distributions considers constraint feasibility as possibilistic generalizations of *chance constraint* methods. The approach used in [27] and [32] is a possibilistic generalization of the *recourse models* in stochastic optimization (see for example [2]), where violations of constraints are considered as allowable up to a maximum but at a cost. The recourse model in the context of non-probabilistic uncertainty has been studied by [25] where interval parameters/coefficients are treated.

Possibilistic optimization historically was introduced by Buckley [3] as an off-shoot of fuzzy optimization. Because this is a possibilistic linear program, the objective function is governed by a possibilistic distribution. To derive the possibilistic objective function value for a particular solution x, Buckley first specifies the possibility that x satisfies each constraint, and takes the minimum to indicate the possibility that x satisfies all constraints. Buckley next constructs $\text{Poss}[Z = z|x]$, which is the conditional possibility that the objective function Z obtains a particular value z, given values for x. This definition of the possibility distribution motivates Buckley's solution method. Recall that because we are dealing with a possibilistic problem, the solution is governed by a possibilistic distribution. Buckley's method depends upon a static α, chosen *a priori*. The decision maker defines an acceptable level of uncertainty in the objective outcome, $0 < \alpha \leq 1$. For a given α, we define the left and right end-points of the α-cut of a fuzzy number $\tilde{x}$ as $x^-(\alpha)$ and $x^+(\alpha)$, respectively. Using these, Buckley defines a new objective function:

$$\begin{aligned} Z(\alpha) &= c^-(\alpha)x \\ A^+(\alpha)x &\geq b^-(\alpha). \end{aligned} \tag{24}$$

This linear program is constrained upon the best-case scenario. That is, for a given α-level, each variable is multiplied by the largest possible coefficient $a_{ij}^+(\alpha)$, and is required to be greater than the smallest possible right-hand side $b_i^-(\alpha)$. We should interpret $Z(\alpha)$ accordingly. If the solution to the linear program is implemented, the possibility that the objective function will attain the level $Z(\alpha)$ is given by α. Stated another way, the best-case scenario is that the objective function attains a value of $Z(\alpha)$, and the possibility of the best case scenario occurring is α.

In the mid 1980s, Tanaka and Asai [47] and Tanaka, Ichahashi, and Asai [50] proposed a technique for dealing with ambiguous coefficients and right-hand sides based upon a possibilistic definition of "greater than zero." First, the objective function is viewed as a goal. As in flexible programming, the goal

becomes a constraint, with the aspiration level for the objective function on the right-hand-side of the inequality. The right-hand sides of the constraints are subtracted so that all the numeric information is contained in a single matrix, which is feasible when it is greater than zero. A fuzzy measure of non-negativity is introduced to gauge optimality of a potential solution x.

Luhanjula's [39] formulation of the possibilistic mathematical program depends upon his concept of "more possible" values. He first defines a possibility distribution Π_X with respect to constraint F as

$$\Pi_X = \mu_F(u),$$

where $\mu_F(u)$ is the degree to which the constraint F is satisfied when u is the value assigned to the solution X. Then the set of more possible values for X, denoted by $V_p(X)$, is given by

$$V_p(X) = \Pi_X^{-1}(\max_u \Pi_X(u)).$$

In other words, $V_p(X)$ contains elements of U which are most compatible with the restrictions defined by Π_X. It follows from intuition and from Luhanjula's formal proof [39] that when Π_X is convex, $V_p(X)$ is a real valued interval, and when Π_X is strongly convex, $V_p(X)$ is a single real number. This formulation varies significantly from the other approaches considered thus far. The possibility of each possibilistic component is maximized individually. Other formulations have required that each possibilistic component $\tilde{c_j}$, $\tilde{A_{ij}}$, and $\tilde{b_i}$ achieve the same possibility level defined by α. This formulation also has a distinct disadvantage over the others presented here. The authors know of no proposed computational method for determining the "more possible" values, V_p, so there appears to be no way to solve the deterministic the problem.

2.4.4 Mixed Fuzzy and Possibilistic Decision Making Using Mixed Possibilistic and Fuzzy Optimization Methods

An optimization problem containing both fuzzy and possibilistic variables is called a *mixed problem* in this chapter. Problems in which one or more possibilistic parameters occurs with one or more fuzzy parameters (or fuzzy inequalities) have been studied (see [24], [36]). Within a quantitative setting, there are two cases for the mixed problem. The first case is a problem that contains both fuzzy and possibilistic parameters (or soft inequalities), but in which each constraint contains exclusively fuzzy or possibilistic parameters. In this case, the fuzzy constraints can be optimized by $\alpha - levels$ (according to [40] or [56]) and the possibilistic constraints by penalized expected averages (according to [27]). For the more complex case in which both fuzzy and possibilistic parameters appear in the same constraint, one approach is to compute the possibilistic distributions of the aggregation of the fuzzy membership functions and optimize over the penalized expected average. The

fuzzy parameter(s) mixed with possibilistic parameter(s) generate a possibilistic distribution that depends on the aggregation operator of the fuzzy membership functions. Another approach is described in Untiedt's chapter in this volume.

3 A Taxonomy of Fuzzy and Possibilistic Optimization for Our Generic Problem

The structure of the generic optimization problem (16) when it is a linear programming problem may be considered to be formed by (i) the *rim* $f(x,c) = c^T x$, and b, e, (ii) the *body* $f(x,a), h(x,d) = Ax$, and (iii) the *relationship*, $\leq, =$ ([21]). For the generic form of the mathematical programming problem (16), we consider a taxonomy based on (i) rim objective function parameters, c, (ii) rim right-hand side parameters, b and e, (iii) body parameters a and d, and (iv) relationship $\leq, =$. For this exposition, a fuzzy/possibilistic optimization problem is considered to be (16) in the presence of data $\{a, b, c, d, e, \leq, =\}$ that is either all or a mixture of real, interval, probabilistic, fuzzy, possibilistic with at least one parameter being fuzzy or possibilistic where soft constraints are assumed to have been translated into fuzzy intervals. If we have a probabilistic optimization problem whose values are known over each $x \in \mathbb{R}$, we would consider it under possibilistic optimization where the upper bound and lower bound (possibility/plausibility and necessity/belief) would be equal. If the probability were known only over sets, then we would have an upper possibility bound and a lower necessity bound as in Figure 3 and do our bound interval-valued possibilistic optimization which is transformed into utility optimization. A right-hand side value that is fuzzy may be interpreted in two ways depending on the context of the problem. First, a fuzzy right-hand side may indicate flexibility. Second, it may indicate (true) decomposable transition modelled by a fuzzy interval. For the former, the constraint becomes a flexible constraint. For the latter, it becomes a possibility.

Note that in the context of interval, fuzzy, possibility, $r \leq s$ and $s \leq r$ does **not** imply $r = s$ as can be seen in the following.

Example 7. *Let*

$$[2,3]x \leq [3,6] \text{ and} \tag{25}$$

$$[2,3]x \geq [3,6]. \tag{26}$$

The solution of (25) is

$$x = (-\infty, 1].$$

The solution of (26) is

$$x = [3, \infty).$$

Thus (25) and (26) imply that $x = \varnothing$. *However,*

$$[2,3][\underline{x}, \overline{x}] = [3,6]$$

means that $[\underline{x}, \overline{x}] = [\frac{3}{2}, 2]$.

This example shows that for the constraint fuzzy/possibilistic linear system, $Ax \leq b$ and $Ax \geq b$ are not equivalent to $Ax = b$.

The types of optimization in the presence of interval, fuzzy interval, possibilistic coefficients, and soft constraints are as follows:

1. **Flexible Programming**
   a. Soft constraints relationships $\leq$ and/or $=$ that take on a flexible meaning (come as close as possible or come as close as possible without violation some hard constraints).
   b. The objective function expresses a target desired (come as close as possible to staying under a budgetary value, or attain at least as much profit as was obtained last year).
   c. The right-hand-side value of a constraint is a fuzzy interval which is semantically a target (deliver as close to zero radiation as possible to healthy cells but absolutely to not exceed a critical threshold value).
2. **Utility Programming**
   a. Interval, fuzzy interval, possibilistic cost coefficients of the objective function rim parameter c with real valued coefficient constraint coefficient $a, b, d, e \in \mathbb{R}$.
   b. The objective function rim parameter $c \in \mathbb{R}$ with interval, probability, possibility, fuzzy interval, and one or two of the following - body parameters interval, fuzzy interval, possibilistic a, d and/or rim right-hand values b, e are possibilistic.
   c. Interval-Valued Probability Measure (IVPM) Programming - any of the coefficients a, b, c, d, e may be interval, fuzzy, possibilistic where there may be a mixture of types within one constraint statement.
3. **Random Set** - any of the coefficients a, b, c, d, e may be interval, fuzzy, possibilistic where there may be a mixture of types within one constraint statement.

One also might classify fuzzy and possibilistic programming according to whether or not the solution is a real valued fuzzy interval vector or a real valued vector. Possibilistic programming methods of Buckley [3] and his colleagues and Delgado [5] and his colleagues have considered fuzzy interval-valued solutions. The methods to obtain fuzzy interval solutions are different than those that obtain a real valued solution. Nevertheless, they fall under possibilistic programming or random set programming of the above taxonomy.

References

1. Bellman, R.E., Zadeh, L.A.: Decision-making in a fuzzy environment. Management Science, Serial B 17, 141–164 (1970)
2. Birge, J.R., Louveux, F.: Introduction to Stochastic Programming. Springer, New York (1997)
3. Buckley, J.J.: Solving possibilistic linear programming problems. Fuzzy Sets and Systems 31, 329–341 (1989)
4. de Cooman, G.: Possibility theory I: The measure- and integral - theoretic groundwork. International Journal of General Systems 25(4), 291–323 (1997)
5. Delgado, M., Verdegay, J.L., Vila, M.A.: A general model for fuzzy linear programming. Fuzzy Sets and Systems 29, 21–29 (1989)
6. Dempster, A.P.: Upper and lower probabilities induced by a multivalued mapping. Annals of Mathematical Statistics 38, 325–339 (1967)
7. Dubois, D.: Linear programming with fuzzy data. In: Bezdek, J. (ed.) Analysis of Fuzzy Information. Applications in Engineering and Science, vol. III, pp. 241–263. CRC Press, Boca Raton (1897)
8. Dubois, D.: Personal communications (2008, 2009)
9. Dubois, D., Prade, H.: Fuzzy Sets and Systems: Theory and Applications. Academic Press, New York (1980)
10. Dubois, D., Prade, H.: The mean value of a fuzzy number. Fuzzy Sets and Systems 24, 279–300 (1987)
11. Dubois, D., Prade, H.: Possibility Theory. Plenum Press, New York (1988)
12. Dubois, D., Prade, H.: The three semantics of fuzzy sets. Fuzzy Sets and Systems 90, 141–150 (1997)
13. Dubois, D., Prade, H. (eds.): Fundamentals of Fuzzy Sets. Kluwer Academic Press, Dordrecht (2000)
14. Dubois, D., Prade, H.: Formal representations of uncertainty. In: IFSA 2009, Lisboa, Portugal (July 2009)
15. Dubois, D., Karre, E., Mesiar, R., Prade, H.: Fuzzy interval analysis. In: Dubois, D., Prade, H. (eds.) Fundamentals of Fuzzy Sets, ch. 10, pp. 483–581. Kluwer Academic Press, Dordrecht (2000)
16. Dubois, D., Nguyen, H., Prade, H.: Possibility theory, probability theory, and fuzzy sets: Misunderstanding, bridges and gaps. In: Dubois, D., Prade, H. (eds.) Fundamentals of Fuzzy Sets, ch. 7, pp. 343–438. Kluwer Academic Press, Dordrecht (2000)
17. Dubois, D., Prade, H., Sabbadin, R.: Decision-theoretic foundations of qualitative possibility theory. European Journal of Operational Research 128, 459–478 (2001)
18. Fandel, G.: PPS-Systeme: Grundlagen, Methoden, Software, Markanalyse. Springer, Heidelberg (1994)
19. Fishburn, P.C.: Stochastic dominance and moments of distributions. Mathematics of Operations Research 5, 94–100 (1980)
20. Gay, D.M.: Solving interval linear equations. SIAM Journal of Numerical Analysis 19(4), 858–870 (1982)
21. Greenberg, H.: Personal communication (2005)
22. Inuiguchi, M.: Stochastic programming problems versus fuzzy mathematical programming problems. Japanese Journal of Fuzzy Theory and Systems 4(1), 97–109 (1992)

23. Inuiguchi, M., Ichihashi, H., Kume, Y.: Relationships between modality constrained programming problems and various fuzzy mathematical programming problems. Fuzzy Sets and Systems 49, 243–259 (1992)
24. Inuiguchi, M., Sakawa, M., Kume, Y.: The usefulness of possibilistic programming in production planning problems. Int. J. of Production Economics 33, 49–52 (1994)
25. Inuiguchi, M., Tanino, T.: Two-stage linear recourse problems under non-probabilistic uncertainty. In: Yoshida, Y. (ed.) Dynamical Aspect in Fuzzy Decision Making, pp. 117–140. Physica-Verlag, Heidelberg (2001)
26. Jamison, K.D.: Modeling uncertainty using probability based possibility theory with applications to optimization. Ph.D. Thesis, UCD Department of Mathematics (1998)
27. Jamison, K.D., Lodwick, W.A.: Fuzzy linear programming using penalty method. Fuzzy Sets and Systems 119, 97–110 (2001)
28. Jamison, K.D., Lodwick, W.A.: The construction of consistent possibility and necessity measures. Fuzzy Sets and Systems 132, 1–10 (2002)
29. Kaymak, U., Sousa, J.M.: Weighting of constraints in fuzzy optimization. Constraints 8, 61–78 (2003)
30. Klir, G., Yuan, B.: Fuzzy Sets and Fuzzy Logic: Theory and Applications. Prentice Hall, New Jersey (1995)
31. Lodwick, W., Bachman, K.: Solving Large Scale Fuzzy Possibilistic Optimization Problems. Fuzzy Optimization and Decision Making 4(4), 257–278 (2005)
32. Lodwick, W.A., Jamison, K.D.: A computational method for fuzzy optimization. In: Ayyub, B., Gupta, M. (eds.) Uncertainty Analysis in Engineering and Sciences: Fuzzy Logic, Statistics, and Neural Network Approach, ch. 19. Kluwer Academic Publishers, Dordrecht (1997)
33. Lodwick, W.A., Jamison, K.D.: Estimating and validating the cumulative distribution of a function of random variables: Toward the development of distribution arithmetic. Reliable Computing 9, 127–141 (2003)
34. Lodwick, W.A., Jamison, K.D.: Theory and semantics for fuzzy and possibilistic optimization. In: Fuzzy Logic, Soft Computing and Computational Intelligence (Eleventh International Fuzzy Systems Association World Congress), Beijing, China, July 28-31, vol. III, pp. 1805–1810 (2005)
35. Lodwick, W.A., Jamison, K.D.: Theory and semantics for fuzzy and possibilistic optimization. Fuzzy Sets and Systems 158(7), 1861–1871 (2007)
36. Lodwick, W.A., Jamison, K.D.: The Use of interval-valued probability measure in optimization under uncertainty for problems containing a mixture of possibilistic, probabilistic and interval uncertainty. In: Melin, P., Castillo, O., Aguilar, L.T., Kacprzyk, J., Pedrycz, W. (eds.) IFSA 2007. LNCS (LNAI), vol. 4529, pp. 361–370. Springer, Heidelberg (2007)
37. Lodwick, W.A., Jamison, K.D.: Interval-Valued Probability in the Analysis of Problems Containing a Mixture of Possibilistic, Probabilistic, and Interval Uncertainty. Fuzzy Sets and Systems 159(1), 2845–2858 (2008)
38. Lodwick, W.A., Neumaier, A., Newman, F.D.: Optimization under uncertainty: methods and applications in radiation therapy. In: Proceedings 10th IEEE International Conference on Fuzzy Systems, vol. 3, pp. 1219–1222 (2001)
39. Luhandjula, M.K.: On possibilistic linear programming. Fuzzy Sets and Systems 18, 15–30 (1986)
40. Neumaier, A.: Fuzzy modeling in terms of surprise. Fuzzy Sets and Systems 135(1), 21–38 (2003)

41. Puri, M.L., Ralescu, D.: Fuzzy measures are not possibility measures. Fuzzy Sets and Systems 7, 311–313 (1982)
42. Rommelfanger, H.J.: The advantages of fuzzy optimization models in practical use. Fuzzy Optimization and Decision Making 3, 295–309 (2004)
43. Rommelfanger, H.J.: Personal communication (April 2009)
44. Russell, B.: Vagueness. Australasian Journal of Psychology and Philosophy 1, 84–92 (1923)
45. Shafer, G.: Mathematical Theory of Evidence. Princeton University Press, Princeton (1976)
46. Simon, H.: The Sciences of the Artificial, 2nd edn. The MIT Press, Cambridge (1969/1981)
47. Tanaka, H., Asai, K.: Fuzzy linear programming with fuzzy numbers. Fuzzy Sets and Systems 13, 1–10 (1984)
48. Tanaka, H., Okuda, T., Asai, K.: On fuzzy mathematical programming. Transactions of the Society of Instrument and Control Engineers 9(5), 607–613 (1973) (in Japanese)
49. Tanaka, H., Okuda, T., Asai, K.: On fuzzy mathematical programming. J. of Cybernetics 3, 37–46 (1974)
50. Tanaka, H., Ichihashi, H., Asai, K.: Fuzzy Decision in linear programming with trapezoid fuzzy parameters. In: Kacpryzk, J., Yager, R.R. (eds.) Management Decision Support Systems Using Fuzzy Sets and Possibility Theory, pp. 146–154. Verlag TUV, Koln (1985)
51. Verdegay, J.L.: Fuzzy mathematical programming. In: Gupta, M.M., Sanchez, E. (eds.) Fuzzy Information and Decision Processes, pp. 231–237. North Holland Company, Amsterdam (1982)
52. Von Neumann, J., Morgenstern, O.: Theory of Games and Economic Behavior. Princeton University Press, Princeton (1947)
53. Yager, R.: On choosing between fuzzy subsets. Kybernetes 9, 151–154 (1980)
54. Yager, R.: A procedure for ordering fuzzy subsets of the unit interval. Information Sciences 24, 143–161 (1981)
55. Zadeh, L.A.: Fuzzy sets as a basis for a theory of possibility. Fuzzy Sets and Systems 1, 3–28 (1978)
56. Zimmermann, H.: Description and optimization of fuzzy systems. International J. of General Systems 2, 209–215 (1976)

Part II
Basic Issues

Aggregation Operators for Evaluating Alternatives

Vicenç Torra

Abstract. This chapter reviews the use of aggregation functions and operators in the field of decision making. We first present an overview of main decision making problems, and, then, we show that aggregation operators are in common use for solving them. Once having presented their interest, we describe the major aggregation operators, their properties and their major differences.

1 Introduction

Decision making [8] is a broad field that encompasses a few slightly different problems. In general, all of them are concerned about the process of selecting an alternative among a set of them. Differences on the problems correspond to differences on the settings around the alternatives.

In this paper we will discuss the use of aggregation operators [18] in the field of decision making. Although these operators have been used for different purposes in this framework, we will focus on their application in multicriteria decision making. In particular, we will discuss their application to the case of a finite set of alternatives, each of them evaluated using numerical utility functions. In this setting, aggregation operators can be used to aggregate the different values of the utility functions, so that an aggregated value (an aggregated criteria) is constructed.

The structure of the paper is as follows. In Section 2 we will describe with some detail decision making problems and explain where aggregation operators can be applied in this setting. Then, in Section 3 we will give an overview of some of the aggregation operators. The paper finishes with a summary.

Vicenç Torra
IIIA-CSIC, Campus UAB s/n, 08193 Bellaterra, Catalonia, Spain
e-mail: vtorra@iiia.csic.es

W.A. Lodwick & J. Kacprzyk (Eds.): Fuzzy Optimization, STUDFUZZ 254, pp. 65–76.
springerlink.com

2 Decision Making and Aggregation Operators

Basically, all decision making problems concern a set of alternatives $A = \{a_1, a_2, \dots\}$ (either finite or infinite), a set of criteria $C = \{c_1, \dots, c_n\}$ that permit us to evaluate in which degree the alternatives satisfy the requirements for their selection, and a group of people $G = \{e_1, \dots, e_g\}$ (e for expert).

Different approaches in decision making correspond to different interactions among these previous elements: alternatives, criteria, and experts. These different approaches have originated a set of different names. A few of them are described below.

Two common terms are *multicriteria decision making* (MCDM) and *multicriteria decision aid* (MCDA). MCDA usually focus on the tools that help a user to capture, understand, and analyze the differences between the alternatives. So, this follows a constructivist approach. In contrast, MCDM pressumes that the decision making process can be formalized, and focuses on tools to describe this process. So, MCDM follows a descriptive approach. The French and the American school are focused, respectively, on the MCDA and MCDM approaches. The French school is after ELECTRE and Roy and Vanderpooten [3, 9, 10]. For a discussion of the two schools see e.g. [5]. See also [1, 20].

Within MCDA and MCDM two main areas can be distinguished. One of them is *Multiobjective decision making* (MODM) [12], which corresponds to the case that there is an infinite number of alternatives. That is, the space of alternatives is a continuum. In contrast, *Multiattribute decision making* (MADM) corresponds to the case of a finite number of alternatives. It is usual to use MCDM as synonymous of MADM. MODM problems are usually expressed in terms of a function to be maximized with some constraints to be satisfied, and these problems are usually solved using optimization tools (as e.g., the SIMPLEX algorithm).

In both MCDM and MADM, the decision making process follows the formalization given above with a predefined set of alternatives $A = \{a_1, a_2, \dots\}$, which is usually finite, and a set of criteria $C = \{c_1, \dots, c_n\}$. Each criterion is assumed to define a preference over the alternatives. Different approaches have been considered in the literature for expressing the preferences for each criteria. The most used ones are the following:

- **Utility functions.** They are functions over alternatives into a certain range. Then, the larger the value attached to an alternative, the more we prefer it.
- **Preference relations.** They are binary relations that are satisfied for a pair of alternatives (e.g., (al_1, al_2)) when the first one is preferred to the second one.

We illustrate these two alternatives with an example.

Example 1. Let us consider the problem of buying a car with the following set of alternatives $A = \{Peugeot206, FordT., \dots\}$ and criteria $c = \{price, quality, comfort\}$. Then, we can consider representing our preference on the alternatives by means of utility functions (e.g., functions $U_{price}, U_{quality}, U_{comfort}$) or by means of preference

relations (e.g., $\mathbf{R}_{price}, \mathbf{R}_{quality}, \mathbf{R}_{comfort}$). For illustration, we use the following definitions:

- **Utility functions**
  - **Ford T:** $U_{price} = 0.2$, $U_{quality} = 0.8$, $U_{comfort} = 0.3$
  - **Peugeot206:** $U_{price} = 0.7$, $U_{quality} = 0.7$, $U_{comfort} = 0.8$
- **Preference relations**
  - **Price:** $R_{price}(P206, FordT)$, $\neg R_{price}(FordT, P206)$
  - **Quality:** $\neg R_{quality}(P206, FordT)$, $R_{quality}(FordT, P206)$
  - **Comfort:** $R_{comfort}(P206, FordT)$, $\neg R_{comfort}(FordT, P206)$

Here, we understand $R_c(a,b)$ as that we prefer a to b with respect to c.

Decision making problems as the one described in Example 1 can be solved constructing an aggregate preference that synthesizes the ones of the different criteria.

In the case of considering a set of criteria $C = \{c_1, \ldots, c_n\}$ represented in terms of preference relations $R = \{R_1, \ldots, R_n\}$ this problem consists of constructing a new relation $R_{\mathbb{C}}$ as a combination of the ones in $R = \{R_1, \ldots, R_n\}$. Note that once this relation $R_{\mathbb{C}}$ is built, the selection of the *best* alternative corresponds to selecting the alternative (or the set of alternatives) that is preferable with respect to $R_{\mathbb{C}}$.

Similarly, if the set of criteria is represented in terms of the utility functions $U = \{U_1, \ldots, U_n\}$, we have that one way to solve the problem is to build a utility function $U_{\mathbb{C}}$ as a combination of the ones in the set U. As before, this aggregated utility can be used to select the *best* alternative.

It has to be said that alternative approaches exist for each of the possible representations for the criteria. E.g., in the case of the utility functions, we can use them to define preference relations (i.e., for each U_i we define $R_i(a,b) := (U_i(a) > U_i(b))$) and, then, aggregate the preference relations.

The functions to combine the utility functions and the functions to combine the preference relations are aggregation functions (or aggregation operators). The former typically are aggregation operators for numerical data [19] and the latter aggregation operators for preferences (see e.g. [4]). Other aggregation operators are of interest here, as utilities can be expressed in domains other than numerical. E.g., we can express them using ordered sets of categories (ordered sets of linguistic labels, or ordinal scales), partial orders, or fuzzy sets. In Section 3 of this paper, we will describe some of the existing functions to aggregate these preferences.

A topic not discussed so far is when a group of people is involved in the decision problem. *Group decision making* (GDM) and social choice correspond to this type of problems. So, in this case, there is more than one e in $G = \{e_1, \ldots, e_g\}$.

3 Aggregation Operators

At present, a large number of aggregation functions have been defined. They are functions that combine N different data into a single datum. Let us express these

functions by $\mathbb{C}$ (from $\mathbb{C}$onsensus), then, the aggregation of $a_1,\ldots,a_N$ in a given domain D is $\mathbb{C}(a_1,\ldots,a_N)$, also in this domain D.

Differences on the functions are due to the type of data they aggregate (different domains D) and the properties these functions satisfy. In general, an aggregation operator is a function that satisfies monotonicity and unanimity (at least for the values at the boundaries). That is, they are functions satisfying:

- **Unanimity or idempotency:** $\mathbb{C}(a,\ldots,a) = a$ for all a in D
- **Monotonicity:** $\mathbb{C}(a_1,\ldots,a_N) \geq \mathbb{C}(a'_1,\ldots,a'_N)$ when $a_i \geq a'_i$

Some variations exist in the literature on what an aggregation operator is. First, naturally, the monotonicity condition can only be satisfied in those domains D where $\geq$ is defined. In addition, some call aggregation operators those functions satisfying monotonicity and idempotency in the boundary of D. E.g., if $D = [0,1]$, idempotency is only required for 0 and 1. That is, it is required $\mathbb{C}(0,\ldots,0) = 0$ and $\mathbb{C}(1,\ldots,1) = 1$. Under this definition, some functions for numerical data as t-norms and t-conorms are aggregation functions. In this case, mean operators correspond to the ones that satisfy unanimity for all a in D.

In addition to the two conditions above, it is not uncommon to require symmetry with respect to the inputs. That is,

- **Symmetry:** For any permutation π on $\{1,\ldots,N\}$ it holds that

$$\mathbb{C}(a_1,\ldots,a_N) = \mathbb{C}(a_{\pi(1)},\ldots,a_{\pi(N)})$$

3.1 Aggregation Operators for Numerical Data

The two simplest and most well known aggregation operators are the arithmetic mean and the weighted mean. While the former combines the data without requiring any parameter, the latter uses a weighting vector **p**. This weighting vector permits us to take into account in the aggregation process some *a prior* information about the information sources. Formally, these operators, as well as the weighting vector, are defined as follows:

Definition 1. *Let $A = (a_1,\ldots,a_N)$ be N data in $\mathbb{R}$. Then, we define a weighting vector, the arithmetic mean ($AM : \mathbb{R}^N \to \mathbb{R}$) of A, and the weighted mean (WM) of A with respect to a weighting vector as follows:*

- *A vector $v = (v_1 \ldots v_N)$ is a* weighting vector *of dimension N if and only if $v_i \in [0,1]$ and $\sum_i v_i = 1$.*
- *AM is an* arithmetic mean*, if $AM(a_1,...,a_N) = (1/N)\sum_{i=1}^{N} a_i$.*
- *WM is the* weighted mean *with respect to a weighting vector* **p***, if $WM_{\mathbf{p}}(a_1,...,a_N) = \sum_{i=1}^{N} p_i a_i$.*

The OWA (Ordered Weighting Averaging) operator is a similar function that combines N values with respect to a weighting vector. Its definition is as follows:

Definition 2. *[21] (see also [22]) Let* $\mathbf{w}$ *be a weighting vector of dimension N; then, a mapping OWA:* $\mathbb{R}^N \to \mathbb{R}$ *is an* Ordered Weighting Averaging (OWA) operator *of dimension N if*

$$OWA_{\mathbf{w}}(a_1,...,a_N) = \sum_{i=1}^{N} w_i a_{\sigma(i)},$$

where $\{\sigma(1),...,\sigma(N)\}$ *is a permutation of* $\{1,...,N\}$ *such that* $a_{\sigma(i-1)} \geq a_{\sigma(i)}$ *for all* $i = \{2,...,N\}$ *(i.e.,* $a_{\sigma(i)}$ *is the ith largest element in the collection* $a_1,...,a_N$*).*

Although the definition of the weighted mean and the one of the OWA are similar, both are a linear combination of the values with respect to the weights, there is a fundamental difference due to the permutation σ included in the definition of the OWA. While the *i*th weight is assigned to the *i*th data in the weighted mean, this is not so in the OWA. In this latter operator, the weight is assigned to the position. Due to this, in a multicriteria decision making application weights in the weighted mean are assigned to the criteria, while in the OWA they are assigned to the value itself, or to the relative position of one value with respect to the other values. This permits the decision maker to represent different types of *prior* knowledge or different types of information in the aggregation process.

For example, when in a decision process we have the set of criteria $C = \{c_1, c_2, \dots\}$ and we know that the criteria c_1 is more important than the criteria c_2, then, we will assign a weight p_1 larger than the one of p_2, and latter apply the weigthed mean as the proper aggregation operator. The larger the importance of c_1, the larger its weight with respect to the others. In this case, the OWA operator is not appropriate.

In contrast, when we have a decision making process where we need to represent some compromise between the different criteria, the OWA operator can be used. For example, let us consider the case of having four criteria c_1, c_2, c_3, and c_4, and let us require for an alternative being acceptable that three of the criteria are good and, in this case, the fourth does not matter. This situation can be modeled using the OWA operator with e.g. a weighting vector $(0,0,1,0)$. Note that in this case, the evaluation of an alternative will be equal to the evaluation of the third worst criteria. That is, the worst criteria will be ignored (no matter its value), and the evaluation of the alternative will be equal to the criteria with the third worst value. If this value is very good, it naturally means that the other two are also very good, and, thus, the alternative will have a large evaluation.

As OWA operates in this way, it is said that this operator permits us to represent certain compensation between the criteria. No compensation corresponds to give a large importance to low values (i.e., a very low value can *damage* the good values of the other criteria). Instead, maximum compensation is achived giving a large importance to high values (i.e., any high value can cause the ignorance of all the other, possibly bad, values).

Due to the significance of the compensation degree in the OWA, a measure has been defined to evaluate in what extent a certain weighting vector allows for compensation. This measure is known as *orness*. The larger the orness, the larger the compensation. The lower the orness, the lower the compensation. It is formally

defined below. Its definition is valid for all aggregation operators $\mathbb{C}$ and for all parameterizations P. For some of the operators, the orness does not depend on the particular parameterization selected, while for others the orness depends on the particular parameterization used. The weighted mean is an example of the former (i.e., the orness of the weighted mean is independent of the parameter used), and the OWA is an example of the latter (i.e., different parameters give different orness for the OWA).

Definition 3. *Let $\mathbb{C}$ be an aggregation operator with parameters P; then, the* orness *of $\mathbb{C}_P$ is defined by*

$$\text{orness}(\mathbb{C}_P) := \frac{AV(\mathbb{C}_P) - AV(\min)}{AV(\max) - AV(\min)}. \tag{1}$$

The orness of the aggregation operators reviewed above is as follows:

- $orness\ (AM) = 1/2$
- $orness(WM_{\mathbf{p}}) = 1/2$
- $orness(OWA_{\mathbf{w}}) = \frac{1}{N-1}\sum_{i=1}^{N}(N-i)w_i$

As stated above, we have that the orness of the weighted mean does not depend on the weighting vector **p** while in the case of the OWA the orness depends on the weighting vector. Note that in the case of the OWA, the maximum orness is 1 when $w_1 = 1$ and $w_i = 0$ for all $i \neq 1$ (in this case the OWA corresponds to the maximum), and the minimum ornes is 0 when $w_N = 1$ and $w_i = 0$ for all $i \neq N$ (in this case the WM corresponds to the minimum).

3.1.1 The WOWA Operator

In the previous section we have seen that both weighted mean and OWA have a similar form as both are a linear combination of the values $a_1, \ldots, a_N$ with respect to weights. Nevertheless, we have discussed that due to the ordering step σ, the meaning of the weights are different and that they permit to represent different aspects. That is, the weighting vector **p** in the weighted mean permits us to represent the importance of the criteria, while the weighting vector **w** in the OWA permits to represent the compensation degree (which can be quantified by means of the orness measure). Note that although we use here different letters **w** and **p**, both weighting vectors have the same mathematical properties. For the sake of clearity, we will use in the rest of this section **w** and **p** to represent, respectively, the weights from the OWA and weighted mean, respectively.

In certain multicriteria decision making problems it is relevant to consider at the same time weights for the different criteria, and a certain degree of compensation. To this end, we can use the WOWA (Weighted OWA) operator. This operator permits us to include in a single application both types of weights. Its definition is given below.

Definition 4. *[13] Let* **p** *and* **w** *be two weighting vectors of dimension N; then, a mapping WOWA:* $\mathbb{R}^N \to \mathbb{R}$ *is a* Weighted Ordered Weighted Averaging (WOWA) operator *of dimension N if*

$$WOWA_{\mathbf{p},\mathbf{w}}(a_1,...,a_N) = \sum_{i=1}^{N} \omega_i a_{\sigma(i)},$$

where σ *is defined as in the case of OWA (*i.e., $a_{\sigma(i)}$ *is the ith largest element in the collection* $a_1,...,a_N$*), and the weight* ω_i *is defined as*

$$\omega_i = w^*(\sum_{j\leq i} p_{\sigma(j)}) - w^*(\sum_{j<i} p_{\sigma(j)}),$$

with w^* *being a nondecreasing function that interpolates the points*

$$\{(i/N, \sum_{j\leq i} w_j)\}_{i=1,...,N} \cup \{(0,0)\}.$$

The function w^* *is required to be a straight line when the points can be interpolated in this way.*

This definition requires an interpolation method for all the points in the set $\{(i/N, \sum_{j\leq i} w_j)\}_{i=1,...,N} \cup \{(0,0)\}$. The original definition used the interpolation method described in [14] although other interpolation methods, as e.g., linear interpolation, have been used. [16] compares different interpolation methods. See also [18, 19] for details on the WOWA operator and its interpolation function.

3.1.2 The Choquet Integral

The weighted mean, the OWA operator, and the WOWA operator are all operators from the same family. They are all particular cases of the Choquet integral. This is an operator that permits us to aggregate a set of data taking into account the interactions between the criteria. In this section we review its definition.

Its definition is based on considering that the importance is not only a function of a single criteria but of a set of them. That is, while in the weighted mean, given the criteria $C = \{c_1, c_2, \ldots, c_{|C|}\}$ we have weights $p_1, \ldots, p_{|C|}$, we have now *weights* for sets of criteria. Formally, in the weighted mean we have weights $p : C \to [0,1]$ such that $\sum_{c\in C} p(c) = 1$, and we use the notation $p_i = p(c_i)$. Thus, $p_i = p(c_i)$ is the importance of criteria c_i.

In contrast, in the Choquet integral, as well as in other fuzzy integrals, we can consider the importance of a set of criteria $\psi \subset C$. To do so, we need a function μ such that $\mu(\psi)$ corresponds to the importance of the set ψ. As in the case of the weighting vectors, $\mu(\psi) \in [0,1]$. That is, μ is a set function $\mu : \wp(C) \to [0,1]$. Formally, the function μ has to satisfy some constraints, and such set functions are known as fuzzy measures. We review their definition below.

Definition 5. *A fuzzy measure μ on a set C is a set function $\mu : \wp(C) \to [0,1]$ satisfying the following axioms:*

(i) $\mu(\emptyset) = 0$, $\mu(C) = 1$ *(boundary conditions)*
(ii) $A \subseteq B$ *implies* $\mu(A) \leq \mu(B)$ *(monotonicity)*

As stated above, μ is a set function into $[0,1]$. In addition, μ satisfies monotonicity. That is, the larger the set, the larger the measure, or, equivalently, the larger the set of criteria, the larger their importance. In addition, the maximum importance (equal to 1) is achieved for the whole set of criteria, and the minimum importance (equal to 0) is achived for the empty set.

Then, given a fuzzy measure, it is of relevance to consider the aggregation of the criteria taking into account the importance of these criteria. The Choquet integral permits us to do so. Nevertheless, the Choquet integral, as all the integrals, integrate functions. To permit to use them in aggregation we need a last transformation. This corresponds to consider the input data a_i as a function of the criteria. That is, let as above be C the set of criteria, then $f : C \to \mathbb{R}$ is the function that assigns a value to each criteria. In other words, if a_i is the value for the alternative to criteria c_i, then $f(c_i) = a_i$.

Using this function f, the evaluation of an alternative with respect to μ can be defined as the Choquet integral of f with respect to μ. We review below the definition of this integral.

Definition 6. *Let μ be a fuzzy measure on X; then, the* Choquet integral *of a function $f : X \to \mathbb{R}^+$ with respect to the fuzzy measure μ is defined by*

$$(C)\int f d\mu = \sum_{i=1}^{N} [f(x_{s(i)}) - f(x_{s(i-1)})]\mu(A_{s(i)}), \tag{2}$$

where $f(x_{s(i)})$ indicates that the indices have been permuted so that $0 \leq f(x_{s(1)}) \leq \cdots \leq f(x_{s(N)}) \leq 1$, and where $f(x_{s(0)}) = 0$ and $A_{s(i)} = \{x_{s(i)}, \ldots, x_{s(N)}\}$.

When no confusion exists, we can use $CI_\mu(a_1, \ldots, a_N) = (C)\int f d\mu$, where, $f(x_i) = a_i$, as before. There are alternative expressions for the Choquet integral that are equivalent to the one given above. The next proposition presents one of them.

Proposition 1. *Let μ be a fuzzy measure on X; then, the Choquet integral of a function $f : X \to \mathbb{R}^+$ with respect to μ can be expressed as*

$$(C)\int f d\mu = \sum_{i=1}^{N} f(x_{\sigma(i)})[\mu(A_{\sigma(i)}) - \mu(A_{\sigma(i-1)})], \tag{3}$$

where $\{\sigma(1), \ldots, \sigma(N)\}$ is a permutation of $\{1, \ldots, N\}$ such that $f(x_{\sigma(1)}) \geq f(x_{\sigma(2)}) \geq \cdots \geq f(x_{\sigma(N)})$, where $A_{\sigma(k)} = \{x_{\sigma(j)} | j \leq k\}$ (or, equivalently, $A_{\sigma(k)} = \{x_{\sigma(1)}, \ldots, x_{\sigma(k)}\}$ when $k \geq 1$ and $A_{\sigma(0)} = \emptyset$).

3.1.3 Other Fuzzy Integrals

The Choquet integral is not the only function that permits us to combine a set of values with respect to a fuzzy measure. There are other functions that permit us to do the same. One of them is the Sugeno integral. Although there are differences on the properties of the two integrals, there is a basic difference on the type of data they can aggregate. While the Choquet integral relies on addition, multiplication, and difference, and, thus, it is required the values to be numerical, the Sugeno integral only relies on maximum and minimum. Therefore, the Sugeno integral can be applied in any scale where maximum and minimum is meaningful. So, it is appropriate when the values are given in an ordinal scale.

From a formal point of view, the Sugeno integral integrates a function with respect to a fuzzy measure. Thus, it is similar to the Choquet integral.

From the point of view of their interpretation, the Choquet integral has a *probabilistic flavor* while the Sugeno integral has a *possibilistic flavor*. This is so, because the Choquet integral can be seen as a generalization of the expectation, and then the fuzzy measure is a generalization of a probability. This interpretation is also supported by the result which establishes that the Choquet integral with respect to an additive fuzzy measure (a fuzzy measure satisfying $\mu(A \cup B) = \mu(A) + \mu(B)$ for $A \cap B = \emptyset$) is equivalent to a weighted mean with respect to this measure. Note that in this case, μ is a probability and $\mu(\{c_i\}) = p(c_i)$ is a probability distribution. In contrast, similar results apply to Sugeno integrals with respect to possibility distributions.

The definition of the Sugeno integral is given below. We include also below an equivalent expression in terms of the permutation σ.

For details on the meaning of the Sugeno integral, and about its interpretation, see [17] and [18].

Definition 7. *[17] Let μ be a fuzzy measure on X; then, the* Sugeno integral *of a function $f : X \to [0,1]$ with respect to μ is defined by*

$$(S)\int f d\mu = \max_{i=1,N} \min(f(x_{s(i)}), \mu(A_{s(i)})), \tag{4}$$

where $f(x_{s(i)})$ indicates that the indices have been permuted so that $0 \leq f(x_{s(1)}) \leq \ldots \leq f(x_{s(N)}) \leq 1$ and $A_{s(i)} = \{x_{s(i)}, \ldots, x_{s(N)}\}$.

Proposition 2. *The Sugeno integral of a function $f : X \to [0,1]$ with respect to a fuzzy measure μ can be equivalently expressed by*

$$(S)\int f d\mu = \max_{i} \min(f(x_{\sigma(i)}), \mu(A_{\sigma(i)})),$$

where $A_{\sigma(k)} = \{x_{\sigma(j)} | j \leq k\}$ (or, equivalently, $A_{\sigma(k)} = \{x_{\sigma(1)}, \ldots, x_{\sigma(k)}\}$ when $k \geq 1$ and $A_{\sigma(0)} = \emptyset$), and where σ is a permutation such that $f(x_{\sigma(i)}) \geq f(x_{\sigma(i+1)})$ for $i \geq 1$.

A few generalizations have been defined for Choquet and Sugeno integrals. One of them is the t-conorm integral and another one is the twofold integral. The definition of the latter is given below. The twofold integral integrates a function with respect to two fuzzy measures μ_S and μ_C corresponding, respectively, to the fuzzy measures of the Sugeno integral and the Choquet integral. Somehow, the twofold integral generalizes the two fuzzy integrals in the same way as WOWA generalizes both weighted mean and OWA. That is, the generalization considers two parameters: **p** and **w** in the case of the WOWA, and μ_S and μ_C in the case of the twofold integral. Also as in the case of the WOWA, both parameters (i.e., both measures μ_S and μ_C) have the same *form/structure* but have a different meaning.

Definition 8. *[15, 6] Let μ_C and μ_S be two fuzzy measures on X; then, the* twofold integral *of a function $f : X \to [0,1]$ with respect to the fuzzy measures μ_S and μ_C is defined by*

$$TI_{\mu_S,\mu_C}(f) = \sum_{i=1}^{n} \Big(\big(\bigvee_{j=1}^{i} f(x_{s(j)}) \wedge \mu_S(A_{s(j)}) \big) \big(\mu_C(A_{s(i)}) - \mu_C(A_{s(i+1)}) \big) \Big),$$

where $f(x_{s(i)})$ indicates that the indices have been permuted so that $0 \le f(x_{s(1)}) \le \cdots \le f(x_{s(n)}) \le 1$, and where $A_{s(i)} = \{x_{s(i)}, \cdots, x_{s(n)}\}$ and $A_{s(n+1)} = \emptyset$.

For details and properties of this integral see [6]. [7] presents a graphical interpretation of the integral.

4 Conclusions

In this paper we have reviewed the use of aggregation operators in the field of multicriteria decision making. We have seen that they can be used to aggregate the values of the different criteria for each alternative. We have reviewed the main aggregation operators for numerical data. In particular, the arithmetic mean, the weighted mean, the OWA and the WOWA operator. We have underlined their differences and shown that the WOWA operator permits the user to combine the criteria taking into account the importance of each of the criteria (as the weighted mean does) and also the degree of compensation between the different criteria (as the OWA operator does).

In addition, we have also reviewed some fuzzy integrals. We have argued that fuzzy measures are suitable for representing the importance of sets of information sources, and then described both Choquet and Sugeno integrals as a way to aggregate the evaluation of the alternatives with respect to the fuzzy measures. We have underlined that the Sugeno integral permits us to use it when data is not represented using numerical scales but ordinal scales. We have reviewed the twofold integral, which generalizes both the Sugeno and the Choquet intregral. The way the generalization is done is somehow similar to the one of the WOWA in the sense that the generalization permits us to use the parameters of both Choquet and Sugeno integrals.

Acknowledgements. Partial support by the Generalitat de Catalunya (2009 SGR 7) and by the Spanish MEC (projects ARES – CONSOLIDER INGENIO 2010 CSD2007-00004 – and eAEGIS – TSI2007-65406-C03-02) is acknowledged.

References

1. Bouyssou, D., Marchant, T., Pirlot, M., Perny, P., Tsoukiàs, A., Vincke, P.: Evaluation and Decision Models: A Critical Perspective. Kluwer's International Series. Kluwer Academic Publishers, Dordrecht (2000)
2. Choquet, G.: Theory of capacities. Ann. Inst. Fourier 5, 131–295 (1953/1954)
3. Figuera, J., Mousseau, V., Roy, B.: Electre Methods. In: Figueira, J., Greco, S., Ehrgott, M. (eds.) Multi Criteria Decision Analysis: State of the Art Surveys. Springer, Heidelberg (2005)
4. Fodor, J., Roubens, M.: Fuzzy Preference Modelling and Multicriteria Decision Support. Kluwer Academic Publishers, Dordrecht (1994)
5. Geldermann, J., Rentz, O.: Bridging the Gap between American and European MADM-Approaches? In: Proc. of the 51st Meeting of the European Working Group "Multicriteria Aid for Decisions", Madrid (2000)
6. Narukawa, Y., Torra, V.: Twofold integral and Multi-step Choquet integral. Kybernetika 40(1), 39–50 (2004)
7. Narukawa, Y., Torra, V.: Graphical interpretation of the twofold integral and its generalization. Int. J. of Unc., Fuzz. and Knowledge Based Systems 13(4), 415–424 (2005)
8. Rapoport, A.: Decision Theory and Decision Behavior. Kluwer Academic Publishers, Dordrecht (1989)
9. Roy, B.: Multicriteria Methodology for Decision Aiding. Kluwer Academic Publishers, Dordrecht (1996)
10. Roy, B., Vanderpooten, D.: The European school of MCDA: Emergence, basic features and current works. J. of Multi-Criteria Decision Analysis 5(1), 22–38 (1998)
11. Sugeno, M.: Theory of Fuzzy Integrals and its Applications. Ph. D. Dissertation, Tokyo Institute of Technology, Tokyo, Japan (1974)
12. Tanino, T., Tanaka, T., Inuiguchi, M. (eds.): Multi-Objective Programming and Goal Programming. Springer, Heidelberg (2002)
13. Torra, V.: The weighted OWA operator. Int. J. of Intel. Syst. 12, 153–166 (1997)
14. Torra, V.: The WOWA operator and the interpolation function W*: Chen and Otto's interpolation method revisited. Fuzzy Sets and Systems 113(3), 389–396 (2000)
15. Torra, V.: La integral doble o *twofold* integral: Una generalització de les integrals de Choquet i Sugeno, Butlletí de l'Associació Catalana d'Inte·igència Artificial 29 13-19. Preliminary version in English: Twofold integral: A generalization of Choquet and Sugeno integral, IIIA Technical Report TR-2003-08 (2003)
16. Torra, V., Lv: On the WOWA operator and its interpolation function. Int. J. of Intel. Systems (2009) (in press), doi:10.1002/int.20371
17. Torra, V., Narukawa, Y.: The interpretation of fuzzy integrals and their application to fuzzy systems. Int. J. of Approx. Reasoning 41(1), 43–58 (2006)
18. Torra, V., Narukawa, Y.: Modeling decisions: information fusion and aggregation operators. Springer, Heidelberg (2007)

19. Torra, V., Narukawa, Y.: Modelització de decisions: fusió d'informació i operadors d'agregació. UAB Press (2007)
20. Valls, A.: ClusDM: A Multiple Criteria Decision Making Method for Heterogeneous Data Sets. Monografies de l'IIIA (2003)
21. Yager, R.R.: On ordered weighted averaging aggregation operators in multi-criteria decision making. IEEE Trans. on Systems, Man and Cybernetics 18, 183–190 (1988)
22. Yager, R.R., Kacprzyk, J. (eds.): The Ordered Weighted Averaging Operators: Theory and Applications. Springer, Heidelberg (1997)

Optimization and Aggregation Functions

Gleb Beliakov

1 Introduction

In this work we will look at connections between aggregation functions and optimization. There are two such connections: 1) aggregation functions are used to transform a multiobjective optimization problem into a single objective problem by aggregating several criteria into one, and 2) construction of aggregation functions often involves an optimization problem.

Aggregation of several inputs into one value arises in combining preferences given by several individuals of a group, aggregation of criteria in multicriteria decision problems, or fusion of possibly uncertain evidence provided by several sources. Aggregation functions, or operators, are functions that combine several input values into one output value, which can be used to rank the alternatives, among other purposes. Weighted arithmetic mean is one example of a commonly used aggregation function, but there are many alternative ways of combining the inputs. The overviews of many different types of aggregation functions are presented in [1, 30, 10, 16, 21]. We concentrate on aggregation functions that take the inputs from a closed interval, for convenience $[0,1]$, and produce the output in the same interval. Such aggregation functions are widely used in decision theory (cf. multiattribute utility functions), fuzzy logic, engineering, expert and decision support systems, and management science.

The choice of an aggregation function is application specific, and is frequently performed in ad hoc manner. One problem here is that the domain experts can rarely specify how they perform aggregation by means of an algebraic formula. For instance, decision support systems in medical domain rely on aggregation of evidence given as various symptoms, but doctors would not specify the exact formula. With today's automatic collection of vast amounts of data, it is possible to

Gleb Beliakov
School of Information Technology, Deakin University
221 Burwood Hwy, Burwood 3125, Australia
e-mail: gleb@deakin.edu.au

W.A. Lodwick & J. Kacprzyk (Eds.): Fuzzy Optimization, STUDFUZZ 254, pp. 77–108.
springerlink.com

extract many aggregation rules from databases. An example here is the use of preferences of e-commerce customers in recommender systems. Such systems recommend customers a number of products that match their preferences, and the strength of recommendation is based on aggregating the degrees to which the preferences are matched by individual products. An analysis of customers' responses to recommendations can provide suitable aggregation rules.

It is possible to construct suitable application-specific aggregation functions from the recorded data by solving, essentially, a regression problem. In the case of weighted mean operators, it boils down to a standard quadratic programming problem, but for other aggregation functions the situation is different. The issue is preservation of semantically important properties of aggregation functions, without which they would provide inconsistent, and even erroneous output. This is the reason why many off-the-shelf nonlinear regression methods, such as neural networks, do not work, as they fail to account for such properties. In this contribution we outline various alternative methods suitable for construction of aggregation functions.

2 Aggregation Functions

The purpose of aggregation functions is to combine inputs that are typically interpreted as degrees of membership in fuzzy sets, degrees of preference, strength of evidence, or support of a hypothesis, and so on. Consider these prototypical examples.

Example 1 (A multicriteria decision making problem)

There are two (or more) alternatives, and n criteria to evaluate each alternative (or rather a preference for each alternative). Denote the scores (preferences) by $x_1, x_2, \ldots, x_n$ and $y_1, y_2, \ldots, y_n$ for the alternatives x and y respectively. The goal is to combine these scores using some aggregation function f, and to compare the values $f(x_1, x_2, \ldots, x_n)$ and $f(y_1, y_2, \ldots, y_n)$ to decide the best alternative.

Example 2 (Connectives in fuzzy logic)

An object d has partial degrees of membership to n fuzzy sets, denoted $\mu_1, \mu_2, \ldots, \mu_n$. The goal is to obtain the overall membership value in the combined fuzzy set $\mu = f(\mu_1, \mu_2, \ldots, \mu_n)$. The combination can be the set operation of union, intersection, or a more complicated (e.g., composite) operation.

Example 3 (A rule based system)

The system contains rules of the form

If t_1 *is* A_1 **AND** t_2 *is* A_2 **AND** $\ldots$ t_n *is* A_n **THEN ...**

$x_1, x_2, \ldots, x_n$ denote the degrees of satisfaction of the rule predicates t_1 is A_1, t_2 is A_2, etc. The goal is to calculate the overall degree of satisfaction of the combined predicate of the rule antecedent $f(x_1, x_2, \ldots, x_n)$.

We will consider aggregation functions defined on the unit interval with $f : [0,1]^n \to [0,1]$, however other choices are possible. The input value 0 is interpreted as no membership, no preference, no evidence, no satisfaction, etc., and naturally, an aggregation of n 0s should yield 0. Similarly, the value 1 is interpreted as full membership (strongest preference, evidence), and an aggregation of 1s should naturally yield 1. Aggregation functions also require monotonicity in each argument, where an increase to any input cannot result in a decrease in the overall score.

Definition 1 (Aggregation function). *An aggregation function is a function of $n > 1$ arguments that maps the (n-dimensional) unit cube onto the unit interval $f : [0,1]^n \to [0,1]$, with the properties*
(i) $f(\underbrace{0,0,\ldots,0}_{n-times}) = 0$ *and* $f(\underbrace{1,1,\ldots,1}_{n-times}) = 1$.
(ii) $\mathbf{x} \leq \mathbf{y}$ *implies* $f(\mathbf{x}) \leq f(\mathbf{y})$ *for all* $\mathbf{x}, \mathbf{y} \in [0,1]^n$.

For some applications, the inputs may have a varying number of components (for instance, some values can be missing). Particularly in the case of automated systems, it may be desirable to utilize functions defined for $n = 2,3,\ldots$ arguments with the same underlying property in order to give consistent aggregation results. Functions satisfying the following definition may then be worth considering.

Definition 2 (Extended aggregation function). *An extended aggregation function is a mapping*

$$F : \bigcup_{n \in \{1,2,\ldots\}} [0,1]^n \to [0,1],$$

such that the restriction of this mapping to the domain $[0,1]^n$ for a fixed n is an n-ary aggregation function f, with the convention $F(x) = x$ for $n = 1$.

2.1 Main Types

Aggregation functions are classed depending on their overall behavior in relation to the inputs. In some cases we require high inputs to compensate for low inputs, or that inputs may average each other. In other situations, it may make more sense that high scores reinforce each other and low inputs are essentially discarded. The four main classes of aggregation functions are *Averaging, Conjunctive, Disjunctive,* and *Mixed* [16, 20, 21]. We provide the mathematical definitions for each below.

Definition 3. *An aggregation function f is said to be*

- *Averaging if for every* $\mathbf{x}$ *it is bounded by*

$$\min(\mathbf{x}) \leq f(\mathbf{x}) \leq \max(\mathbf{x}).$$

- *Conjunctive if for every* $\mathbf{x}$ *it is bounded by*

$$f(\mathbf{x}) \leq \min(\mathbf{x}) = \min(x_1, x_2, \ldots, x_n).$$

- *Disjunctive if for every* $\mathbf{x}$ *it is bounded by*

$$f(\mathbf{x}) \geq \max(\mathbf{x}) = \max(x_1, x_2, \ldots, x_n).$$

- *Mixed if it does not belong to any of the above classes, i.e., it exhibits different types of behavior on different parts of the domain.*

The use of averaging functions is very prominent. Weighted means (arithmetic, geometric, power means) are typical examples, but there are many other averaging functions, such as medians, ordered weighted averaging (OWA), discrete Choquet and Sugeno integrals and many others [15]. The semantics of some mixed aggregation functions makes their use appealing, for instance, when we want inputs of only high scores to reinforce each other while scores all below a given threshold are penalized. MYCIN [14] is a classical expert system used to diagnose and treat rare blood diseases and utilizies precisely this type of aggregation.

There is also wide potential for the use of conjunctive or disjunctive functions in fuzzy rule-based recommender systems. Conjunctive and disjunctive aggregation functions are often studied in parallel, as they can satisfy very similar properties, just viewed from a different angle. The properties of conjunctive functions can be studied and then mapped to their disjunctive equivalents by using the concept of a dual aggregation function. There are also aggregation functions that are self-dual, in particular many averaging functions and uninorms. Before giving these definitions, we need first to define the concept a strong negation.

Definition 4 (Strong negation). *A univariate function* N *defined on* $[0,1]$ *is called a strong negation, if it is strictly decreasing and involutive (i.e.,* $N(N(t)) = t$ *for all* $t \in [0,1]$*).*

The most commonly used strong negation is the standard negation

$$N(t) = 1 - t.$$

Definition 5 (Dual aggregation function). *Let* $N : [0,1] \to [0,1]$ *be a strong negation and* $f : [0,1]^n \to [0,1]$ *an aggregation function. Then the aggregation function* f_d *given by*

$$f_d(x_1, \ldots, x_n) = N(f(N(x_1), N(x_2), \ldots, N(x_n)))$$

is called the dual of f *with respect to* N*, or, for short, the* N*-dual of* f*. When using the standard negation,* f_d *is given by*

$$f_d(x_1, \ldots, x_n) = 1 - f(1 - x_1, \ldots, 1 - x_n)$$

and we will simply say that f_d *is the dual of* f*.*

The dual of a conjunctive aggregation function is disjunctive, and vice versa. Averaging functions, uninorms and nullnorms are closed under duality. Some functions are *self-dual*.

Definition 6 (Self-dual aggregation function). *Given a strong negation N, an aggregation function f is self-dual with respect to N (for short, N-self-dual or N-invariant), if*

$$f(\mathbf{x}) = N(f(N(\mathbf{x}))),$$

where $N(\mathbf{x}) = (N(x_1), \ldots, N(x_n))$. For the standard negation we have

$$f(\mathbf{x}) = 1 - f(\mathbf{1} - \mathbf{x}),$$

and it is simply said that f is self-dual.

2.2 Main Properties

There are several studied properties that can be satisfied by aggregation functions, making them useful in certain situations. We provide definitions for those that are frequently referred to in the literature.

Definition 7 (Idempotency). *An aggregation function f is called idempotent if for every input $\mathbf{x} = (t,t,\ldots,t), t \in [0,1]$ the output is $f(t,t\ldots,t) = t$.*

Definition 8 (Symmetry/Anonymity). *An aggregation function f is called symmetric, if its value does not depend on the permutation of the arguments, i.e.,*

$$f(x_1,x_2,\ldots,x_n) = f(x_{P(1)},x_{P(2)},\ldots,x_{P(n)}),$$

for every $\mathbf{x}$ and every permutation $P = (P(1),P(2),\ldots,P(n))$ of $(1,2\ldots,n)$.

Definition 9 (Neutral element). *An aggregation function f has a neutral element $e \in [0,1]$, if for every $t \in [0,1]$ in any position it holds*

$$f(e,\ldots,e,t,e,\ldots,e) = t.$$

Definition 10 (Absorbing element (annihilator)). *An aggregation function f has an absorbing element $a \in [0,1]$ if*

$$f(x_1,\ldots,x_{i-1},a,x_{i+1},\ldots,x_n) = a,$$

for every $\mathbf{x}$ such that $x_i = a$ with a in any position.

Definition 11 (Associativity). *A two-argument function f is associative if $f(f(x_1,x_2),x_3) = f(x_1,f(x_2,x_3))$ holds for all x_1,x_2,x_3 in its domain.*

Associative functions are convenient when working with high or varying dimensions, as computation of any number of arguments can be automatically calculated from the two variate case. The triangular norms and triangular conorms, uninorms and nullnorms (presented in Sect. 2.3) are examples of aggregation functions which have this property.

Another property that is useful for aggregation functions is continuity, which ensures stability in the outputs. In particular, the Lipschitz condition ensures that the change in the aggregated value of the function is bounded. We present the following definitions:

Definition 12 (Lipschitz continuity). *An aggregation function f is called Lipschitz continuous if there is a positive number M, such that for any two vectors $\mathbf{x},\mathbf{y}$ in the domain of definition of f:*

$$|f(\mathbf{x}) - f(\mathbf{y})| \leq Md(\mathbf{x},\mathbf{y}), \tag{1}$$

where $d(\mathbf{x},\mathbf{y})$ is a distance between $\mathbf{x}$ and $\mathbf{y}$ The smallest such number M is called the Lipschitz constant of f (in the distance d).

Typically the distance is the Euclidean distance between vectors, but it can be chosen as any norm $d(\mathbf{x},\mathbf{y}) = ||\mathbf{x}-\mathbf{y}||$. Typically it is chosen as a p-norm, $p \geq 1$: $||\mathbf{x}||_p = \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p}$, for finite p, and $||\mathbf{x}||_\infty = \max_{i=1\ldots,n} |x_i|$. We will denote the class of functions with the Lipschitz constant at most M in the norm $||\cdot||$ by $Lip(M,||\cdot||)$.

We pay attention to the rate of change of a function because of the ever present input inaccuracies. If the aggregation function receives an inaccurate input $\tilde{\mathbf{x}} = (x_1+\delta_1,\ldots,x_n+\delta_n)$, contaminated with some error $(\delta_1,\ldots,\delta_n)$, we do not expect the output $f(\tilde{\mathbf{x}})$ to be substantially different from $f(\mathbf{x})$. The Lipschitz constant M bounds the factor by which the error is magnified.

Definition 13. *An aggregation function f is called 1-Lipschitz if it is p-stable with $p=1$, i.e., for all $\mathbf{x},\mathbf{y}$:*

$$|f(\mathbf{x}) - f(\mathbf{y})| \leq |x_1-y_1| + |x_2-y_2| + \ldots + |x_n-y_n|.$$

An aggregation function f is called kernel if it is p-stable with $p=\infty$, i.e., for all $\mathbf{x},\mathbf{y}$:

$$|f(\mathbf{x}) - f(\mathbf{y})| \leq \max_{i=1,\ldots,n} |x_i - y_i|.$$

2.3 Examples

2.3.1 Averaging

Means

Means are averaging aggregation functions. Formally, a mean is simply a function f with the property [15] $\min(\mathbf{x}) \leq f(\mathbf{x}) \leq \max(\mathbf{x})$. Still there are other properties that define one or another family of means. The following definition of a weighting vector will assist us in defining some of these families.

Definition 14 (Weighting vector). *A vector* $\mathbf{w} = (w_1, \ldots, w_n)$ *is called a weighting vector if* $w_i \in [0,1]$ *and* $\sum_{i=1}^{n} w_i = 1$.

We provide here the weighted definitions for some of the families of means. Where the weights are equal, the standard function will be returned and satisfy the symmetry property. Usually the weight allocated to a particular input is indicative of the importance of that particular input.

Definition 15 (Weighted arithmetic mean). *Given a weighting vector* $\mathbf{w}$*, the weighted arithmetic mean is the function*

$$M_{\mathbf{w}}(\mathbf{x}) = w_1 x_1 + w_2 x_2 + \ldots + w_n x_n = \sum_{i=1}^{n} w_i x_i.$$

Definition 16 (Weighted geometric mean). *Given a weighting vector* $\mathbf{w}$*, the weighted geometric mean is the function*

$$G_{\mathbf{w}}(\mathbf{x}) = \prod_{i=1}^{n} x_i^{w_i}.$$

Definition 17 (Weighted power mean). *Given a weighting vector* $\mathbf{w}$*and* $r \in \Re$*, the weighted power mean is the function*

$$M_{\mathbf{w},[r]}(\mathbf{x}) = \left(\sum_{i=1}^{n} w_i x_i^r \right)^{1/r},$$

if $r \neq 0$*, and* $M_{\mathbf{w},[0]}(\mathbf{x}) = G_{\mathbf{w}}(\mathbf{x})$.

Definition 18 (Weighted quasi-arithmetic mean). *For a given strictly monotone and continuous function* $g : [0,1] \to [-\infty, +\infty]$*, called a generating function or generator, and a weighting vector* $\mathbf{w}$*, the weighted quasi-arithmetic mean is the function*

$$M_{\mathbf{w},g}(\mathbf{x}) = g^{-1}\left(\sum_{i=1}^{n} w_i g(x_i) \right). \tag{2}$$

The weighted quasi-arithmetic mean is an example of a *generated function*, defined by use of a generator g and its inverse g^{-1}. Other generated functions include Archimedean triangular norms and representable uninorms, which will be discussed further on.

Medians

Medians are often used where the aggregation may be skewed by outliers. A common example is in median housing prices, where it is desired that the aggregated value not be altered significantly by only a few very expensive houses.

Definition 19 (Weighted median). *Let* $\mathbf{w}$ *be a weighting vector, and let* $\mathbf{u}$ *denote the vector obtained from* $\mathbf{w}$ *by arranging its components in the order induced by the components of the input vector* $\mathbf{x}$*, such that* $u_k = w_i$ *if* $x_i = x_{(k)}$ *is the k-th largest component of* $\mathbf{x}$*. The lower weighted median is the function*

$$Med_{\mathbf{w}}(\mathbf{x}) = x_{(k)}, \tag{3}$$

where k is the index obtained from the condition

$$\sum_{j=1}^{k-1} u_j < \frac{1}{2} \text{ and } \sum_{j=1}^{k} u_j \geq \frac{1}{2}. \tag{4}$$

The upper weighted median is the function in Eq. (3) where k is the index obtained from the condition

$$\sum_{j=1}^{k-1} u_j \leq \frac{1}{2} \text{ and } \sum_{j=1}^{k} u_j > \frac{1}{2}.$$

Ordered Weighted Averaging

Ordered weighted averaging functions (OWA) are also averaging aggregation functions, which associate a weight not with a particular input, but rather with its relative value or order compared to others. They have been introduced by Yager [43] and have become very popular in the fuzzy sets community.

Definition 20 (OWA). *Given a weighting vector* $\mathbf{w}$*, the OWA function is*

$$OWA_{\mathbf{w}}(\mathbf{x}) = \sum_{i=1}^{n} w_i x_{(i)},$$

where the $(.)$ *notation denotes the components of* $\mathbf{x}$ *being arranged in non-increasing order* $x_{(1)} \geq x_{(2)} \geq \ldots \geq x_{(n)}$.

Special cases of the OWA operator, depending on the weighting vector $\mathbf{w}$ include:

- The *arithmetic mean* where all the weights are equal, i.e. all $w_i = \frac{1}{n}$.
- The *maximum* function for $\mathbf{w} = (1,0,...,0)$ and the *minimum* for $\mathbf{w} = (0,...,0,1)$.
- The *median* function for $w_i = 0$ for all $i \neq k$, $w_k = 1$ if $n = 2k+1$ is odd, and $w_i = 0$ for all $i \neq k, k+1$, $w_k = w_{k+1} = 0.5$ if $n = 2k$ is even.

Choquet and Sugeno Integrals

These are two classes of averaging aggregation functions defined with respect to a fuzzy measure. They are useful to model interactions between the variables x_i.

Definition 21 (Fuzzy measure). *Let* $\mathcal{N} = \{1,2,\ldots,n\}$*. A discrete fuzzy measure is a set function* $v : 2^{\mathcal{N}} \to [0,1]$ *which is monotonic (i.e.* $v(A) \leq v(B)$ *whenever* $A \subseteq B$*) and satisfies* $v(\emptyset) = 0, v(\mathcal{N}) = 1$.

Definition 22 (Choquet integral). *The discrete Choquet integral with respect to a fuzzy measure v is given by*

$$C_v(\mathbf{x}) = \sum_{i=1}^{n} x_{(i)}[v(\{j|x_j \geq x_{(i)}\}) - v(\{j|x_j \geq x_{(i+1)}\})], \tag{5}$$

where $(.)$ in this case denotes the components of $\mathbf{x}$ being arranged in non-decreasing order such that $(x_{(1)} \leq x_{(2)} \leq \cdots \leq x_{(n)})$ (note that this is opposite to OWA).

By rearranging the terms of the sum, Eq. (5) can also be written as

$$C_v(\mathbf{x}) = \sum_{i=1}^{n} \left[x_{(i)} - x_{(i-1)}\right] v(H_i). \tag{6}$$

where $x_{(0)} = 0$ by convention, and $H_i = \{(i), \ldots, (n)\}$ is the subset of indices of $n - i + 1$ largest components of $\mathbf{x}$.

The class of Choquet integrals includes weighted arithmetic means and OWA functions as special cases. The Choquet integral is a piecewise linear idempotent function, uniquely defined by its values at the vertices of the unit cube $[0,1]^n$. Note that there are 2^n such points, the same as the number of values that determine the fuzzy measure v.

The fuzzy measure used to define the Choquet integral can be interpreted as a weight allocation, not merely to individual inputs but rather to each subset of inputs. It may be that there are redundancies among the inputs, or that certain inputs complement each other. The following properties of fuzzy measures are useful for identifying special cases of the Choquet integral and certain behavior.

Certain indices have been introduced in order to better understand the behavior of the Choquet integral. In particular, the Shapley value gives an indication of the overall importance of a given input, while the interaction index between two inputs shows to what extent they are redundant or complimentary.

Definition 23 (Shapley value)
Let v be a fuzzy measure. The Shapley index for every $i \in \mathcal{N}$ is

$$\phi(i) = \sum_{A \subseteq \mathcal{N} \setminus \{i\}} \frac{(n - |A| - 1)!|A|!}{n!}[v(A \cup \{i\}) - v(A)].$$

The Shapley value is the vector $\phi(v) = (\phi(1), \ldots, \phi(n))$.

Definition 24 (Interaction index). *Let v be a fuzzy measure. The interaction index for every pair $i, j \in \mathcal{N}$ is*

$$I_{ij} = \sum_{A \subseteq \mathcal{N} \setminus \{i,j\}} \frac{(n - |A| - 2)!|A|!}{(n-1)!}[v(A \cup \{i,j\}) - v(A \cup \{i\}) - v(A \cup \{j\}) + v(A)].$$

Where the interaction index is negative, there is some redundancy between the two inputs. Where it is positive, the inputs complement each other to some degree and their weight together is worth more than their combined individual weights.

The Choquet integral has been predominantly used for numerical inputs, the Sugeno integral defined below is useful where the inputs are ordinal. It also uses fuzzy measures for its definition.

Definition 25 (Sugeno integral). *The Sugeno integral with respect to a fuzzy measure v is given by*

$$S_v(\mathbf{x}) = \max_{i=1,\ldots,n} \min\{x_{(i)}, v(H_i)\}, \tag{7}$$

where $(.)$ denotes a non-decreasing permutation of the inputs such that $(x_{(1)} \leq x_{(2)} \leq \cdots \leq x_{(n)})$ (the same as with the Choquet integral), and $H_i = \{(i), \ldots, (n)\}$.

In the special case of a symmetric fuzzy measure (i.e., when $v(H_i) = v(|H_i|)$ depends only on the cardinality of the set H_i), Sugeno integral becomes the median $S_v(\mathbf{x}) = Med(x_1, \ldots, x_n, 1, v(n-1), v(n-2), \ldots, v(1))$.

2.3.2 Conjunctive and Disjunctive Functions

The prototypical examples of conjunctive and disjunctive aggregation functions are so-called triangular norms and conorms respectively (t-norms and t-conorms) [30]. T-conorms are dual to t-norms in the sense of Def. 5. Triangular norms are associative, symmetric with the neutral element $e = 1$, whereas trinagular conorms are associative, symmetric and have the neutral element $e = 0$.

Definition 26. *The four basic t-norms, T_{min}, T_P, T_L and T_D are given by*

$$\begin{aligned} T_{min}(x_1,x_2) &= \min(x_1,x_2), && (minimum)\\ T_P(x_1,x_2) &= x_1x_2, && (product)\\ T_L(x_1,x_2) &= \max(x_1+x_2-1,0), && (\text{Łukasiewicz } t-norm) \end{aligned}$$

$$T_D(x_1,x_2) = \begin{cases} 0, & if\ (x_1,x_2) \in [0,1[^2,\\ \min(x_1,x_2) & otherwise. \end{cases} \quad (drastic\ product)$$

Definition 27 (The four basic t-conorms). *The four basic t-conorms, S_{max}, S_P, S_L and S_D are given by*

$$\begin{aligned} S_{max}(x_1,x_2) &= \max(x_1,x_2), && (maximum)\\ S_P(x_1,x_2) &= x_1+x_2-x_1x_2, && (probabilistic sum)\\ S_L(x_1,x_2) &= \min(x_1+x_2,1), && (\text{Łukasiewicz } t-conorm) \end{aligned}$$

$$S_D(x_1,x_2) = \begin{cases} 1, & if\ (x_1,x_2) \in]0,1]^2,\\ \max(x_1,x_2) & otherwise. \end{cases} \quad (drastic\ sum)$$

There are families of parameterized t-norms and t-conorms that include the above as special or limiting cases. These families are defined with respect to generating functions and are known as Archimedean t-norms.

Definition 28 (Archimedean t-norm). *A t-norm is called Archimedean if for each* $(a,b) \in]0,1[^2$ *there is an* $n = \{1,2,...\}$ *with* $T(\overbrace{a,...,a}^{n-times}) < b$.

For t-conorms, the inequality is reversed, i.e. the t-conorm $S > b$. Continuous Archimedean t-norms can be expressed by use of their generators as

$$T(x_1,...,x_n) = g^{(-1)}(g(x_1) + ... + g(x_n))$$

Where $g : [0,1] \to [0,\infty]$ with $g(1) = 0$ is a continuous, strictly decreasing function and $g^{(-1)}$ is the pseudo inverse of g, i.e.,

$$g^{(-1)}(x) = g^{-1}(\min(g(1), \max(g(0), x)))$$

Archimedean families include Schweizer-Sklar, Hamacher, Frank, Yager, Dombi, Aczel-Alsina, Mayor-Torrens and Weber-Sugeno t-norms and t-conorms. The Einstein sum given below in its bivariate case belongs to the Hamacher family

$$f(x_1,x_2) = \frac{x_1 + x_2}{1 + x_1 x_2}.$$

There do, of course, exist other conjunctive and disjunctive functions that are not t-norms or t-conorms. For instance, the function

$$f(x_1,x_2) = x_1 x_2^2$$

is a conjunctive ($x_1 x_2^2 \leq x_1 x_2 \leq \min(x_1,x_2)$), asymmetric aggregation function. It is not a t-norm.

2.3.3 Mixed Aggregation

In some situations, it may be required that high input values reinforce each other whereas low values pull the overall output down. In other words, the aggregation function has to be disjunctive for high values, conjunctive for low values, and perhaps averaging if some values are high and some are low. This is typically the case when high values are interpreted as "positive" information, and low values as "negative" information. Uninorms and nullnorms are typical examples of such aggregation functions, but there are many others.

Definition 29 (Nullnorm). *A nullnorm is a bivariate aggregation function* $V : [0,1]^2 \to [0,1]$ *which is associative, symmetric, such that there exists an element a belonging to the open interval* $]0,1[$ *verifying*

$$\forall t \in [0,a], \quad V(t,0) = t,$$

$$\forall t \in [a,1], \quad V(t,1) = t.$$

Definition 30 (Uninorm). *A uninorm is a bivariate aggregation function* $U : [0,1]^2 \to [0,1]$ *which is associative, symmetric and has a neutral element e belonging to the open interval [0, 1].*

Some uninorms can be built from generating functions in a similar way to quasi-arithmetic means and Archimedean t-norms. These are called representable uninorms.

Definition 31 (Representable uninorm). *Let* $u : [0,1] \to [-\infty, +\infty]$ *be a strictly increasing bijection verifying* $g(0) = -\infty, g(1) = +\infty$ *such that* $g(e) = 0$ *for some* $e \in]0,1[$.

- *The function given by*

$$U(x,y) = \begin{cases} g^{-1}(g(x)+g(y)), & \text{if } (x,y) \in [0,1]^2 \backslash \{(0,1),(1,0)\}, \\ 0, & \text{otherwise.} \end{cases}$$

  is a conjunctive uninorm with the neutral element e, known as a conjunctive representable uninorm.
- *The function given by*

$$U(x,y) = \begin{cases} g^{-1}(g(x)+g(y)), & \text{if } (x,y) \in [0,1]^2 \backslash \{(0,1),(1,0)\}, \\ 1, & \text{otherwise.} \end{cases}$$

  is a disjunctive uninorm with the neutral element e, known as a disjunctive representable uninorm.

The $3-\Pi$ function is an example of a representable uninorm [44]. It uses a generating function $g(x) = \ln(\frac{x}{1-x})$ and is used by the expert system PROSPECTOR [22] for combining uncertainty factors.

$$f(\mathbf{x}) = \frac{\prod_{i=1}^{n} x_i}{\prod_{i=1}^{n} x_i + \prod_{i=1}^{n} (1-x_i)},$$

with the convention $\frac{0}{0} = 0$. It is conjunctive on $[0, \frac{1}{2}]^n$, disjunctive on $[\frac{1}{2}, 1]^n$ and averaging elsewhere. It is associative, with the neutral element $e = \frac{1}{2}$, and discontinuous on the boundaries of $[0,1]^n$.

3 Aggregation Functions in Optimization Problems

A typical multiobjective optimization problem looks like this [36]:

$$\begin{aligned} &\text{minimize} && \{F_1(\mathbf{x}), F_2(\mathbf{x}), \ldots F_n(\mathbf{x})\} \\ &\text{subject to: linear or nonlinear constraints on } \mathbf{x}. \end{aligned}$$

One way of solving such problems is to transform the multiojective problem to a single aggregate objective problem

$$\begin{aligned} &\text{minimize} && f(F_1(\mathbf{x}), F_2(\mathbf{x}), \ldots F_n(\mathbf{x})) \\ &\text{subject to: linear or nonlinear constraints on } \mathbf{x}, \end{aligned}$$

where f is an aggregation function. Notice the importance of monotonicity of f. The aggregation function used most widely is the weighted arithmetic mean.

We note that this approach is subjective, as the solution depends on the weights chosen by the decision maker. Further, it cannot identify all non-dominant solutions. However its simplicity, and the fact that off-the-shelf single objective optimization methods can be used, provide it with a strong advantage.

Clearly, the use of weighted mean is warranted only if the objectives are independent, which is not always true in practice. For objectives which are mutually dependent, the use of Choquet integral is justified. Here one can account not only for importance of individual criteria, but also for importance of groups of criteria (coalitions).

Another application of aggregation functions is in constrained optimization problems

$$\begin{aligned} &\text{minimize} && F(\mathbf{x}) \\ &\text{subject to: } && \mathbf{g}(\mathbf{x}) \leq \mathbf{0}. \end{aligned}$$

Penalty function approach transforms this problem into

$$\text{minimize } F(\mathbf{x}) + a_1 g_1(\mathbf{x}) + \ldots a_m g_m(\mathbf{x}).$$

Clearly we can use any other aggregation function f to combine the objective and the constraints

$$\text{minimize } f(F(\mathbf{x}), g_1(\mathbf{x}), \ldots, g_m(\mathbf{x})).$$

4 Optimization Methods for Construction of Aggregation Functions

In this section we review methods of construction of aggregation functions based on optimization. In applications one faces a difficult task of choosing the most appropriate aggregation function for a specific problem. One has to discriminate among different aggregation functions based on some criteria. The criteria may relate to the mathematical properties of an aggregation function, or to some data, such as the desired or observed values of an aggregation function at some points.

4.1 Determination of Weighting Vectors

One important class of aggregation functions are OWA functions (Definition 20). There are two quantities associated with OWA weighting vectors, the orness value and the entropy. The orness value specifies how far is the OWA function from the *max* function.

Definition 32. *The orness value of an OWA function is*

$$orness(OWA_{\mathbf{w}}) = \sum_{i=1}^{n} w_i \frac{n-i}{n-1} = OWA_{\mathbf{w}}(1, \frac{n-2}{n-1}, \ldots, \frac{1}{n-1}, 0). \tag{8}$$

Clearly $orness(\max) = 1$, $orness(\min) = 0$, and for the arithmetic mean $orness(M) = \frac{1}{2}$. If the weighting vector is non-decreasing, i.e., $w_i \leq w_{i+1}, i = 1, \ldots, n-1$, then $orness(OWA_{\mathbf{w}}) \in [\frac{1}{2}, 1]$. If the weighting vector is non-increasing, then $orness(OWA_{\mathbf{w}}) \in [0, \frac{1}{2}]$.

Definition 33. *The weights dispersion (entropy) of an OWA function is*

$$Disp(\mathbf{w}) = -\sum_{i=1}^{n} w_i \log w_i.$$

It measures the degree to which all the information (i.e., all the inputs) is used in the aggregation process.

- If the orness is not specified, the maximum of $Disp$ is achieved at $w_i = \frac{1}{n}$, i.e., the arithmetic mean, and $Disp(\frac{1}{n}, \ldots, \frac{1}{n}) = \log n$.
- The minimum value of $Disp$, 0, is achieved if and only if $w_i = 0$, $i \neq k$, and $w_k = 1$, i.e., the order statistic.
- The entropy of an OWA and its dual (reverse OWA) coincide, $Disp(\mathbf{w}) = Disp(\mathbf{w}_d)$.

4.1.1 Maximum Entropy OWA

One approach to choosing OWA weights was proposed in [37] and followed in [23]. It uses various measures of weight entropy or dispersion. The idea is to choose for a given n such a vector of weights that maximizes the dispersion $Disp(\mathbf{w})$.

It is formulated as an optimization problem

$$\begin{aligned} \min \quad & \sum_{i=1}^{n} w_i \log w_i \\ \text{s.t.} \quad & \sum_{i=1}^{n} w_i = 1, \\ & \sum_{i=1}^{n} w_i \tfrac{n-i}{n-1} = \alpha, \\ & w_i \geq 0, i = 1, \ldots, n. \end{aligned} \tag{9}$$

The solution is provided in [23] and is called Maximum Entropy OWA (MEOWA). Using the method of Lagrange multipliers, the authors obtain the following expressions for w_i:

$$w_i = (w_1^{n-i} w_n^{i-1})^{\frac{1}{n-1}}, i = 2, \ldots, n-1, \tag{10}$$

$$w_n = \frac{((n-1)\alpha - n)w_1 + 1}{(n-1)\alpha + 1 - nw_1},$$

and w_1 being the unique solution to the equation

$$w_1[(n-1)\alpha + 1 - nw_1]^n = ((n-1)\alpha)^{n-1}[((n-1)\alpha - n)w_1 + 1] \tag{11}$$

on the interval $(0, \frac{1}{n})$. For $n = 3$, we obtain $w_2 = \sqrt{w_1 w_3}$ independently of the value of α.

A different representation of the same solution was given in [19]. Let t be the (unique) positive solution to the equation

$$dt^{n-1} + (d+1)t^{n-2} + \ldots + (d+n-2)t + (d+n-1) = 0, \tag{12}$$

with $d = -\alpha(n-1)$. Then the MEOWA weights are identified from

$$w_i = \frac{t^i}{T}, \; i = 1, \ldots, n, \quad \text{where } T = \sum_{j=1}^{n} t^j. \tag{13}$$

It is not difficult to check that both (10) and (13) represent the same set of weights, noting that $t = \sqrt[n-1]{\frac{w_n}{w_1}} = -\frac{1-d-nw_1}{d}$, or $w_1 = \frac{1+td-d}{n}$, and that substituting w_1 into (11) yields

$$1 - t^n = \frac{n(1-t)}{1-d(1-t)},$$

which translates into

$$\frac{1-t^n}{1-t} - d(1-t^n) - n = 0,$$

and then into

$$dt^n + t^{n-1} + t^{n-2} + \ldots + t + (1-d-n) = 0.$$

After factoring out $(t-1)$ we obtain (12).

4.1.2 Minimum Variance OWA

Another popular characteristic of weighting vector is weights variance [24]

$$D^2(\mathbf{w}) = \frac{1}{n}\sum_{i=1}^{n}(w_i - M(\mathbf{w}))^2 = \frac{1}{n}\sum_{i=1}^{n} w_i^2 - \frac{1}{n^2}, \tag{14}$$

where $M(\mathbf{w})$ is the arithmetic mean of $\mathbf{w}$.

Here one minimizes $D^2(\mathbf{w})$ subject to given orness value. The resulting OWA function is called Minumum Variance OWA (MVOWA). Since adding a constant to the objective function does not change the minimizer, this is equivalent to the problem

$$\begin{aligned} \min \quad & \sum_{i=1}^{n} w_i^2 \\ \text{s.t.} \quad & \sum_{i=1}^{n} w_i \tfrac{n-i}{n-1} = \alpha, \\ & \sum_{i=1}^{n} w_i = 1, w_i \geq 0, i = 1, \ldots, n. \end{aligned} \tag{15}$$

For $\alpha = \frac{1}{2}$ the optimal solution is always $w_j = \frac{1}{n}, j = 1, \ldots, n$. It is also worth noting that the optimal solution to (15) for $\alpha > \frac{1}{2}$, $\mathbf{w}^*$, is related to the optimal solution for $\alpha < \frac{1}{2}$, $\mathbf{w}$, by $w_i^* = w_{n-i+1}$, i.e., it gives the reverse OWA. Thus it is sufficient to establish the optimal solution in the case $\alpha < \frac{1}{2}$.

The optimal solution [24,32] for $\alpha < \frac{1}{2}$ is given as the vector $\mathbf{w} = (0, 0, \ldots, 0, w_r, \ldots, w_n)$, i.e., $w_j = 0$ if $j < r$, and

$$w_r = \frac{6(n-1)\alpha - 2(n-r-1)}{(n-r+1)(n-r+2)},$$

$$w_n = \frac{2(2n-2r+1) - 6(n-1)\alpha)}{(n-r+1)(n-r+2)},$$

and

$$w_j = w_r + \frac{j-r}{n-r}(w_n - w_r),\ r < j < n.$$

The index r depends on the value of α, and is found from the inequalities

$$n-3(n-1)\alpha-1<r\leq n-3(n-1)\alpha.$$

Recently it was established [32] that the solution to the minimum variance OWA weights problem is equivalent to that of minimax disparity [42], i.e., the solution to

$$\begin{aligned}\min \quad & \left\{\max_{i=1,\ldots,n-1}|w_i-w_{i-1}|\right\} \\ \text{s.t.} \quad & \sum_{i=1}^{n} w_i\tfrac{n-i}{n-1}=\alpha, \\ & \sum_{i=1}^{n} w_i=1, w_i\geq 0, i=1,\ldots,n.\end{aligned} \tag{16}$$

The weights of OWA functions obtained as solutions to maximum entropy or minimum variance problems are fixed for any given n and orness value, and can be precomputed. However, both criteria are also useful for data driven weights identification, if there are multiple optimal solutions. Then the solution maximizing $Disp(\mathbf{w})$ or minimizing $D(\mathbf{w})$ is chosen. Torra [39] proposes to solve an auxiliary univariate optimization problem to maximize weights dispersion, subject to a given value of entropy. On the other hand, one can fit the orness value α of MEOWA or MVOWA to empirical data, using a univariate nonlinear optimization method, in which at each iteration the vector $\mathbf{w}$ is computed using analytical solutions to problems (9) and (15).

4.2 Construction of Aggregation Functions from Data

In this section we assume that there is a set of data $\mathscr{D}=\{(\mathbf{x}_k,y_k)\}_{k=1}^{K}$, with $\mathbf{x}_k\in[0,1]^n, y_k\in[0,1]$ for all $k=1,\ldots,K$, and that our goal is to construct an aggregation function (general, or from a given class), that fits the data best. There could however, be variations: a) some components of vectors $\mathbf{x}_k$ may be missing, b) vectors $\mathbf{x}_k$ may have varying dimension by construction, and c) the outputs y_k could be specified as a range of values (i.e., the interval $[\underline{y}_k,\overline{y}_k]$).

The selection of an aggregation function can be stated formally as follows:

Let us have a number of mathematical properties $P_1,P_2,\ldots$ and the data $\mathscr{D}$. Choose an aggregation function f consistent with $P_1,P_2,\ldots$, and satisfying $f(\mathbf{x}_k)\approx y_k, k=1,\ldots,K$.

We can also vary the problem to accommodate a fitting to intervals, i.e. we require $f(\mathbf{x}_k)\in[\underline{y}_k,\overline{y}_k]$. How these values are specified will depend on the application. In some cases it may be possible to fit the function exactly without violating any of the desired properties, however most of the time we merely want to minimize the error of approximation.

Mathematically, the satisfaction of approximate equalities $f(\mathbf{x}_k) \approx y_k$ can be translated into the following minimization problem.

$$\begin{aligned} &\text{minimize } ||\mathbf{r}|| \\ &\text{subject to } f \text{ satisfies } \mathscr{P}_1, \mathscr{P}_2, \ldots, \end{aligned} \tag{17}$$

where $||\mathbf{r}||$ is the norm of the residuals, i.e., $\mathbf{r} \in R^K$ is the vector of the differences between the predicted and observed values $r_k = f(\mathbf{x}_k) - y_k$. There are many ways to choose the norm, and the most popular are the least squares norm, the least absolute deviation norm, the Chebyshev norm or their weighted analogues. Another possibility is to use Huber-type functions which appear in M-estimates to increase robustness with respect to outliers.

Example 4 (Fitting a weighted arithmetic mean)

Suppose that we have a data set $\mathscr{D}$ and we want to define a weighted arithmetic mean using the least squares approach. So we have

$$\begin{aligned} \text{minimize } & \sum_{k=1}^{K} \left(\sum_{i=1}^{n} w_i x_{ik} - y_k \right)^2 \\ \text{subject to } & \sum_{i=1}^{n} w_i = 1, \\ & w_1, \ldots, w_n \geq 0. \end{aligned}$$

This is a quadratic programming problem, which is solved by a number of standard methods.

It may be that the actual numerical value of the output $f(\mathbf{x}_k)$ is not so much important, but rather the ranking of the outputs [29]. That is, if $y_k \leq y_l$, then we want the application to return $f(\mathbf{x}_k) \leq f(\mathbf{x}_l)$. In providing the data, an individual may give more reliable information by providing ranked alternatives, or specifying their preference for one item over another, than by estimating overall scores. In order to preserve the ranking of the outputs, constraints $f(\mathbf{x}_k) \leq f(\mathbf{x}_l)$ if $y_k \leq y_l$ for all pairs k, l can be imposed.

The approximation problem thus far described may turn out to be a general nonlinear optimization problem, or a problem from a special class. Clearly, the complexity of the problem depends on the properties of the chosen norm $||\mathbf{r}||$ and on the properties P_i. Certain optimization problems have well researched solution techniques and proven algorithms, so it is important to recognize the possibility for this where it exists.

Some optimization problems utilize a convex objective function or variant of this, in which case the difficulty is not so much in this step, but rather in defining the constraints. Fitting the Choquet integral, for instance has an exponential number of constraints which need to be defined. Many problems, however can be specified as linear or quadratic programming problems, which have been extensively studied with many solution techniques available.

4.2.1 Learning Weights from Data

We exemplify the process of learning weights of weighted aggregation functions from data using the least absolute deviation (LAD) criterion. This criterion leads to linear programming formulation of regression problems. If we use the least squares criterion, we would obtain corresponding quadratic programming problems.

We remind LAD problem (17) is converted to an LP problem by using the auxiliary variables $r_k^-, r_k^+ \geq 0 : r_k^+ - r_k^- = f(\mathbf{x}_k; w) - y^k$, in which case $r_k^+ + r_k^- = |f(\mathbf{x}_k; w) - y_k|$. Then problem (17) converts to an LP problem [12]

$$\begin{aligned} &\text{minimize} && \sum_{k=1}^{K} r_k^+ + r_k^- \\ &\text{subject to } r_k^+ - r_k^- - f(\mathbf{x}_k; w) = -y_k,\ k = 1, \ldots, K \\ &\qquad \text{other linear constraints on } w \\ &\qquad r_k^-, r_k^+ \geq 0. \end{aligned} \tag{18}$$

Weighted Arithmetic Means
The problem of weights identification (18) takes the special form

$$\begin{aligned} &\text{minimize} && \sum_{k=1}^{K} r_k^+ + r_k^- \\ &\text{subject to } r_k^+ - r_k^- - \sum_{i=1}^{n} w_i x_{ki} = -y_k,\ k = 1, \ldots, K \\ &\qquad \sum_{i=1}^{n} w_i = 1 \\ &\qquad r_k^-, r_k^+ \geq 0, w_i \geq 0. \end{aligned} \tag{19}$$

OWA Functions

The problem of weights identification (18) takes the form

$$\begin{aligned} &\text{minimize} && \sum_{k=1}^{K} r_k^+ + r_k^- \\ &\text{subject to } r_k^+ - r_k^- - \sum_{i=1}^{n} w_i x_{k(i)} = -y_k,\ k = 1, \ldots, K \\ &\qquad \sum_{i=1}^{n} w_i = 1,\ \tfrac{1}{n-1} \sum_{i=1}^{n} w_i (n-i) = \alpha \\ &\qquad r_k^-, r_k^+ \geq 0, w_i \geq 0. \end{aligned} \tag{20}$$

The second constraint involving w is optional, its aim is to obtain a specified orness value α.

Choquet Integrals
Discrete Choquet integrals are conveniently written in terms of the Möbius transform of the corresponding fuzzy measure μ as (see [27, 26]).

$$f(x;w) = \sum_{\mathscr{A}\subseteq\mathscr{N}} w_{\mathscr{A}} \min_{i\in\mathscr{A}} x_i,$$

where w is an array of size $2^n - 1$ encoding the values of the Möbius representation of μ, and $w_{\mathscr{A}}$ is the element corresponding to the subset $\mathscr{A} \subseteq \mathscr{N} = \{1,2,\ldots,n\}$. The problem (18) takes the form (see [10], p.115)

$$\begin{aligned} \text{minimize} \quad & \sum_{k=1}^{K} r_k^+ + r_k^- \\ \text{subject to} \quad & r_k^+ - r_k^- - \sum_{\mathscr{A}\subseteq\mathscr{N}} w_{\mathscr{A}} \min_{i\in\mathscr{A}} x_i = -y_k,\ k = 1,\ldots,K \\ & \sum_{\mathscr{B}\subseteq\mathscr{A}|i\in\mathscr{B}} w_{\mathscr{B}} \geq 0, \text{ for all } \mathscr{A} \subseteq \mathscr{N}, |\mathscr{A}| > 1 \text{ and all } i \in \mathscr{A}, \\ & \sum_{\mathscr{A}\subseteq\mathscr{N}} w_{\mathscr{A}} = 1, \\ & r_k^-, r_k^+ \geq 0, w_{\{i\}} \geq 0, \text{ other } w_{\mathscr{A}} \text{ unrestricted.} \end{aligned} \tag{21}$$

The linear constraints express the conditions of monotonicity fuzzy measures and the bound $\mu_{\mathscr{N}} = 1$. If desired, other conditions (bounds on Shapley values, interaction indices or orness level) can be added as additional linear constraints (see [26, 10]). Furthermore, conditions of k-additivity can also be imposed as linear constraints $w_{\mathscr{A}} = 0$ for all $\mathscr{A}|\ |\mathscr{A}| > k$, or even by eliminating the corresponding variables.

Quasi-Arithmetic Means and Generalized OWA
Consider the situation where the generating function g is fixed (by the decision maker), and the task is to identify w. The problem is solved by linearization [10], by solving the problem

$$\begin{aligned} \text{minimize} \quad & \sum_{k=1}^{K} r_k^+ + r_k^- \\ \text{subject to } & r_k^+ - r_k^- - \sum_{i=1}^{n} w_i g(x_{ki}) = -g(y_k),\ k = 1,\ldots,K \\ & \sum_{i=1}^{n} w_i = 1 \\ & r_k^-, r_k^+ \geq 0, w_i \geq 0, \end{aligned} \tag{22}$$

and similarly for generalized OWA (by replacing x_{ki} with $x_{k(i)}$).

4.2.2 Learning Generators from Data

Quasi-Arithmetic Means
In this case we assume that the weights w are fixed, and the task is to identify the generating function g. We use a linear spline as the generating function, written as a linear combination of some basis functions $B_j, j = 1,\ldots,J$, related to B-splines [3]

$$g(x) = \sum_{j=1}^{J} c_j B_j(x).$$

Since the generating function is defined up to an arbitrary linear transformation, we add the constraints $g(0) = 0, g(1) = 1$ which fix a specific g, if its range is finite. If it is semiinfinite, then its asymptotic behavior is modeled as in [4,9].

The goal is to identify the coefficients c_j, which satisfy $c_j \geq 0$ in order for g to be monotone increasing [3,4]. Then we have

$$\sum_{i=1}^{n} w_i g(x_{ki}) - g(y_k) = \sum_{j=1}^{J} c_j \left(\sum_{i=1}^{n} w_i B_j(x_{ki}) - B_j(y_k) \right).$$

Hence the LAD fitting problem translates into

$$\begin{aligned} &\text{minimize} && \sum_{k=1}^{K} r_k^+ + r_k^- && (23)\\ &\text{subject to } r_k^+ - r_k^- - \sum_{j=1}^{J} c_j \left(\sum_{i=1}^{n} w_i B_j(x_{ki}) - B_j(y_k) \right) = 0, \\ &&& k = 1, \ldots, K, \\ &&& \sum_{j=1}^{J} c_j B_j(0) = 0, \ \sum_{j=1}^{J} c_j B_j(1) = 1, \\ &&& r_k^-, r_k^+ \geq 0, c_j \geq 0, \ w_i \text{ fixed}. \end{aligned}$$

When both w and c are variables, it is possible to set the problem as a bi-level optimization problem, for details see [4,9].

Triangular Norms and Conorms

We consider only t-norms, as equivalent results for t-conorms are obtained by duality. Continuous Archimedean t-norms are dense in the set of all continuous t-norms. They also possess additive generators, monotone decreasing continuous functions $g : [0,1] \to [0,\infty], g(1) = 0$, and can be either strict ($g(0) = \infty$) or nilpotent ($g(0) < \infty$).

By using a very similar method to that we used for quasi-arithmetic means, approximating an additive generator using a linear spline, we have the LAD problem written in the form (in case of nilpotent t-norms)

$$\begin{aligned} &\text{minimize} && \sum_{k=1}^{K} r_k^+ + r_k^- \\ &\text{subject to } r_k^+ - r_k^- - \sum_{j=1}^{J} c_j \left(\sum_{i=1}^{n} B_j(x_{ki}) - B_j(y_k) \right) = 0, \\ &&& k = 1, \ldots, K, \\ &&& \sum_{j=1}^{J} c_j B_j(1) = 0, \ \sum_{j=1}^{J} c_j B_j(1) = 1, \\ &&& r_k^-, r_k^+ \geq 0, c_j \leq 0. \end{aligned} \tag{24}$$

Note that $c_j \leq 0$ because g is monotone decreasing. In case of strict t-norms, the asymptotic behavior is modeled as in [4,9].

Uninorms and Other Generated Functions

Representable uninorms possess additive generators g which satisfy $g(e) = 0$, e being the neutral element, and $g(0) = -\infty$, $g(1) = \infty$. In contrast, continuous generated functions with a neutral element [31,17,10] have generators that are finite at either or both ends of $[0,1]$. They are not associative.

The LAD problem is dealt with very similarly to the case of t-norms, with slightly different constraints on c, namely $c_j \geq 0$ and $\sum_{j=1}^{J} c_j B_j(e) = 0$, plus the constraints resulting from the values $g(0)$ and $g(1)$ (the asymptotes, if needed, are modeled as in [4,9]).

T-S Functions

A T-S function [38] is composed from a t-norm T, a t-conorm S and a weighted quasi-arithmetic mean M_w (with the weighting vector $w = (1-\gamma, \gamma), \gamma \in [0,1]$) by means of

$$f(x; w) = M_w(T(x), S(x)).$$

Special cases are linear convex (M_w is a weighted arithmetic mean) and exponential convex (M_w is a weighted geometric mean) T-S functions. γ-operators [46] are a special case of exponential convex T-S functions with T being the product and S the dual product.

For fixed T, S and M_w and variable parameter γ, the LAD problem is

$$\begin{aligned} &\text{minimize} && \sum_{k=1}^{K} r_k^+ + r_k^- \\ &\text{subject to } r_k^+ - r_k^- - w_1 g(T(x_k)) - w_2 g(S(x_k)) = -g(y_k), \\ &&& k = 1, \ldots, K, \\ &&& w_1 + w_2 = 1, \\ &&& r_k^-, r_k^+ \geq 0, w_1, w_2 \geq 0, \end{aligned} \tag{25}$$

where g is the generating function of M_w and $w = (1-\gamma, \gamma)$.

If the data are interval-valued, or there is a requirement to preserve the ordering of the outputs, there are methods analogous to those mentioned in this section, which also convert the problem to a linear programming problem [7].

4.2.3 Method of Optimal Interpolation

The method of monotone optimal Lipschitz interpolation was proposed in [5] and applied to aggregation functions in [8, 6]. Denote by Mon the set of monotone non-decreasing functions on $[0,1]^n$. Then the set of general Lipschitz n-ary aggregation functions with Lipschitz constant M is characterized as

$$A_{M,||\cdot||} = \{f \in Lip(M, ||\cdot||) \cap Mon : f(\mathbf{0}) = 0, f(\mathbf{1}) = 1\}.$$

We assume that the data set is consistent with the class $A_{M,||\cdot||}$. If not, there are ways of smoothing the data, discussed in [5]. Our goal is to determine the *best* element of $A_{M,||\cdot||}$ which interpolates the data. The *best* is understood in the sense of optimal interpolation [41]: it is the function which minimizes the worst case error, i.e., solves the following Problem.

Optimal Interpolation Problem

$$\min_{f \in A_{M,||\cdot||}} \max_{g \in A_{M,||\cdot||}} \max_{\mathbf{x} \in [0,1]^n} |f(\mathbf{x}) - g(\mathbf{x})|$$

$$\text{s.t. } f(\mathbf{x}_k) = y_k, k = 1, \ldots, K.$$

The solution to this problem will be an aggregation function f which is the "center" of the set of all possible aggregation functions in this class consistent with the data. The method of computing f is based on the following result [5].

Theorem 1. *Let $\mathcal{D}$ be a data set compatible with the conditions $f \in Lip(M, ||\cdot||) \cap Mon$. Then for any $\mathbf{x} \in [0,1]^n$, the values $f(\mathbf{x})$ are bounded by $\sigma_l(\mathbf{x}) \leq f(\mathbf{x}) \leq \sigma_u(\mathbf{x})$, with*

$$\begin{aligned} \sigma_u(\mathbf{x}) &= \min_k \{y_k + M||(\mathbf{x} - \mathbf{x}_k)_+||\}, \\ \sigma_l(\mathbf{x}) &= \max_k \{y_k - M||(\mathbf{x}_k - \mathbf{x})_+||\}, \end{aligned} \tag{26}$$

where $\mathbf{z}_+$ denotes the positive part of vector $\mathbf{z}$: $\mathbf{z}_+ = (\bar{z}_1, \ldots, \bar{z}_n)$, with

$$\bar{z}_i = \max\{z_i, 0\}.$$

The optimal interpolant is given by

$$f(\mathbf{x}) = \frac{1}{2}(\sigma_l(\mathbf{x}) + \sigma_u(\mathbf{x})). \tag{27}$$

Computation of the function f is straightforward, it requires computation of both bounds, and all the functions, σ_l, σ_u and f belong to $Lip(M, ||\cdot||) \cap Mon$. Thus, in

addition to the optimal function f, one obtains as a by-product the strongest and the weakest aggregation functions from the mentioned class.

It is also useful to consider infinite data sets

$$\mathscr{D} = \{(\mathbf{t}, v(\mathbf{t})) : \mathbf{t} \in \Omega \subset [0,1]^n, v : \Omega \to [0,1]]\}$$

in which case the bounds translate into

$$\begin{aligned} B_u(\mathbf{x}) &= \inf_{\mathbf{t}\in\Omega}\{v(\mathbf{t}) + M||(\mathbf{x}-\mathbf{t})_+||\}, \\ B_l(\mathbf{x}) &= \sup_{\mathbf{t}\in\Omega}\{v(\mathbf{t}) - M||(\mathbf{t}-\mathbf{x})_+||\}. \end{aligned} \tag{28}$$

We can make use of these bounds when considering special properties of aggregation functions, such as idempotency or neutral element.

The function f given in Theorem 1 is not yet an aggregation function, because we did not take into account the conditions $f(\mathbf{0}) = 0, f(\mathbf{1}) = 1$. By adding these conditions, we obtain the following generic construction of Lipschitz aggregation functions

$$f(\mathbf{x}) = \frac{1}{2}(\underline{A}(\mathbf{x}) + \overline{A}(\mathbf{x})). \tag{29}$$

$$\underline{A}(\mathbf{x}) = \max\{\sigma_l(\mathbf{x}), B_l(\mathbf{x})\}, \ \overline{A}(\mathbf{x}) = \min\{\sigma_u(\mathbf{x}), B_u(\mathbf{x})\}), \tag{30}$$

where the additional bounds B_l and B_u are due to specific properties of aggregation functions, considered in the next section. At the very least we have (because of $f(\mathbf{0}) = 0, f(\mathbf{1}) = 1$)

$$\begin{aligned} B_u(\mathbf{x}) &= \min\{M||\mathbf{x}||, 1\}, \\ B_l(\mathbf{x}) &= \max\{0, 1 - M||\mathbf{1}-\mathbf{x}||\}, \end{aligned} \tag{31}$$

but other conditions will tighten these bounds.

We note that as a special case of Equations (26)-(31) we obtain p-stable aggregation functions, which have Lipschitz constant $M = 1$ in the norm $||\cdot||_p$. In this case the bounds (31) become Yager t–norm and t–conorm respectively.

4.3 Penalty Based Aggregation

Penalty based aggregation functions have been studied by several authors [45, 18, 34, 28, 2, 25, 11, 33, 35]. The results on the arithmetic means and the median (tt is known that the weighted arithmetic and geometric means, the median and the mode are functions that minimize some simple penalty functions), were already known to Laplace (quoted from [40], p.15), see also [25]. The main motivation is the following. Let $\mathbf{x}$ be the inputs and y the output. If all the inputs coincide $x = x_1 = \ldots = x_n$, then the output is $y = x$, and we have a unanimous vote. If some input $x_i \neq y$, then we impose a "penalty" for this disagreement. The larger the disagreement, and the more inputs disagree with the output, the larger (in general) is the penalty. We

look for an aggregated value which minimizes the penalty; we look for a consensus which minimizes the disagreement.

We start with a very broad definition of penalties, and then particularize it and obtain many known aggregation functions as special cases. Let us consider a vector of inputs $\mathbf{x}$ and the vector $\mathbf{y} = (y, y, \ldots, y)$.

Definition 34. *The function $P : X^{n+1} \to \bar{\Re}_+ = [0, \infty]$ is a penalty function if and only if it satisfies:*

i) $P(\mathbf{x}, y) \geq 0$ *for all* $\mathbf{x}, y$*;*
ii) $P(\mathbf{x}, y) = 0$ *if* $\mathbf{x} = \mathbf{y}$*;*
iii) For every fixed $\mathbf{x}$*, the set of minimizers of* $P(\mathbf{x}, y)$ *is either a singleton or an interval.*

The penalty based function is

$$f(\mathbf{x}) = \arg\min_y P(\mathbf{x}, y),$$

if y is the unique minimizer, and $y = \frac{a+b}{2}$ if the set of minimizers is the interval (a, b) (open or closed).

The first two conditions have useful interpretations: no penalty is imposed if there is full agreement, and no negative penalties are allowed. However, since adding a constant to P does not change its minimizers, technically they can be relaxed: P just needs to reach its absolute minimum when $\mathbf{x} = \mathbf{y}$. Condition iii) ensures that the function f is well defined. If P is quasiconvex in y, then iii) is automatically satisfied. We should also note that a penalty based function is necessarily idempotent, but it is not always monotone.

Definition 35. *A penalty based function f, which is monotone increasing in all components of* $\mathbf{x}$ *is called penalty based aggregation function.*

Next we establish a few general results.

Proposition 1. *Let f be a penalty based aggregation function on X^n, such that y^* is the unique minimizer of $P(\mathbf{x}, y)$. Let h be a continuous strictly monotone function $Y \to X$. Then $f_h(\mathbf{x}) = h^{-1}(f(h(\mathbf{x})))$ is also a penalty based aggregation function on Y^n, with $P_h(\mathbf{x}, y) = P(h(\mathbf{x}), h(y))$.*

Theorem 2. *Let $f : X^n \to X$ be an idempotent function. Then there exists a penalty function $P : X^{n+1} \to \bar{\Re}_+$, such that*

$$f(\mathbf{x}) = \arg\min_y P(\mathbf{x}, y).$$

Proof. The function $P(\mathbf{x}, y) = (f(\mathbf{x}) - y)^2$ is one such penalty function. In fact, any strictly convex (or quasi-convex) univariate function of $t = f(\mathbf{x}) - y$ can serve as such a penalty function.

Corollary 1. *Any averaging aggregation function f can be expressed as a penalty based aggregation function.*

Not every penalty based function is monotone. But for some types of penalty based functions we can establish when monotonicity holds.

A special class of penalty functions was considered in [18]. Let P be given as

$$P(\mathbf{x},y)=\sum_{i=1}^{n} w_i p(x_i,y), \tag{32}$$

where $p: X^2 \to \Re_+$ is a dissimilarity function with the properties

1) $p(t,s)=0$ if and only if $t=s$, and
2) $p(t_1,s) \geq p(t_2,s)$ whenever $t_1 \geq t_2 \geq s$ or $t_1 \leq t_2 \leq s$,

and $\mathbf{w}$ is a weighting vector. Note that the condition 2) is weaker than that in [45], which is $p(t_1,s) \geq p(t_2,s)$ if $|t_1-s| > |t_2-s|$.

The resulting penalty based function, if it exists, is idempotent, but it need not be monotone. To ensure that y^* is unique, and f is an aggregation function, the authors in [18] use the so called "faithful" penalty function.

Definition 36. *The function $p: X^2 \to \Re_+$ is called faithful penalty function, if it satisfies 1) and can be represented as $p(t,s)=K(h(t),h(s))$, where $h: X \to \Re$ is some continuous monotone function (scaling function) and $K: \Re^2 \to \Re_+$ is convex.*

Definition 37. *Let the penalty function P be given by (32), where $p: X^2 \to \Re_+$ is a faithful penalty function. The function*

$$f(\mathbf{x})=y^*=\arg\min_y P(\mathbf{x},y)$$

is a faithful penalty based aggregation function.

A special class of faithful penalty based functions was considered in [33, 35] (dissimilarity functions). The (faithful) penalties p are expressed as

$$p(t,s)=K(h(t)-h(s)), \tag{33}$$

where $K: \Re^2 \to \Re$ is convex (shape function) with the unique minimum $K(0)=0$, and h is the scaling function.

Theorem 3. *[33] The penalty based function with the penalty expressed in (32) and (33) is an aggregation function.*

Example 5. There are several well known aggregation functions that are faithful penalty based aggregation functions.

1. Weighted arithmetic mean with $p(t,s)=(t-s)^2$.
2. Weighted median with $p(t,s)=|t-s|$.
3. Weighted quasi-arithmetic means with the generator h: $p(t,s)=(h(t)-h(s))^2$.

4. Weighted quasi-median with the generator h, $f(\mathbf{x}) = h^{-1}(Med_{\mathbf{w}}(h(\mathbf{x})))$: $p(t,s) = |h(t) - h(s)|$.
5. Let $P(\mathbf{x},y) = \sum_{i=1}^{n} w_i p(x_{(i)},y)$, where $x_{(i)}$ is the i-th largest component of $\mathbf{x}$. We obtain the ordered weighted counterparts of the means in the previous examples, namely the OWA, ordered weighted median and generalized OWA.

Example 6. Let $p(t,s) = 1, t \neq s$ and $p(t,t) = 0$. p is not a faithful penalty function, but it does satisfy conditions 1) and 2). Then the minimizer of (32) is the mode. The mode is not monotone non-decreasing, hence mode is not an aggregation function.

Definition 38 (Deviation mean). *(see [15], p.316) Let $d : X \to \Re$ be a continuous function strictly increasing with respect to the second argument, and satisfying $d(t,t) = 0$ for all $t \in X$. The equation*

$$\sum_{i=1}^{n} w_i d(x_i,y) = 0 \tag{34}$$

has the unique solution $y*$, which is the value of the function $f(\mathbf{x})$ called the deviation mean.

If $d(t,s) = h(s) - h(t)$ for some continuous strictly monotone function h, one recovers the class of weighted quasi-arithmetic means with the generator h.

Theorem 4. *Let the penalty function P be defined as*

$$P(\mathbf{x},y) = \sum_{i=1}^{n} w_i d(x_i,y)^2,$$

where d is a deviation function. Then the penalty based aggregation function is the deviation mean.

Proof: Of course, the equation (34) is the necessary condition of a minimum, which is unique since d^2 is strictly quasiconvex with respect to y.

Hence all deviation means can be represented as penalty based functions but not vice versa (because P needs not be differentiable with respect to y, nor strictly convex in y). They form a subclass of penalty based functions.

Definition 39 (Entropic mean). *[11] Let $\phi : \Re_+ \to \Re$ be a strictly convex differentiable function with $(0,1] \subset dom\ \phi$ and such that $\phi(1) = \phi'(1) = 0$, and $\mathbf{w}$ is a weighting vector. The penalty d_ϕ is defined as*

$$d_\phi(x,y) = x\phi(y/x).$$

The entropic mean is the function

$$f(\mathbf{x}) = y^* = \arg\min_{y \in \Re_+} \sum_{i=1}^{n} w_i d_\phi(x_i,y).$$

It turns out that $d_\phi(\alpha,\cdot)$ is strictly convex for any $\alpha > 0$, and $d_\phi(\alpha,\beta) \geq 0$ with equality if and only if $\alpha = \beta$. The differentiability assumption can be relaxed, see [11]. However entropic means are not a subset of faithful penalty based aggregation functions. On the other hand, all entropic means are homogeneous aggregation functions [11].

Definition 40. *[13] Let $\psi : \Re \to \Re$ be a strictly convex differentiable function. Then the Bregman loss function $D_\psi : \Re \times \Re^n \to \Re$ is defined as*

$$D_\psi(x,y) = \psi(x) - \psi(y) - (x-y)\psi'(y). \tag{35}$$

Definition 41. *Let $\psi : \Re \to \Re$ be a strictly convex differentiable function. Then Bregman penalty based aggregation function is*

$$f(\mathbf{x}) = y^* = \arg\min_{y\in\Re_+} \sum_{i=1}^n w_i D_\psi(y,x_i).$$

Taking the partial derivative of D_ψ [11]:

$$\frac{\partial D_\psi(x,y)}{\partial x} = \psi'(x) - \psi'(y),$$

from which it follows that y^* satisfies

$$\psi'(y^*) = \sum_{i=1}^n w_i \psi'(x_i).$$

Because ψ is strictly convex, the minimum is unique, ψ' is strictly increasing, and $f(\mathbf{x}) = y^*$ is a weighted quasi-arithmetic mean with the generator $h(t) = \psi'(t)$.

Example 7. Penalties based on Bregman loss function [11]:

1. $\psi(t) = (t-1)^2$, then $D_\psi(x,y) = (x-y)^2$ and f is WAM;
2. $\psi(t) = t\log t$, then $D_\psi(x,y) = y\log x/y$, and $f = M_g$ is WGM;
3. $\psi(t) = t\log t - (1+t)\log(1+t)$, then $h(t) = \psi'(t) = \log(\frac{t}{1+t})$, $h^{-1}(t) = \frac{e^t}{1+e^t}$, and
$$f(\mathbf{x}) = \frac{\prod x_i^{w_i}}{\prod(1+x_i)^{w_i} - \prod x_i^{w_i}}.$$

Let us now consider more general penalties, based on the formula

$$P(\mathbf{x},y) = \sum_{i=1}^n p_i(x_i,y). \tag{36}$$

Here, in addition to the weights, we can vary the contribution of the i-th input based on the functional form of the corresponding penalty $p_i(x_i,y)$. This is useful in the

following context. Consider the inputs of different sensors, which need to be averaged (e.g., temperature sensors). The inputs from sensors are random variables with different distributions (e.g., normal, Laplace or another member of exponential family). Then taking the weighted arithmetic mean or median is not appropriate, because sensors are heterogeneous. We can take into account the diversity of inputs errors distributions by means of different penalty functions. The following example presents the penalty suitable when the first distribution is Laplace, and the second is normal.

Example 8. [34,33] Let $n = 2$ and the penalty be

$$P(\mathbf{x}, y) = |x_1 - y| + (x_2 - y)^2.$$

Solving the equation of the necessary condition for a minimum, and taking into account that P is convex, we obtain

$$f(x_1, x_2) = Med(x_1, x_2 - \frac{1}{2}, x_2 + \frac{1}{2}).$$

Extending Example 8 we have

Example 9

$$P(\mathbf{x}, y) = \sum_{i=1}^{n} w_i |x_i - y|^i.$$

We cannot provide a closed form solution in this case, but note that P is convex with respect to y, and a numerical solution is easily obtained using the method of golden section.

Example 10. Let $P(\mathbf{x}, y) = \max(0, y - x_{(1)}) + \max(0, y - x_{(2)}) + \sum_{i=3}^{n} |x_{(i)} - y|$, and $x_{(i)}$ is the i-th largest component of $\mathbf{x}$. The first two terms penalize solutions y exceeding the largest and the second largest inputs. As the result, we discard the two largest values of $\mathbf{x}$. The solution is equivalent to a weighted median with the weighting vector $\mathbf{w} = (0, 0, \frac{1}{n-2}, \ldots, \frac{1}{n-2})$. By changing the absolute value to the squared differences, we obtain an OWA function with the same weighting vector.

Example 11. Let $P(\mathbf{x}, y) = \sum_{i=1}^{n-1} w_i (x_i - y)^2 + w_n \max(0, y - x_n)^2$. The meaning of the last term is the following. Suppose the n-th input (e.g., the n-th expert) usually underestimates the result y. Then we wish to penalize $y > x_n$ but not $y < x_n$. So the n-th input is discarded only if $y < x_n$. The resulting penalty P is a piecewise quadratic function whose minimum is easily found: it is the minimum of the weighted arithmetic means of the first $n - 1$ and of all components of $\mathbf{x}$, $f(\mathbf{x}) = \min(A(x_1, \ldots, x_{n-1}), A(x_1, \ldots, x_n))$.

5 Conclusion

Aggregation functions are widely used in decision and management sciences, expert and decision support systems, recommender systems, internet search engines and many other areas, in which consistent combinations of several inputs into one output value are needed. We discussed various method of construction of aggregation functions based on optimization problems. The construction methods include weights determination based on dispersion or entropy, learning weights from data, learning generating functions, pointwise construction based on optimal interpolation, as well as penalty based aggregation functions. We also looked at how aggregation functions can be used in the context of multiobjective and constrained optimization. We see that there are multiple connections between the areas of optimization and aggregation functions, and there may be interesting new connections worth exploring.

References

1. Alsina, C., Frank, M.J., Schweizer, B.: Associative Functions: Triangular Norms And Copulas. World Scientific, Singapore (2006)
2. Barral Souto, J.: El modo y otras medias, casos particulares de una misma expresion matematica. Boletin Matematico 11, 29–41 (1938)
3. Beliakov, G.: Shape preserving approximation using least squares splines. Approximation Theory and Applications 16, 80–98 (2000)
4. Beliakov, G.: Fitting triangular norms to empirical data. In: Klement, E.P., Mesiar, R. (eds.) Logical, Algebraic, Analytic, and Probabilistic Aspects of Triangular Norms, pp. 255–265. Elsevier, New York (2005)
5. Beliakov, G.: Monotonicity preserving approximation of multivariate scattered data. BIT 45, 653–677 (2005)
6. Beliakov, G.: Construction of aggregation operators for automated decision making via optimal interpolation and global optimization. J. of Industrial and Management Optimization 3, 193–208 (2007)
7. Beliakov, G.: Construction of aggregation functions from data using linear programming. Fuzzy Sets and Systems 160, 65–75 (2009)
8. Beliakov, G., Calvo, T., Lázaro, J.: Pointwise construction of Lipschitz aggregation operators with specific properties. Int. J. of Uncertainty, Fuzziness and Knowledge-Based Systems 15, 193–223 (2007)
9. Beliakov, G., Mesiar, R., Valášková, L.: Fitting generated aggregation operators to empirical data. Int. J. of Uncertainty, Fuzziness and Knowledge-Based Systems 12, 219–236 (2004)
10. Beliakov, G., Pradera, A., Calvo, T.: Aggregation Functions: A Guide for Practitioners. Springer, Heidelberg (2007)
11. Ben-Tal, A., Charnes, A., Teboulle, M.: Entropic means. J. Math. Anal. Appl. 139, 537–551 (1989)
12. Bloomfield, P., Steiger, W.L.: Least Absolute Deviations. Theory, Applications and Algorithms. Birkhauser, Basel (1983)
13. Bregman, L.M.: The relaxation method of finding the common point of convex sets and its application to the solution of problems in convex programming. USSR Comput. Math. Phys. 7, 200–217 (1967)

14. Buchanan, B., Shortliffe, E.: Rule-based Expert Systems. In: The MYCIN Experiments of the Stanford Heuristic Programming Project. Addison-Wesley, Reading (1984)
15. Bullen, P.S.: Handbook of Means and Their Inequalities. Kluwer, Dordrecht (2003)
16. Calvo, T., Kolesárová, A., Komorníková, M., Mesiar, R.: Aggregation operators: properties, classes and construction methods. In: Calvo, T., Mayor, G., Mesiar, R. (eds.) Aggregation Operators. New Trends and Applications, pp. 3–104. Physica-Verlag, Heidelberg (2002)
17. Calvo, T., Mesiar, R.: Continuous generated associative aggregation operators. Fuzzy Sets and Systems 126, 191–197 (2002)
18. Calvo, T., Mesiar, R., Yager, R.: Quantitative weights and aggregation. IEEE Trans. on Fuzzy Systems 12, 62–69 (2004)
19. Carbonell, M., Mas, M., Mayor, G.: On a class of monotonic extended OWA operators. In: 6th IEEE International Conference on Fuzzy Systems, Barcelona, Spain, vol. III, pp. 1695–1700 (1997)
20. Dubois, D., Prade, H.: Fundamentals of Fuzzy Sets. Kluwer, Boston (2000)
21. Dubois, D., Prade, H.: On the use of aggregation operations in information fusion processes. Fuzzy Sets and Systems 142, 143–161 (2004)
22. Duda, R., Hart, P., Nilsson, N.: Subjective Bayesian methods for rule-based inference systems. In: Proc. Nat. Comput. Conf. (AFIPS), vol. 45, pp. 1075–1082 (1976)
23. Fuller, R., Majlender, P.: An analytic approach for obtaining maximal entropy OWA operator weights. Fuzzy Sets and Systems 124, 53–57 (2001)
24. Fuller, R., Majlender, P.: On obtaining minimal variability OWA operator weights. Fuzzy Sets and Systems 136, 203–215 (2003)
25. Gini, C.: Le Medie. Unione Tipografico-Editorial Torinese, Milan (Russian translation, Srednie Velichiny, Statistica, Moscow, 1970) (1958)
26. Grabisch, M., Kojadinovic, I., Meyer, P.: A review of methods for capacity identification in Choquet integral based multi-attribute utility theory. Europ. J. Operations Research 186, 766–785 (2008)
27. Grabisch, M., Murofushi, T., Sugeno, M. (eds.): Fuzzy Measures and Integrals. Theory and Applications. Physica-Verlag, Heidelberg (2000)
28. Jackson, D.: Note on the median of a set of numbers. Bulletin of the Americam Math. Soc. 27, 160–164 (1921)
29. Kaymak, U., van Nauta Lemke, H.R.: Selecting an aggregation operator for fuzzy decision making. In: 3rd IEEE Intl. Conf. on Fuzzy Systems, vol. 2, pp. 1418–1422 (1994)
30. Klement, E.P., Mesiar, R., Pap, E.: Triangular Norms. Kluwer, Dordrecht (2000)
31. Komorníková, M.: Aggregation operators and additive generators. Int. J. of Uncertainty, Fuzziness and Knowledge-Based Systems 9, 205–215 (2001)
32. Liu, X.: The solution equivalence of minimax disparity and minimum variance problems for OWA operators. Int. J. Approx. Reasoning 45, 68–81 (2007)
33. Mesiar, R.: Fuzzy set approach to the utility, preference relations, and aggregation operators. Europ. J. Oper. Res. 176, 414–422 (2007)
34. Mesiar, R., Komornikova, M., Kolesarova, A., Calvo, T.: Aggregation functions: A revision. In: Bustince, H., Herrera, F., Montero, J. (eds.) Fuzzy Sets and Their Extensions: Representation, Aggregation and Models. Springer, Heidelberg (2008)
35. Mesiar, R., Špirková, J., Vavríková, L.: Weighted aggregation operators based on minimization. Inform. Sci. 178, 1133–1140 (2008)
36. Miettinen, K.: Nonlinear Multiobjective Optimization. Springer, New York (1998)
37. O'Hagan, M.O.: Aggregating template or rule antecedents in real-time expert systems with fuzzy set logic. In: 22nd Annual IEEE Asilomar Conf. Signals, Systems, Computers, Pacific Grove, CA, pp. 681–689 (1988)

38. Pradera, A., Trillas, E., Calvo, T.: A general class of triangular norm-based aggregation operators: quasi-linear T-S operators. Intl. J. of Approximate Reasoning 30, 57–72 (2002)
39. Torra, V.: Learning weights for the quasi-weighted means. IEEE Trans. on Fuzzy Systems 10, 653–666 (2002)
40. Torra, Y., Narukawa, V.: Modeling Decisions. Information Fusion and Aggregation Operators. Springer, Heidelberg (2007)
41. Traub, J.F., Wozniakowski, H.: A General Theory of Optimal Algorithms. Academic Press, New York (1980)
42. Wang, Y.M., Parkan, C.: A minimax disparity approach for obtaining OWA operator weights. Inform. Sci. 175, 20–29 (2005)
43. Yager, R.: On ordered weighted averaging aggregation operators in multicriteria decision making. IEEE Trans. on Systems, Man and Cybernetics 18, 183–190 (1988)
44. Yager, R., Rybalov, A.: Uninorm aggregation operators. Fuzzy Sets and Systems 80, 111–120 (1996)
45. Yager, R., Rybalov, A.: Understanding the median as a fusion operator. Int. J. General Syst. 26, 239–263 (1997)
46. Zimmermann, H.-J., Zysno, P.: Latent connectives in human decision making. Fuzzy Sets and Systems 4, 37–51 (1980)

Chebyshev Approximation of Inconsistent Fuzzy Relational Equations with Max-T Composition*

Pingke Li and Shu-Cherng Fang

Abstract. This paper considers resolving the inconsistency of a system of fuzzy relational equations with max-T composition by simultaneously modifying the coefficient matrix and the right hand side vector. We show that resolving the inconsistency of fuzzy relational equations with max-T composition by means of Chebyshev approximation is closely related to the generalized solvability of interval-valued fuzzy relational equations with max-T composition. An efficient procedure is proposed to obtain a consistent system with the smallest perturbation in the sense of Chebyshev distance.

Keywords: Fuzzy optimization, fuzzy relational equations, Chebyshev approximation.

1 Introduction

A system of fuzzy relational equations with max-T composition is of the form

$$\max_{j\in N} T(a_{ij}, x_j) = b_i, \quad \forall\ i \in M, \tag{1}$$

where $M = \{1, 2, \cdots, m\}$ and $N = \{1, 2, \cdots, n\}$ are two index sets, $A = (a_{ij})_{m\times n} \in [0,1]^{mn}$, $\boldsymbol{b} = (b_1, b_2, \cdots, b_m)^T \in [0,1]^m$, $\boldsymbol{x} = (x_1, x_2, \cdots, x_n)^T \in [0,1]^n$ and $T : [0,1]^2 \to [0,1]$ is a triangular norm (t-norm for short). A

Pingke Li · Shu-Cherng Fang
Edward P. Fitts Department of Industrial and Systems Engineering,
North Carolina State University, Raleigh, NC 27695-7906, USA
e-mail: pli@ncsu.edu,fang@eos.ncsu.edu

Shu-Cherng Fang
Department of Mathematical Sciences and Department of Industrial Engineering,
Tsinghua University, Beijing 100084, China

* The work is supported by US NSF Grant #DMI-0553310.

W.A. Lodwick & J. Kacprzyk (Eds.): Fuzzy Optimization, STUDFUZZ 254, pp. 109–124.
springerlink.com

system of the above form in (1) is also called a system of max-T equations for short and denoted as $A \circ_T \boldsymbol{x} = \boldsymbol{b}$ in the matrix form where "$\circ_T$" stands for the max-T composition. Typically, the t-norm T involved in a system $A \circ_T \boldsymbol{x} = \boldsymbol{b}$ is required to be continuous, i.e., continuous as a function of two arguments.

The resolution of a system of max-T equations is to determine the unknown vector $\boldsymbol{x}$ for a given coefficient matrix A and a right hand side vector $\boldsymbol{b}$ such that $A \circ_T \boldsymbol{x} = \boldsymbol{b}$. The set of all solutions to a system of max-T equations $A \circ_T \boldsymbol{x} = \boldsymbol{b}$ is denoted by $S(A, \boldsymbol{b})$, i.e., $S(A, \boldsymbol{b}) = \{\boldsymbol{x} \in [0,1]^n \mid A \circ_T \boldsymbol{x} = \boldsymbol{b}\}$. A system $A \circ_T \boldsymbol{x} = \boldsymbol{b}$ is called consistent if $S(A, \boldsymbol{b}) \neq \emptyset$, otherwise, it is inconsistent.

Fuzzy relational equations were first investigated by Sanchez [26, 27] under the max-T_M composition where T_M is the *minimum* operator, i.e., $T_M(x, y) = \min(x, y)$. Since then, solving various types of fuzzy relational equations has become one of the most appealing issues in fuzzy set theory. It has been pointed out that fuzzy relational equations play an important role as a uniform platform in many applications of fuzzy sets and fuzzy systems. See, e.g., Pedrycz [22, 24], Mordeson and Malik [18] and Peeva and Kyosev [25].

The resolution of a system of max-T equations has been investigated by Pedrycz [20, 21], Miyakoshi and Shimbo [17], Di Nola *et al.* [6, 7, 8], Klir and Yuan [12], De Baets [5], etc. It is well known that the consistency of a system of max-T equations can be verified in polynomial time by constructing and checking a potential maximum solution. The set of all solutions, when it is nonempty, is a finitely generated root system which can be fully determined by a unique maximum solution and a finite number of minimal solutions. However, the detection of all minimal solutions is an NP-hard problem. Similar conclusions can be drawn for a system of max-T inequalities. The reader may refer to Li and Fang [14, 15] and references therein for more details.

Although the consistency of a system of max-T equations can be readily verified and its solution set can be well characterized, related investigations are meaningful only when the system under consideration is consistent. However, due to the inaccuracy and deficiency in data or the inappropriate choice of the t-norm, it happens quite often that the system of max-T equations obtained in modeling a real situation turns out to be inconsistent. Moreover, the consistency of a system of max-T equations could be very sensitive to the data, i.e., small perturbations in the data could lead a consistent system to become inconsistent.

To deal with the impreciseness of the data and resolve the inconsistency of the system, one possible approach is to consider the interval-valued max-T equations, i.e., each entry in the matrix A and the vector $\boldsymbol{b}$ is replaced by a closed interval of possible values in $[0, 1]$. A system of interval-valued max-T equations can be represented in the form $\tilde{A} \circ_T \boldsymbol{x} = \tilde{\boldsymbol{b}}$ where $\tilde{A} = (\tilde{a}_{ij})_{m \times n}$ is an interval-valued matrix with $\tilde{a}_{ij} = [\underline{a}_{ij}, \overline{a}_{ij}] \subseteq [0, 1]$ and $\tilde{\boldsymbol{b}} = (\tilde{b}_i)_{m \times 1}$ is

an interval-valued vector with $\tilde{b}_i = [\underline{b}_i, \overline{b}_i] \subseteq [0,1]$. Denote $\underline{A} = (\underline{a}_{ij})_{m\times n}$, $\overline{A} = (\overline{a}_{ij})_{m\times n}$ and similarly, $\underline{\boldsymbol{b}} = (\underline{b}_i)_{m\times 1}$, $\overline{\boldsymbol{b}} = (\overline{b}_i)_{m\times 1}$. By extending the natural order in a componentwise manner, $\tilde{A}$ and $\tilde{\boldsymbol{b}}$ induce the sets $[\underline{A}, \overline{A}] \triangleq \{A \mid \underline{A} \leq A \leq \overline{A}\}$ and $[\underline{\boldsymbol{b}}, \overline{\boldsymbol{b}}] \triangleq \{\boldsymbol{b} \mid \underline{\boldsymbol{b}} \leq \boldsymbol{b} \leq \overline{\boldsymbol{b}}\}$, respectively. The matrices $\underline{A}$ and $\overline{A}$ are referred to as the lower and upper bounds of $\tilde{A}$, respectively, and similarly, the vectors $\underline{\boldsymbol{b}}$ and $\overline{\boldsymbol{b}}$ the lower and upper bounds of $\tilde{\boldsymbol{b}}$, respectively. A system of interval-valued max-T equations $\tilde{A} \circ_T \boldsymbol{x} = \tilde{\boldsymbol{b}}$ is understood as the family of all systems of max-T equations $A \circ_T \boldsymbol{x} = \boldsymbol{b}$ with $A \in [\underline{A}, \overline{A}]$ and $\boldsymbol{b} \in [\underline{\boldsymbol{b}}, \overline{\boldsymbol{b}}]$. Without loss of generality, we may always assume that $\underline{A} \leq \overline{A}$ and $\underline{\boldsymbol{b}} \leq \overline{\boldsymbol{b}}$ such that the system $\tilde{A} \circ \boldsymbol{x} = \tilde{\boldsymbol{b}}$ is properly defined. Interval-valued max-T equations have been investigated by Wagenknecht and Hartmann [28, 29], Wang and Chang [30], Li and Fang [13], Wang *et al.* [31] and Li and Fang [16].

Another approach to resolving the inconsistency of a system of max-T equations is to perturb as slightly as possible either the coefficient matrix, the right hand side vector or both to reach a consistent system. Based on the notion of "minimal distortions", Pedrycz [23] proposed a procedure to modify the right hand side vector of an inconsistent system of max-T_M equations. However, as indicated by Cuninghame-Green and Cechlárová [3], the procedure is not given in a precise algorithmic form and hence would be difficult to implement in a computer. Cuninghame-Green and Cechlárová [3] presented an algorithm to obtain a consistent system of max-T_M equations with the smallest perturbation of the right hand side vector in the sense of Chebyshev distance, whereas Cechlárová [2] proposed an analogous algorithm to resolve the inconsistency of a system of max-T_M equations by modifying the coefficient matrix. Both algorithms are essentially based on the idempotency property of T_M, i.e., $T_M(x,x) = \min(x,x) = x$, $\forall\ x \in [0,1]$. Hence, neither of them can be generalized for general max-T equations since T_M has been proved to be the unique t-norm which possesses the idempotency property. Moreover, no algorithm is known that resolves the inconsistency of a system of max-T_M equations by simultaneously modifying the coefficient matrix and the right hand side vector. To the best of our knowledge, resolving the inconsistency of a system of max-T equations by means of Chebyshev approximation remains to be an open problem.

In this paper, we show that resolving the inconsistency of a system of max-T equations by means of Chebyshev approximation is closely related to the generalized solvability of interval-valued max-T equations. A bisection method is proposed for an inconsistent system of max-T equations to obtain the smallest perturbation bound of both the coefficient matrix and the right hand side vector. The construction of a Chebyshev approximation is introduced thereafter and illustrated by numerical examples. The proposed procedure remains valid with necessary modifications if only the coefficient matrix or the right hand side vector, but not both, can be perturbed.

2 Preliminaries

In this section, we recall some basic concepts and results associated with fuzzy relational equations, which are indispensable for the introduction of the Chebyshev approximation approach in this context. All proofs in this section are omitted to make the paper succinct and readable. The reader may refer to the monograph of Klement *et al.* [11] for a comprehensive discussion on triangular norms, and Li and Fang [15, 16] and reference therein for a detailed discussion on max-T equations and interval-valued max-T equations.

2.1 Triangular Norms

Although originally introduced in the framework of probabilistic metric spaces, t-norms have been proposed as natural generalizations of the logical conjunction in fuzzy logic and played an important role in the construction of fuzzy systems which may be described by fuzzy relational equations.

Definition 2.1. *A t-norm is a binary operator* $T : [0,1]^2 \to [0,1]$ *such that for all* $x, y, z \in [0,1]$ *the following four axioms are satisfied:*

(T1) $T(x,y) = T(y,x)$. *(commutativity)*
(T2) $T(x, T(y,z)) = T(T(x,y), z)$. *(associativity)*
(T3) $T(x,y) \le T(x,z)$, *whenever* $y \le z$. *(monotonicity)*
(T4) $T(x,1) = x$. *(boundary condition)*

A t-norm is said to be continuous if it is continuous as a function of two arguments. Due to its commutativity and monotonicity properties, a t-norm is continuous if and only if it is continuous in one of its arguments. Analogously, a t-norm is said to be left- or right-continuous if it is left- or right-continuous, respectively, in one of its arguments.

The most frequently used continuous t-norm is the *minimum* operator $T_M(x,y) = \min(x,y)$. Other important continuous t-norms include the *product* operator $T_P(x,y) = xy$ and the *bounded difference* operator $T_L(x,y) = \max(x + y - 1, 0)$, a.k.a., Łukasiewicz t-norm. Note that T_M is the largest t-norm while T_P and T_L are prototypical examples of two important classes of continuous t-norms, i.e., strict t-norms and nilpotent t-norms, respectively.

Definition 2.2. *Let* $T : [0,1]^2 \to [0,1]$ *be a left-continuous t-norm. The associated residual implicator is a binary operator* $I_T : [0,1]^2 \to [0,1]$ *such that*

$$I_T(x,y) = \sup\{z \in [0,1] \mid T(x,z) \le y\}, \quad \forall\ (x,y) \in [0,1]^2. \tag{2}$$

Residual implicators are also known as φ-operators which were introduced by Pedrycz [19, 20] in a different approach. The connection between a φ-operator and its corresponding t-norm has been investigated in full generality by Gottwald [9, 10], Miyakoshi and Shimbo [17] and Di Nola *et al.* [8].

The residual implicators with respect to the *minimum* operator T_M, the *product* operator T_P and the Łukasiewicz t-norm T_L are, respectively,

$$I_{T_M}(x,y) = \begin{cases} 1, & \text{if } x \le y \\ y, & \text{otherwise,} \end{cases} \qquad \text{(Gödel implicator)}$$

$$I_{T_P}(x,y) = \begin{cases} 1, & \text{if } x \le y \\ y/x, & \text{otherwise,} \end{cases} \qquad \text{(Goguen implicator)}$$

$$I_{T_L}(x,y) = \min(1-x+y, 1). \qquad \text{(Łukasiewicz implicator)}$$

Lemma 2.3. *The residual implicator I_T with respect to a left-continuous t-norm T is left-continuous and decreasing in its first argument as well as right-continuous and increasing in its second argument.*

Lemma 2.4. *Let T be a left-continuous t-norm and I_T its residual implicator. The inequality $T(a, I_T(a,b)) \le b$ holds for all $a, b \in [0,1]$. Moreover, $T(a,x) \le b$ if and only if $x \le I_T(a,b)$.*

Lemma 2.4 plays a crucial role in the resolution of max-T equations, which is actually a special scenario of the general theory of Galois connections [1].

2.2 Fuzzy Relational Equations

Let $A \circ_T \boldsymbol{x} = \boldsymbol{b}$ be a system of max-T equations with T being a continuous t-norm. Due to the monotonicity of the t-norm T, we have $A \circ_T \boldsymbol{x}^1 \le A \circ_T \boldsymbol{x}^2$ whenever $\boldsymbol{x}^1 \le \boldsymbol{x}^2$. Hence, $\boldsymbol{x} \in S(A, \boldsymbol{b})$ if $\boldsymbol{x}^1,\ \boldsymbol{x}^2 \in S(A, \boldsymbol{b})$ and $\boldsymbol{x}^1 \le \boldsymbol{x} \le \boldsymbol{x}^2$. Therefore, we may focus on the extremal solutions as defined below.

Definition 2.5. *A solution $\check{\boldsymbol{x}} \in S(A, \boldsymbol{b})$ is called a minimal solution if for any $\boldsymbol{x} \in S(A, \boldsymbol{b})$, the relation $\boldsymbol{x} \le \check{\boldsymbol{x}}$ implies $\boldsymbol{x} = \check{\boldsymbol{x}}$. A solution $\hat{\boldsymbol{x}} \in S(A, \boldsymbol{b})$ is called a maximum solution if $\boldsymbol{x} \le \hat{\boldsymbol{x}}$, $\forall\ \boldsymbol{x} \in S(A, \boldsymbol{b})$.*

Lemma 2.6 *Let T be a left-continuous t-norm and I_T its residual implicator. For any $A \in [0,1]^{mn}$ and $\boldsymbol{b} \in [0,1]^m$, it holds that $A \circ_T (A^T \circ_\varphi \boldsymbol{b}) \le \boldsymbol{b}$ where "$\circ_\varphi$" stands for the min-I_T composition and $A^T \circ_\varphi \boldsymbol{b} \in [0,1]^n$ is the vector with its components being defined by*

$$(A^T \circ_\varphi \boldsymbol{b})_j = \min\{I_T(a_{ij}, b_i) \mid i \in M\}, \quad \forall\ j \in N. \tag{3}$$

Moreover, $A \circ_T \boldsymbol{x} \le \boldsymbol{b}$ if and only if $\boldsymbol{x} \le A^T \circ_\varphi \boldsymbol{b}$.

Theorem 2.7 *Let $A \circ_T \boldsymbol{x} = \boldsymbol{b}$ be a system of max-T equations with T being a left-continuous t-norm. The system is consistent if and only if $A^T \circ_\varphi \boldsymbol{b}$ is a solution to $A \circ_T \boldsymbol{x} = \boldsymbol{b}$. Moreover, if T is also right-continuous and hence*

continuous, the solution set $S(A, \boldsymbol{b})$, when it is nonempty, can be fully determined by one maximum solution and a finite number of minimal solutions, i.e.,

$$S(A, \boldsymbol{b}) = \bigcup_{\check{\boldsymbol{x}} \in \check{S}(A, \boldsymbol{b})} \left\{ \boldsymbol{x} \in [0,1]^n \mid \check{\boldsymbol{x}} \leq \boldsymbol{x} \leq \hat{\boldsymbol{x}} \right\}, \tag{4}$$

where $\check{S}(A, \boldsymbol{b})$ is the set of all minimal solutions of $A \circ_T \boldsymbol{x} = \boldsymbol{b}$ and $\hat{\boldsymbol{x}} = A^T \circ_\varphi \boldsymbol{b}$ is the maximum solution.

Lemma 2.6 is a direct result of Lemma 2.4. The solvability criteria of max-T equations were investigated by Sanchez [26], Pedrycz [20, 21] and Miyakoshi and Shimbo [17] while the structure of the solution set was characterized by Sanchez [27] and Di Nola *et al.* [6, 7, 8]. The particular structure of $S(A, \boldsymbol{b})$ is called a finitely generated root system by De Baets [4, 5]. Note that the intersection of two finitely generated root systems, when it is nonempty, remains to be a finitely generated root system.

According to Theorem 2.7, the consistency of a system of max-T equations $A \circ_T \boldsymbol{x} = \boldsymbol{b}$ can be verified by constructing and checking the potential maximum solution $\hat{\boldsymbol{x}} = A^T \circ_\varphi \boldsymbol{b}$ in a time complexity of $O(mn)$. However, the detection of all minimal solutions is a complicated and challenging issue for investigation. The reader may refer to Li and Fang [15] and references therein for more detailed discussion.

2.3 Interval-Valued Fuzzy Relational Equations

Let $\tilde{A}$ be an interval-valued matrix with the lower bound $\underline{A} \in [0,1]^{mn}$ and the upper bound $\overline{A} \in [0,1]^{mn}$, and $\tilde{\boldsymbol{b}}$ an interval-valued vector with the lower bound $\underline{\boldsymbol{b}} \in [0,1]^m$ and the upper bound $\overline{\boldsymbol{b}} \in [0,1]^m$. We now consider a system of interval-valued max-T equations $\tilde{A} \circ_T \boldsymbol{x} = \tilde{\boldsymbol{b}}$ with T being a continuous t-norm. The following two lemmas are crucial in dealing with the system $\tilde{A} \circ_T \boldsymbol{x} = \tilde{\boldsymbol{b}}$, both of which simply rely on the monotonicity and continuity properties of the t-norm T.

Lemma 2.8. *Let $\tilde{A}$ be an interval-valued matrix and T a continuous t-norm. Given a vector $\boldsymbol{x} \in [0,1]^n$, then for each vector $\boldsymbol{b} \in [\underline{A} \circ_T \boldsymbol{x}, \overline{A} \circ_T \boldsymbol{x}]$ there exists $A \in [\underline{A}, \overline{A}]$ such that $A \circ_T \boldsymbol{x} = \boldsymbol{b}$.*

Lemma 2.9. *Let $\tilde{A}$ be an interval-valued matrix and T a continuous t-norm. For any vector $\boldsymbol{x} \in [0,1]^n$, we have*

$$\{A \circ_T \boldsymbol{x} \mid A \in [\underline{A}, \overline{A}]\} = [\underline{A} \circ_T \boldsymbol{x}, \overline{A} \circ_T \boldsymbol{x}]. \tag{5}$$

Definition 2.10. *Let $\tilde{A} \circ_T \boldsymbol{x} = \tilde{\boldsymbol{b}}$ be a system of interval-valued max-T equations with T being a continuous t-norm. A vector $\boldsymbol{x} \in [0,1]^n$ is called a united solution of $\tilde{A} \circ_T \boldsymbol{x} = \tilde{\boldsymbol{b}}$ if there exist $A \in [\underline{A}, \overline{A}]$ and $\boldsymbol{b} \in [\underline{\boldsymbol{b}}, \overline{\boldsymbol{b}}]$ such that $A \circ_T \boldsymbol{x} = \boldsymbol{b}$.*

Denote $S_u(\tilde{A}, \tilde{\boldsymbol{b}})$ the set of united solutions for a system of interval-valued max-T equations $\tilde{A} \circ_T \boldsymbol{x} = \tilde{\boldsymbol{b}}$. By Lemma 2.9, we immediately have

$$S_u(\tilde{A}, \tilde{\boldsymbol{b}}) = \{\boldsymbol{x} \in [0,1]^n \mid [\underline{A} \circ_T \boldsymbol{x}, \overline{A} \circ_T \boldsymbol{x}] \cap [\underline{\boldsymbol{b}}, \overline{\boldsymbol{b}}] \neq \emptyset\} \tag{6}$$

$$= \{\boldsymbol{x} \in [0,1]^n \mid \underline{A} \circ_T \boldsymbol{x} \leq \overline{\boldsymbol{b}},\ \overline{A} \circ_T \boldsymbol{x} \geq \underline{\boldsymbol{b}}\} \tag{7}$$

and hence, by Lemma 2.6, the following straightforward result.

Theorem 2.11. *Let $\tilde{A} \circ_T \boldsymbol{x} = \tilde{\boldsymbol{b}}$ be a system of interval-valued max-T equations with T being a continuous t-norm. The set of united solutions $S_u(\tilde{A}, \tilde{\boldsymbol{b}}) \neq \emptyset$ if and only if $\overline{A} \circ_T (\underline{A}^T \circ_\varphi \overline{\boldsymbol{b}}) \geq \underline{\boldsymbol{b}}$.*

Definition 2.12. *Let $\tilde{A} \circ_T \boldsymbol{x} = \tilde{\boldsymbol{b}}$ be a system of interval-valued max-T equations with T being a continuous t-norm. A vector $\boldsymbol{x} \in [0,1]^n$ is called a tolerable solution of $\tilde{A} \circ_T \boldsymbol{x} = \tilde{\boldsymbol{b}}$ if for each $A \in [\underline{A}, \overline{A}]$ there exists $\boldsymbol{b} \in [\underline{\boldsymbol{b}}, \overline{\boldsymbol{b}}]$ such that $A \circ_T \boldsymbol{x} = \boldsymbol{b}$. Similarly, a vector $\boldsymbol{x} \in [0,1]^n$ is called a controllable solution of $\tilde{A} \circ_T \boldsymbol{x} = \tilde{\boldsymbol{b}}$ if for each $\boldsymbol{b} \in [\underline{\boldsymbol{b}}, \overline{\boldsymbol{b}}]$ there exists $A \in [\underline{A}, \overline{A}]$ such that $A \circ_T \boldsymbol{x} = \boldsymbol{b}$.*

Denote $S_t(\tilde{A}, \tilde{\boldsymbol{b}})$ and $S_c(\tilde{A}, \tilde{\boldsymbol{b}})$ the sets of tolerable solutions and controllable solutions, respectively, for a system of interval-valued max-T equations $\tilde{A} \circ_T \boldsymbol{x} = \tilde{\boldsymbol{b}}$. By Lemma 2.9, we have

$$S_t(\tilde{A}, \tilde{\boldsymbol{b}}) = \{\boldsymbol{x} \in [0,1]^n \mid [\underline{A} \circ_T \boldsymbol{x}, \overline{A} \circ_T \boldsymbol{x}] \subseteq [\underline{\boldsymbol{b}}, \overline{\boldsymbol{b}}]\} \tag{8}$$

$$= \{\boldsymbol{x} \in [0,1]^n \mid \overline{A} \circ_T \boldsymbol{x} \leq \overline{\boldsymbol{b}},\ \underline{A} \circ_T \boldsymbol{x} \geq \underline{\boldsymbol{b}}\} \tag{9}$$

and

$$S_c(\tilde{A}, \tilde{\boldsymbol{b}}) = \{\boldsymbol{x} \in [0,1]^n \mid [\underline{A} \circ_T \boldsymbol{x}, \overline{A} \circ_T \boldsymbol{x}] \supseteq [\underline{\boldsymbol{b}}, \overline{\boldsymbol{b}}]\} \tag{10}$$

$$= \{\boldsymbol{x} \in [0,1]^n \mid \underline{A} \circ_T \boldsymbol{x} \leq \underline{\boldsymbol{b}},\ \overline{A} \circ_T \boldsymbol{x} \geq \overline{\boldsymbol{b}}\}. \tag{11}$$

Theorem 2.13. *Let $\tilde{A} \circ_T \boldsymbol{x} = \tilde{\boldsymbol{b}}$ be a system of interval-valued max-T equations with T being a continuous t-norm. The set of tolerable solutions $S_t(\tilde{A}, \tilde{\boldsymbol{b}}) \neq \emptyset$ if and only if $\underline{A} \circ_T (\overline{A}^T \circ_\varphi \overline{\boldsymbol{b}}) \geq \underline{\boldsymbol{b}}$ while the set of controllable solutions $S_c(\tilde{A}, \tilde{\boldsymbol{b}}) \neq \emptyset$ if and only if $\overline{A} \circ_T (\underline{A}^T \circ_\varphi \underline{\boldsymbol{b}}) \geq \overline{\boldsymbol{b}}$.*

It is clear that $S_t(\tilde{A}, \tilde{\boldsymbol{b}}) \subseteq S_u(\tilde{A}, \tilde{\boldsymbol{b}})$ and $S_c(\tilde{A}, \tilde{\boldsymbol{b}}) \subseteq S_u(\tilde{A}, \tilde{\boldsymbol{b}})$. Moreover, $S_t(\tilde{A}, \tilde{\boldsymbol{b}}) \cap S_c(\tilde{A}, \tilde{\boldsymbol{b}}) = \{\boldsymbol{x} \in [0,1]^n \mid \underline{A} \circ_T \boldsymbol{x} = \underline{\boldsymbol{b}},\ \overline{A} \circ_T \boldsymbol{x} = \overline{\boldsymbol{b}}\}$.

By Theorems 2.11 and 2.13, the existence of a united solution can be verified in a time complexity of $O(mn)$ as well as the existence of a tolerable solution and controllable solution, respectively. Furthermore, as will be shown in Section 3, the notion of united solutions bridges the gap between an inconsistent system of max-T equations and a system of interval-valued max-T equations. The notions of tolerable solutions and controllable solutions are the key to the construction of a Chebyshev approximation of an inconsistent system of max-T equations.

3 The Chebyshev Approximation

In this section, we consider an inconsistent system of max-T equations $A \circ_T \boldsymbol{x} = \boldsymbol{b}$ with T being a continuous t-norm and resolve its inconsistency by means of Chebyshev approximation. Without loss of generality, we may always assume that the system $A \circ_T \boldsymbol{x} = \boldsymbol{b}$ is in its normal form, i.e., the equations are arranged in the way such that $b_1 \geq b_2 \geq \cdots \geq b_m \geq 0$. Note that the equations corresponding to the index set $M_0 = \{i \in M \mid b_i = 0\}$ should be taken into consideration whenever $S(A, \boldsymbol{b}) = \emptyset$ while they can be discarded with necessary modifications on the remaining equations in case of consistency.

For notational convenience, the infix notations "$\wedge$" and "$\vee$" are used to denote the *minimum* and *maximum* operators, respectively, i.e., $x \wedge y = \min(x, y)$ and $x \vee y = \max(x, y)$. Analogously, we denote $A^1 \wedge A^2 = (a^1_{ij} \wedge a^2_{ij})_{m\times n}$ and $A^1 \vee A^2 = (a^1_{ij} \vee a^2_{ij})_{m\times n}$ for any $A^1, A^2 \in [0,1]^{mn}$, and $\boldsymbol{b}^1 \wedge \boldsymbol{b}^2 = (b^1_i \wedge b^2_i)_{m\times 1}$ and $\boldsymbol{b}^1 \vee \boldsymbol{b}^2 = (b^1_i \vee b^2_i)_{m\times 1}$ for any $\boldsymbol{b}^1, \boldsymbol{b}^2 \in [0,1]^m$.

Denote $\mathscr{C}$ the set of all pairs of a coefficient matrix $A' \in [0,1]^{mn}$ and a right hand side vector $\boldsymbol{b}' \in [0,1]^m$ such that the corresponding system of max-T equations $A' \circ_T \boldsymbol{x} = \boldsymbol{b}'$ is consistent, i.e., $\mathscr{C} = \{(A', \boldsymbol{b}') \mid S(A', \boldsymbol{b}') \neq \emptyset\}$. It is clear that $(A, \boldsymbol{b}) \notin \mathscr{C}$ for the inconsistent system $A \circ_T \boldsymbol{x} = \boldsymbol{b}$. The Chebyshev distance between the pair $(A, \boldsymbol{b})$ and a pair $(A', \boldsymbol{b}') \in \mathscr{C}$ is defined as

$$\rho((A, \boldsymbol{b}), (A', \boldsymbol{b}')) = \max\Big(\max_{i,j} |a_{ij} - a'_{ij}|,\ \max_i |b_i - b'_i|\Big). \tag{12}$$

Definition 3.1. *A system of max-T equations $A' \circ_T \boldsymbol{x} = \boldsymbol{b}'$ is said to be a δ-approximation of the system $A \circ_T \boldsymbol{x} = \boldsymbol{b}$ if $(A', \boldsymbol{b}') \in \mathscr{C}$ and $\rho((A, \boldsymbol{b}), (A', \boldsymbol{b}')) \leq \delta$.*

Definition 3.2. *A system of max-T equations $A^\dagger \circ_T \boldsymbol{x} = \boldsymbol{b}^\dagger$ is said to be a Chebyshev approximation of the system $A \circ_T \boldsymbol{x} = \boldsymbol{b}$ if $(A^\dagger, \boldsymbol{b}^\dagger) \in \mathscr{C}$ and*

$$\rho((A, \boldsymbol{b}), (A^\dagger, \boldsymbol{b}^\dagger)) = \inf_{(A', \boldsymbol{b}') \in \mathscr{C}} \rho((A, \boldsymbol{b}), (A', \boldsymbol{b}')). \tag{13}$$

Clearly, it suffices to consider δ-approximations with $\delta \in [0,1]$ for the system $A \circ_T \boldsymbol{x} = \boldsymbol{b}$. A 1-approximation of the system $A \circ_T \boldsymbol{x} = \boldsymbol{b}$ always exists, while a Chebyshev approximation is a δ-approximation with the smallest possible value of δ. For a given $\delta \in [0,1]$, denote $\tilde{A}(\delta)$ the interval-valued matrix with the lower bound $\underline{A}(\delta)$ and the upper bound $\overline{A}(\delta)$ where

$$\underline{A}(\delta) = ((a_{ij} - \delta) \vee 0)_{m\times n} \quad \text{and} \quad \overline{A}(\delta) = ((a_{ij} + \delta) \wedge 1)_{m\times n}, \tag{14}$$

respectively. Similarly, denote $\tilde{\boldsymbol{b}}(\delta)$ the interval-valued vector with the lower bound $\underline{\boldsymbol{b}}(\delta)$ and the upper bound $\overline{\boldsymbol{b}}(\delta)$ where

$$\underline{\boldsymbol{b}}(\delta) = ((b_i - \delta) \vee 0)_{m\times 1} \quad \text{and} \quad \overline{\boldsymbol{b}}(\delta) = ((b_i + \delta) \wedge 1)_{m\times 1}, \tag{15}$$

respectively. Consequently, we obtain a properly defined system of interval-valued max-T equations $\tilde{A}(\delta) \circ_T \boldsymbol{x} = \tilde{\boldsymbol{b}}(\delta)$. It is clear that a matrix $A' \in [\underline{A}(\delta), \overline{A}(\delta)]$ if and only if $A' = (a'_{ij})_{m\times n} \in [0,1]^{mn}$ and $\max_{i,j} |a_{ij} - a'_{ij}| \leq \delta$. Similarly, a vector $\boldsymbol{b}' \in [\underline{\boldsymbol{b}}(\delta), \overline{\boldsymbol{b}}(\delta)]$ if and only if $\boldsymbol{b}' = (b'_i)_{m\times 1} \in [0,1]^m$ and $\max_i |b_i - b'_i| \leq \delta$.

Theorem 3.3 *The system of max-T equations $A \circ_T \boldsymbol{x} = \boldsymbol{b}$ has a δ-approximation if and only if the system of interval-valued max-T equations $\tilde{A}(\delta) \circ_T \boldsymbol{x} = \tilde{\boldsymbol{b}}(\delta)$ has a united solution.*

Proof: If the system $A \circ_T \boldsymbol{x} = \boldsymbol{b}$ has a δ-approximation $A' \circ_T \boldsymbol{x} = \boldsymbol{b}'$ such that $S(A', \boldsymbol{b}') \neq \emptyset$, it is clear that $\max_{i,j} |a_{ij} - a'_{ij}| \leq \delta$ and $\max_i |b_i - b'_i| \leq \delta$, respectively. Hence, $A' \in [\underline{A}(\delta), \overline{A}(\delta)]$, $\boldsymbol{b}' \in [\underline{\boldsymbol{b}}(\delta), \overline{\boldsymbol{b}}(\delta)]$ and consequently, $S_u(\tilde{A}(\delta), \tilde{\boldsymbol{b}}(\delta)) \neq \emptyset$.

Conversely, if $S_u(\tilde{A}(\delta), \tilde{\boldsymbol{b}}(\delta)) \neq \emptyset$, there exist $A' \in [\underline{A}(\delta), \overline{A}(\delta)]$ and $\boldsymbol{b}' \in [\underline{\boldsymbol{b}}(\delta), \overline{\boldsymbol{b}}(\delta)]$ such that the system $A' \circ_T \boldsymbol{x} = \boldsymbol{b}'$ is consistent. It is clear that $\rho((A, \boldsymbol{b}), (A', \boldsymbol{b}')) \leq \delta$ and hence $A' \circ_T \boldsymbol{x} = \boldsymbol{b}'$ is a δ-approximation of the system $A \circ_T \boldsymbol{x} = \boldsymbol{b}$. □

Theorem 3.4 *If the system of interval-valued max-T equations $\tilde{A}(\delta) \circ_T \boldsymbol{x} = \tilde{\boldsymbol{b}}(\delta)$ has a united solution for some $\delta \in [0,1]$, then the system $\tilde{A}(\delta') \circ_T \boldsymbol{x} = \tilde{\boldsymbol{b}}(\delta')$ has a united solution for any $\delta' \in [\delta, 1]$.*

Proof: It is straightforward from the observation of $[\underline{A}(\delta), \overline{A}(\delta)] \subseteq [\underline{A}(\delta'), \overline{A}(\delta')]$ and $[\underline{\boldsymbol{b}}(\delta), \overline{\boldsymbol{b}}(\delta)] \subseteq [\underline{\boldsymbol{b}}(\delta'), \overline{\boldsymbol{b}}(\delta')]$ for $\delta \leq \delta'$. □

Theorem 3.5 *If the system of interval-valued max-T equations $\tilde{A}(\delta) \circ_T \boldsymbol{x} = \tilde{\boldsymbol{b}}(\delta)$ has a united solution for all $\delta \in (\delta', 1]$, then the system $\tilde{A}(\delta') \circ_T \boldsymbol{x} = \tilde{\boldsymbol{b}}(\delta')$ also has a united solution.*

Proof: By Theorem 2.11, the system $\tilde{A}(\delta) \circ_T \boldsymbol{x} = \tilde{\boldsymbol{b}}(\delta)$ has a united solution if and only if $\overline{A}(\delta) \circ_T (\underline{A}^T(\delta) \circ_\varphi \overline{\boldsymbol{b}}(\delta)) \geq \underline{\boldsymbol{b}}(\delta)$. Notice that each component of the vector $\overline{A}(\delta) \circ_T (\underline{A}^T(\delta) \circ_\varphi \overline{\boldsymbol{b}}(\delta)) - \underline{\boldsymbol{b}}(\delta)$ is right-continuous with respect to δ since all involved operations are continuous except that the residual implicator I_T is left-continuous in its first argument and right-continuous in its second argument. Hence, the system $\tilde{A}(\delta') \circ_T \boldsymbol{x} = \tilde{\boldsymbol{b}}(\delta')$ has a united solution as long as the system $\tilde{A}(\delta) \circ_T \boldsymbol{x} = \tilde{\boldsymbol{b}}(\delta)$ has a united solution for all $\delta \in (\delta', 1]$. □

Theorem 3.3 indicates that the existence of a δ-approximation of a system of max-T equations is equivalent to the existence of a united solution of a corresponding system of interval-valued max-T equations which, by Theorem 2.11, can be verified in polynomial time. Theorems 3.4 and 3.5 guarantee that a Chebyshev approximation does exist for a system of max-T equations and also suggest a bisection method to obtain the smallest perturbation bound δ^* with the existence of a δ^*-approximation.

Algorithm

Step 1. Specify ϵ as the required level of precision.

Step 2. Set $\underline{\delta} = 0$, $\overline{\delta} = 1$ and $\delta = (\underline{\delta} + \overline{\delta})/2$.

Step 3. Verify $\overline{A}(\delta) \circ_T (\underline{A}^T(\delta) \circ_\varphi \overline{\boldsymbol{b}}(\delta)) \geq \underline{\boldsymbol{b}}(\delta)$. If it holds, set $\overline{\delta} = \delta$, otherwise, set $\underline{\delta} = \delta$. Set $\delta = (\underline{\delta} + \overline{\delta})/2$.

Step 4. If $\overline{\delta} - \underline{\delta} \leq \epsilon$, output $\delta^* = \overline{\delta}$ and stop. Otherwise, go to Step 3.

The above algorithm offers an ϵ-optimal value for δ^* in a time complexity of $O(mn \log(1/\epsilon))$ where ϵ is the predetermined level of precision. Once we obtain the value of δ^*, the remaining problem is to construct a Chebyshev approximation, i.e., a δ^*-approximation for the system $A \circ_T \boldsymbol{x} = \boldsymbol{b}$.

Denote $\boldsymbol{x}(\delta^*) = \underline{A}^T(\delta^*) \circ_\varphi \overline{\boldsymbol{b}}(\delta^*)$. By Theorem 2.11, $\boldsymbol{x}(\delta^*)$ is a united solution of the system $\tilde{A}(\delta^*) \circ_T \boldsymbol{x} = \tilde{\boldsymbol{b}}(\delta^*)$ and hence $[\underline{A}(\delta^*) \circ_T \boldsymbol{x}(\delta^*), \overline{A}(\delta^*) \circ_T \boldsymbol{x}(\delta^*)] \cap [\underline{\boldsymbol{b}}(\delta^*), \overline{\boldsymbol{b}}(\delta^*)] \neq \emptyset$. Denote $\tilde{\boldsymbol{b}}^\dagger(\delta^*)$ the interval-valued vector with the lower bound $\underline{\boldsymbol{b}}^\dagger(\delta^*)$ and the upper bound $\overline{\boldsymbol{b}}^\dagger(\delta^*)$ where

$$\underline{\boldsymbol{b}}^\dagger(\delta^*) = \underline{A}(\delta^*) \circ_T \boldsymbol{x}(\delta^*) \vee \underline{\boldsymbol{b}}(\delta^*) \text{ and } \overline{\boldsymbol{b}}^\dagger(\delta^*) = \overline{A}(\delta^*) \circ_T \boldsymbol{x}(\delta^*) \wedge \overline{\boldsymbol{b}}(\delta^*), \quad (16)$$

respectively. Notice that $[\underline{\boldsymbol{b}}^\dagger(\delta^*), \overline{\boldsymbol{b}}^\dagger(\delta^*)] \subseteq [\underline{A}(\delta^*) \circ_T \boldsymbol{x}(\delta^*), \overline{A}(\delta^*) \circ_T \boldsymbol{x}(\delta^*)]$ and hence $\boldsymbol{x}(\delta^*)$ is also a controllable solution of the system $\tilde{A}(\delta^*) \circ_T \boldsymbol{x} = \tilde{\boldsymbol{b}}^\dagger(\delta^*)$ which means that for each $\boldsymbol{b}' \in [\underline{\boldsymbol{b}}^\dagger(\delta^*), \overline{\boldsymbol{b}}^\dagger(\delta^*)]$ there exists $A' \in [\underline{A}(\delta^*), \overline{A}(\delta^*)]$ such that $A' \circ_T \boldsymbol{x}(\delta^*) = \boldsymbol{b}'$.

Therefore, denote $\tilde{A}^\dagger(\delta^*)$ the interval-valued matrix with the lower bound $\underline{A}^\dagger(\delta^*)$ and the upper bound $\overline{A}^\dagger(\delta^*)$ where

$$\underline{A}^\dagger(\delta^*) = (\boldsymbol{x}(\delta^*) \circ_\varphi (\underline{\boldsymbol{b}}^\dagger(\delta^*))^T)^T \wedge \overline{A}(\delta^*) \text{ and } \overline{A}^\dagger(\delta^*) = (\boldsymbol{x}(\delta^*) \circ_\varphi (\overline{\boldsymbol{b}}^\dagger(\delta^*))^T)^T \wedge \overline{A}(\delta^*) \quad (17)$$

respectively. Since $\boldsymbol{x}(\delta^*)$ is a controllable solution of the system $\tilde{A}(\delta^*) \circ_T \boldsymbol{x} = \tilde{\boldsymbol{b}}^\dagger(\delta^*)$, by Lemma 2.6 and Theorem 2.7, we have $\underline{A}^\dagger(\delta^*) \circ_T \boldsymbol{x}(\delta^*) = \underline{\boldsymbol{b}}^\dagger(\delta^*)$ and $\overline{A}^\dagger(\delta^*) \circ_T \boldsymbol{x}(\delta^*) = \overline{\boldsymbol{b}}^\dagger(\delta^*)$, respectively. Therefore, $\boldsymbol{x}(\delta^*)$ is simultaneously a tolerable and controllable solution of the system $\tilde{A}^\dagger(\delta^*) \circ_T \boldsymbol{x} = \tilde{\boldsymbol{b}}^\dagger(\delta^*)$. Moreover, since $[\underline{A}^\dagger(\delta^*), \overline{A}^\dagger(\delta^*)] \subseteq [\underline{A}(\delta^*), \overline{A}(\delta^*)]$ and $[\underline{\boldsymbol{b}}^\dagger(\delta^*), \overline{\boldsymbol{b}}^\dagger(\delta^*)] \subseteq [\underline{\boldsymbol{b}}(\delta^*), \overline{\boldsymbol{b}}(\delta^*)]$, any pair of $A^\dagger \in [\underline{A}^\dagger(\delta^*), \overline{A}^\dagger(\delta^*)]$ and $\boldsymbol{b}^\dagger = A^\dagger \circ_T \boldsymbol{x}(\delta^*) \in [\underline{\boldsymbol{b}}^\dagger(\delta^*), \overline{\boldsymbol{b}}^\dagger(\delta^*)]$ defines a Chebyshev approximation $A^\dagger \circ_T \boldsymbol{x} = \boldsymbol{b}^\dagger$ of the system $A \circ_T \boldsymbol{x} = \boldsymbol{b}$.

The proposed procedure remains valid with necessary modifications if we are allowed to perturb the coefficient matrix only or the right hand side vector only. In the algorithm for determining the smallest perturbation bound, we can simply keep the matrix A unchanged, i.e., $\underline{A}(\delta) = \overline{A}(\delta) = A$ for any $\delta \in [0, 1]$, if only $\boldsymbol{b}$ can be modified. Once the smallest perturbation bound δ_1^* is obtained for $\boldsymbol{b}$, the system $A \circ_T \boldsymbol{x} = \boldsymbol{b}^\dagger$ with $\boldsymbol{b}^\dagger = A \circ_T (A^T \circ_\varphi \overline{\boldsymbol{b}}(\delta_1^*))$ is a Chebyshev approximation of the inconsistent system $A \circ_T \boldsymbol{x} = \boldsymbol{b}$. The situation is analogous if only A can be modified. We can keep the vector $\boldsymbol{b}$ unchanged,

i.e., $\underline{\boldsymbol{b}}(\delta) = \overline{\boldsymbol{b}}(\delta) = \boldsymbol{b}$ for any $\delta \in [0,1]$, to obtain the smallest perturbation bound δ_2^* for A. Thereafter, the system $A^\dagger \circ_T \boldsymbol{x} = \boldsymbol{b}$ with $A^\dagger = ((\underline{A}^T(\delta_2^*) \circ_\varphi \boldsymbol{b}) \circ_\varphi \boldsymbol{b}^T)^T \wedge \overline{A}(\delta_2^*)$ is a Chebyshev approximation of the inconsistent system $A \circ_T \boldsymbol{x} = \boldsymbol{b}$. Besides, it is obvious that $\delta^* \leq \min(\delta_1^*, \delta_2^*)$ for an inconsistent system of max-T equations. Deeper relations among these perturbation bounds are subject to further investigation.

4 Numerical Examples

In this section, we provide a few numerical examples to illustrate the proposed procedure and compare with the known results in Cuninghame-Green and Cechlárová [3] and Cechlárová [2].

Example 1. Consider the system of max-T_M equations $A \circ_{T_M} \boldsymbol{x} = \boldsymbol{b}$ with

$$A = \begin{pmatrix} 0.7 & 0.5 & 0.3 & 0.5 \\ 1 & 0.4 & 0.5 & 0.7 \\ 0.2 & 1 & 1 & 0.6 \\ 0.4 & 0.5 & 0.5 & 0.8 \end{pmatrix}, \qquad \boldsymbol{b} = \begin{pmatrix} 1 \\ 0.4 \\ 0.2 \\ 0 \end{pmatrix}.$$

This example was originally presented by Pedrycz [23] and also investigated by Cuninghame-Green and Cechlárová [3].

The system is inconsistent since the potential maximum solution $\hat{\boldsymbol{x}} = (0,0,0,0)^T$ is clearly not a solution. By perturbing A and $\boldsymbol{b}$ simultaneously, our algorithm obtains $\delta^* = 0.3$ and hence

$$\tilde{A}(\delta^*) = \begin{pmatrix} [0.4,1] & [0.2,0.8] & [0,0.6] & [0.2,0.8] \\ [0.7,1] & [0.1,0.7] & [0.2,0.8] & [0.4,1] \\ [0,0.5] & [0.7,1] & [0.7,1] & [0.3,0.9] \\ [0.1,0.7] & [0.2,0.8] & [0.2,0.8] & [0.5,1] \end{pmatrix}, \qquad \boldsymbol{b}(\delta^*) = \begin{pmatrix} [0.7,1] \\ [0.1,0.7] \\ [0,0.5] \\ [0,0.3] \end{pmatrix}.$$

Moreover, we have $\boldsymbol{x}(\delta^*) = (1, 0.5, 0.5, 0.3)^T$. Consequently,

$$\tilde{\boldsymbol{b}}^\dagger(\delta^*) = \begin{pmatrix} [0.4,1] \\ [0.7,1] \\ [0.5,0.5] \\ [0.3,0.7] \end{pmatrix} \bigcap \begin{pmatrix} [0.7,1] \\ [0.1,0.7] \\ [0,0.5] \\ [0,0.3] \end{pmatrix} = \begin{pmatrix} [0.7,1] \\ 0.7 \\ 0.5 \\ 0.3 \end{pmatrix},$$

$$\underline{A}^{\dagger}(\delta^*) = \left(\begin{pmatrix} 1 \\ 0.5 \\ 0.5 \\ 0.3 \end{pmatrix} \circ_{\varphi} (0.7, 0.7, 0.5, 0.3) \right)^T \bigwedge \begin{pmatrix} 1 & 0.8 & 0.6 & 0.8 \\ 1 & 0.7 & 0.8 & 1 \\ 0.5 & 1 & 1 & 0.9 \\ 0.7 & 0.8 & 0.8 & 1 \end{pmatrix} = \begin{pmatrix} 0.7 & 0.8 & 0.6 & 0.8 \\ 0.7 & 0.7 & 0.8 & 1 \\ 0.5 & 1 & 1 & 0.9 \\ 0.3 & 0.3 & 0.3 & 1 \end{pmatrix}$$

and

$$\overline{A}^{\dagger}(\delta^*) = \left(\begin{pmatrix} 1 \\ 0.5 \\ 0.5 \\ 0.3 \end{pmatrix} \circ_{\varphi} (1, 0.7, 0.5, 0.3) \right)^T \bigwedge \begin{pmatrix} 1 & 0.8 & 0.6 & 0.8 \\ 1 & 0.7 & 0.8 & 1 \\ 0.5 & 1 & 1 & 0.9 \\ 0.7 & 0.8 & 0.8 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0.8 & 0.6 & 0.8 \\ 0.7 & 0.7 & 0.8 & 1 \\ 0.5 & 1 & 1 & 0.9 \\ 0.3 & 0.3 & 0.3 & 1 \end{pmatrix}.$$

Hence, any $A^{\dagger} \in [\underline{A}^{\dagger}(\delta^*), \overline{A}^{\dagger}(\delta^*)]$ and $\boldsymbol{b}^{\dagger} = A^{\dagger} \circ_{T_M} \boldsymbol{x}(\delta^*)$ define a Chebyshev approximation $A^{\dagger} \circ_{T_M} \boldsymbol{x} = \boldsymbol{b}^{\dagger}$, for instance,

$$A^{\dagger} = \begin{pmatrix} 0.8 & 0.8 & 0.6 & 0.8 \\ 0.7 & 0.7 & 0.8 & 1 \\ 0.5 & 1 & 1 & 0.9 \\ 0.3 & 0.3 & 0.3 & 1 \end{pmatrix}, \qquad \boldsymbol{b}^{\dagger} = \begin{pmatrix} 0.8 \\ 0.7 \\ 0.5 \\ 0.3 \end{pmatrix}.$$

If we want to resolve the inconsistency by modifying the right hand side vector only, the corresponding smallest perturbation bound becomes $\delta_1^* = 0.4$. Hence we have

$$\tilde{\boldsymbol{b}}(\delta_1^*) = \begin{pmatrix} [0.6, 1] \\ [0, 0.8] \\ [0, 0.6] \\ [0, 0.4] \end{pmatrix}, \qquad \boldsymbol{b}^{\dagger} = A \circ_{T_M} (A^T \circ_{\varphi} \overline{\boldsymbol{b}}(\delta_1^*)) = \begin{pmatrix} 0.7 \\ 0.8 \\ 0.4 \\ 0.4 \end{pmatrix}.$$

Consequently, the system $A \circ_{T_M} \boldsymbol{x} = \boldsymbol{b}^{\dagger}$ is a Chebyshev approximation of $A \circ_{T_M} \boldsymbol{x} = \boldsymbol{b}$ in this case, which is exactly the same as that given by Cuninghame-Green and Cechlárová [3]. Note that $A \circ_{T_M} \boldsymbol{x} = \boldsymbol{b}^{\dagger}$ has a maximum solution $(0.8, 0.4, 0.4, 0.4)^T$.

On the other hand, the smallest perturbation bound for A becomes $\delta_2^* = 0.6$, if we keep $\boldsymbol{b}$ unchanged. Hence we have

$$\tilde{A}(\delta_2^*) = \begin{pmatrix} [0.1, 1] & [0, 1] & [0, 0.9] & [0, 1] \\ [0.4, 1] & [0, 1] & [0, 1] & [0.1, 1] \\ [0, 0.8] & [0.4, 1] & [0.4, 1] & [0, 1] \\ [0, 1] & [0, 1] & [0, 1] & [0.2, 1] \end{pmatrix},$$

$$A^{\dagger} = ((\underline{A}^T(\delta_2^*) \circ_{\varphi} \boldsymbol{b}) \circ_{\varphi} \boldsymbol{b}^T)^T \wedge \overline{A}(\delta_2^*) = \begin{pmatrix} 1 & 1 & 0.9 & 1 \\ 0.4 & 1 & 1 & 1 \\ 0.2 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Consequently, the system $A^{\dagger} \circ_{T_M} \boldsymbol{x} = \boldsymbol{b}$ is a Chebyshev approximation of $A \circ_{T_M} \boldsymbol{x} = \boldsymbol{b}$ in this case. Note that $A^{\dagger} \circ_{T_M} \boldsymbol{x} = \boldsymbol{b}$ has a maximum solution $(1, 0.2, 0.2, 0)^T$.

Example 2. Consider the system of max-T_M equations $A \circ_{T_M} \boldsymbol{x} = \boldsymbol{b}$ with

$$A = \begin{pmatrix} 0.6 & 0.2 & 0.9 & 0.1 & 0.6 \\ 0.5 & 0.7 & 0.3 & 0.8 & 0.7 \\ 0.3 & 0.6 & 0.7 & 0.4 & 0.2 \\ 0.3 & 0.8 & 0.5 & 0.4 & 0.2 \end{pmatrix}, \qquad \boldsymbol{b} = \begin{pmatrix} 0.9 \\ 0.5 \\ 0.4 \\ 0.3 \end{pmatrix}.$$

This example was originally presented by Cechlárová [2].

The system is inconsistent since the potential maximum solution $\hat{\boldsymbol{x}} = (1, 0.3, 0.3, 0.3, 0.5)^T$ is not a solution. The smallest perturbation bound for A is $\delta_2^* = 0.3$, if we are required to resolve the inconsistency by modifying A only. Hence we have

$$\tilde{A}(\delta_2^*) = \begin{pmatrix} [0.3, 0.9] & [0, 0.5] & [0.6, 1] & [0, 0.4] & [0.3, 0.9]] \\ [0.2, 0.8] & [0.4, 1] & [0, 0.6] & [0.5, 1] & [0.4, 1] \\ [0, 0.6] & [0.5, 1] & [0.2, 0.8] & [0.1, 0.7] & [0, 0.5] \\ [0, 0.6] & [0.3, 0.9] & [0.4, 1] & [0.1, 0.7] & [0, 0.5] \end{pmatrix},$$

$$A^{\dagger} = ((\underline{A}^T(\delta_2^*) \circ_{\varphi} \boldsymbol{b}) \circ_{\varphi} \boldsymbol{b}^T)^T \wedge \overline{A}(\delta_2^*) = \begin{pmatrix} 0.9 & 0.5 & 0.9 & 0.4 & 0.9 \\ 0.5 & 1 & 0.5 & 0.5 & 0.5 \\ 0.4 & 0.9 & 0.4 & 0.4 & 0.4 \\ 0.3 & 1 & 0.3 & 0.3 & 0.3 \end{pmatrix}.$$

Consequently, the system $A^{\dagger} \circ_{T_M} \boldsymbol{x} = \boldsymbol{b}$ is a Chebyshev approximation of $A \circ_{T_M} \boldsymbol{x} = \boldsymbol{b}$ in this case. Note that $A^{\dagger} \circ_{T_M} \boldsymbol{x} = \boldsymbol{b}$ has a maximum solution $(1, 0.3, 1, 1, 1)^T$. Since a different method is used to construct a Chebyshev approximation, the matrix $A^{\dagger}$ offered by our procedure is slightly different from that given by Cechlárová [2], but both matrices share the same Chebyshev distance.

Now we present an example to illustrate that the proposed procedure works for general max-T equations with T being a continuous t-norm.

Example 3. Consider the system of max-T_L equations $A \circ_{T_L} \boldsymbol{x} = \boldsymbol{b}$ with

$$A = \begin{pmatrix} 0.2 & 0.9 & 0.8 & 0.4 \\ 0.8 & 0.3 & 0.4 & 0.8 \\ 0.5 & 0.7 & 0.1 & 0.6 \end{pmatrix}, \qquad \boldsymbol{b} = \begin{pmatrix} 0.8 \\ 0.6 \\ 0.2 \end{pmatrix}.$$

The potential maximum solution is $\hat{\boldsymbol{x}} = (0.7, 0.5, 1, 0.6)^T$ and

$$\begin{pmatrix} 0.2 & 0.9 & 0.8 & 0.4 \\ 0.8 & 0.3 & 0.4 & 0.8 \\ 0.5 & 0.7 & 0.1 & 0.6 \end{pmatrix} \circ_{T_L} \begin{pmatrix} 0.7 \\ 0.5 \\ 0.1 \\ 0.6 \end{pmatrix} = \begin{pmatrix} 0.8 \\ 0.5 \\ 0.2 \end{pmatrix} \neq \begin{pmatrix} 0.8 \\ 0.6 \\ 0.2 \end{pmatrix}.$$

Therefore, the system is inconsistent. By perturbing A and $\boldsymbol{b}$ simultaneously, our algorithm obtains $\boldsymbol{x}(\delta^*) = (0.7500, 0.5500, 1.0000, 0.6500)^T$ with $\delta^* = 0.0250$, and consequently,

$$\underline{\boldsymbol{b}}^{\dagger}(\delta^*) = \begin{pmatrix} 0.7750 \\ 0.5750 \\ 0.2250 \end{pmatrix}, \quad \overline{\boldsymbol{b}}^{\dagger}(\delta^*) = \begin{pmatrix} 0.8250 \\ 0.5750 \\ 0.2250 \end{pmatrix}$$

and

$$\underline{A}^{\dagger}(\delta^*) = \begin{pmatrix} 0.2250 & 0.9250 & 0.7750 & 0.4250 \\ 0.8250 & 0.3250 & 0.4250 & 0.8250 \\ 0.4750 & 0.6750 & 0.1250 & 0.5750 \end{pmatrix}, \ \overline{A}^{\dagger}(\delta^*) = \begin{pmatrix} 0.2250 & 0.9250 & 0.8250 & 0.4250 \\ 0.8250 & 0.3250 & 0.4250 & 0.8250 \\ 0.4750 & 0.6750 & 0.1250 & 0.5750 \end{pmatrix}.$$

Hence, any $A^{\dagger} \in [\underline{A}^{\dagger}(\delta^*), \overline{A}^{\dagger}(\delta^*)]$ and $\boldsymbol{b}^{\dagger} = A^{\dagger} \circ_{T_L} \boldsymbol{x}(\delta^*)$ define a Chebyshev approximation $A^{\dagger} \circ_{T_L} \boldsymbol{x} = \boldsymbol{b}^{\dagger}$, for instance,

$$A^{\dagger} = \begin{pmatrix} 0.2250 & 0.9250 & 0.8000 & 0.4250 \\ 0.8250 & 0.3250 & 0.4250 & 0.8250 \\ 0.4750 & 0.6750 & 0.1250 & 0.5750 \end{pmatrix}, \qquad \boldsymbol{b}^{\dagger} = \begin{pmatrix} 0.8000 \\ 0.5750 \\ 0.2250 \end{pmatrix}.$$

5 Concluding Remarks

We have shown that the existence of a δ-approximation of a system of max-T equations is equivalent to the existence of a united solution of a corresponding system of interval-valued max-T equations. Consequently, the smallest perturbation bound can be obtained by repeatedly constructing a system of interval-valued max-T equations and verifying its solvability condition. As illustrated by our numerical examples, a Chebyshev approximation can be constructed readily once the smallest perturbation bound is obtained. It

is clear that the Chebyshev approximation may not necessarily be unique. In this case, we may be interested in obtaining a Chebyshev approximation of some special quality, for instance, the one with the smallest number of modifications in the coefficient matrix and the right hand side vector. This new challenge goes beyond the scope of this paper and subject to further investigation.

References

1. Blyth, T.S., Janowitz, M.F.: Residuation Theory. Pergamon Press, Oxford (1972)
2. Cechlárová, K.: A note on unsolvable systems of max-min (fuzzy) equations. Linear Algebra and its Applications 310, 123–128 (2000)
3. Cuninghame-Green, R.A., Cechlárová, K.: Residuation in fuzzy algebra and some applications. Fuzzy Sets and Systems 71, 227–239 (1995)
4. De Baets, B.: Oplossen van vaagrelationele vergelijkingen: een ordetheoretische benadering. Ph.D. Dissertation, University of Gent (1995)
5. De Baets, B.: Analytical solution methods for fuzzy relational equations. In: Dubois, D., Prade, H. (eds.) Fundamentals of Fuzzy Sets. The Handbooks of Fuzzy Sets Series, vol. 1, pp. 291–340. Kluwer, Dordrecht (2000)
6. Di Nola, A., Pedrycz, W., Sessa, S.: On solution of fuzzy relational equations and their characterization. BUSEFAL 12, 60–71 (1982)
7. Di Nola, A., Pedrycz, W., Sessa, S., Wang, P.Z.: Fuzzy relation equations under triangular norms: A survey and new results. Stochastica 8, 99–145 (1984)
8. Di Nola, A., Sessa, S., Pedrycz, W., Sanchez, E.: Fuzzy Relation Equations and Their Applications to Knowledge Engineering. Kluwer, Dordrecht (1989)
9. Gottwald, S.: Characterizations of the solvability of fuzzy equations. Elektron. Informationsverarb. Kybernet. 22, 67–91 (1986)
10. Gottwald, S.: Fuzzy Sets and Fuzzy Logic: The Foundations of Application - from a Mathematical Point of View, Vieweg, Wiesbaden (1993)
11. Klement, E.P., Mesiar, R., Pap, E.: Triangular Norms. Kluwer, Dordrecht (2000)
12. Klir, G., Yuan, B.: Fuzzy Sets and Fuzzy Logic: Theory and Applications. Prentice Hall, Upper Saddle River (1995)
13. Li, G., Fang, S.-C.: Solving interval-valued fuzzy relation equations. IEEE Transactions on Fuzzy Systems 6, 321–324 (1998)
14. Li, P., Fang, S.-C.: A survey on fuzzy relational equations, Part I: Classification and solvability. Fuzzy Optimization and Decision Making 8 (2009) doi:10.1007/s10700-009-9059-0
15. Li, P., Fang, S.-C.: On the resolution and optimization of a system of fuzzy relational equations with sup-T composition. Fuzzy Optimization and Decision Making 7, 169–214 (2008)
16. Li, P., Fang, S.-C.: A note on solution sets of interval-valued fuzzy relational equations. Fuzzy Optimization and Decision Making 8 (2009) doi:10.1007/s10700-009-9055-4
17. Miyakoshi, M., Shimbo, M.: Solutions of composite fuzzy relational equations with triangular norms. Fuzzy Sets and Systems 16, 53–63 (1985)
18. Mordeson, J.N., Malik, D.S.: Fuzzy Automata and Languages: Theory and Applications. Chapman & Hall/CRC, Boca Raton (2002)

19. Pedrycz, W.: Some aspects of fuzzy decision making. Kybernetes 11, 297–301 (1982)
20. Pedrycz, W.: Fuzzy relational equations with triangular norms and their resolutions. BUSEFAL 11, 24–32 (1982)
21. Pedrycz, W.: On generalized fuzzy relational equations and their applications. Journal of Mathematical Analysis and Applications 107, 520–536 (1985)
22. Pedrycz, W.: Fuzzy Control and Fuzzy Systems. Research Studies Press/Wiely, New York (1989)
23. Pedrycz, W.: Inverse problem in fuzzy relational equations. Fuzzy Sets and Systems 36, 277–291 (1990)
24. Pedrycz, W.: Processing in relational structures: Fuzzy relational equations. Fuzzy Sets and Systems 40, 77–106 (1991)
25. Peeva, K., Kyosev, Y.: Fuzzy Relational Calculus: Theory, Applications and Software. World Scientific, New Jersey (2004)
26. Sanchez, E.: Resolution of composite fuzzy relation equation. Information and Control 30, 38–48 (1976)
27. Sanchez, E.: Solutions in composite fuzzy relation equations: application to medical diagnosis in Brouwerian logic. In: Gupta, M.M., Saridis, G.N., Gaines, B.R. (eds.) Fuzzy Automata and Decision Processes, pp. 221–234. North-Holland, Amsterdam (1977)
28. Wagenknecht, M., Hartmann, K.: Fuzzy modelling with tolerances. Fuzzy Sets and Systems 20, 325–332 (1986)
29. Wagenknecht, M., Hartmann, K.: On direct and inverse problems for fuzzy equation systems with tolerances. Fuzzy Sets and Systems 24, 93–102 (1987)
30. Wang, H.-F., Chang, Y.-C.: Resolution of composite interval-valued fuzzy relation equations. Fuzzy Sets and Systems 44, 227–240 (1991)
31. Wang, S., Fang, S.-C., Nuttle, H.L.W.: Solution sets of interval-valued fuzzy relational equations. Fuzzy Optimization and Decision Making 2, 41–60 (2003)

Part III
Various Types of Fuzzy Optimization and Fuzzy Mathematical Programming Models

A Survey of Fuzzy Convex Programming Models

Ricardo C. Silva, Carlos Cruz, José L. Verdegay, and Akebo Yamakami

Abstract. Optimization is a procedure of finding and comparing feasible solutions until no better solution can be found. It can be divided into several fields, one of which is the Convex Optimization. It is characterized by a convex objective function and convex constraint functions over a convex set which is the set of the decision variables. This can be viewed, on the one hand, as a particular case of nonlinear programming and, on the other hand, as a general case of linear programming. Convex optimization has applications in a wide range of real-world applications, whose data often cannot be formulate precisely. Hence it makes perfect sense to apply fuzzy set theory as a way to mathematically describe this vagueness. In this paper we review the theory about this topic and describe some flexible and possibilistic programming models to solve fuzzy convex programming problems. Flexible programming uses fuzzy sets to represent the vagueness of the decision maker's aspirations and constraints, while possibilistic programming models imprecise or ambiguous data by possibility distributions.

1 Introduction

Mathematical programming is used to solve problems, achieving the best outcome of the objective function in a function domain that can be constrained or not. This kind of problem is called an optimization problem or

Carlos Cruz · José L. Verdegay
Department of Computer Science and Artificial Intelligence, University of Granada, E-18071, Granada, Spain
e-mail: {carloscruz,verdegay}@decsai.ugr.es

Ricardo C. Silva · Akebo Yamakami
Department of Telematics, School of Electrical and Computer Engineering, P.O. Box 6101, University of Campinas, 13083-970, Campinas, SP, Brazil
e-mail: {rcoelhos,akebo}@dt.fee.unicamp.br

W.A. Lodwick & J. Kacprzyk (Eds.): Fuzzy Optimization, STUDFUZZ 254, pp. 127–143.
springerlink.com

a mathematical programming problem, in which the aim is to find the best of all possible solutions. More formally, find a solution in the feasible region which has the minimum (or maximum) value of the objective function. If all the functions are linear, we obviously have a linear program. Otherwise, the problem is called a nonlinear program. However, many realistic problems cannot be adequately represented or approximated as a linear program owing to the nature of the non-linearity of the objective function and/or the non-linearity of any of the constraints. As it is well known Convex Programming represents a special class of mathematical programming in which the objective function is convex and the set of constraints are formed by convex functions over a convex decision space.

Thus, on the one hand, it is clear that convex programming encompasses all linear programming problems, including applications in scheduling, planning and flow computations, and they may be used to solve some interesting combinatorial optimization problems. On the other hand, it can be viewed as a particular case of nonlinear programming and it is more general than quadratic programming. Nowadays we can use highly efficient and robust algorithms and software for convex programming which are important tools for solving problems in diverse fields. However in many real practical applications one lacks of exact knowledge [14], and only approximate, vague and imprecise values are known. Experience shows that the best way of modeling these kinds of problems is using Soft Computing methodologies [33].

In recent years, Soft Computing, and Fuzzy Logic in particular, has shown great potential for modeling systems which are non-linear, complex, ill-defined and not well understood. Fuzzy Logic has found numerous and different applications due to its easy implementation, flexibility, tolerant nature to imprecise data, low cost implementations and ability to model non-linear behavior of arbitrary complexity because of its basis in terms of natural language.

In the fuzzy environment, as it happens in the case of linear programming problems, a variety of fuzzy convex programming problems can be defined: Convex programming problems with a fuzzy objective, i.e., with fuzzy numbers defining the costs of the objective function, convex programming problems with a fuzzy goal, i.e., with some fuzzy value to be attained in the objective, convex programming problems with fuzzy numbers defining the coefficients of the technological matrix and, finally, with a fuzzy constraint set, i.e., with a feasible set defined by fuzzy constraints.

Thus, fuzzy convex programming is applied in a wide range of disciplines, such as: control systems problems [13, 21, 29], production planning and scheduling problems in the complex industrial systems [26, 27], modeling multi-product aggregate production planning (APP) problems with fuzzy demands and fuzzy capacities [28], regression models [8, 25, 7], portfolio selection problem [1, 12, 16, 24, 31, 35, 34]. Some others interesting papers where various authors apply soft computing methodologies to convex programming are [2, 15, 17, 23, 26, 36].

With this in mind, the objective of this paper is to review the theory about this topic and to describe some flexible and possibilistic approaches to solve fuzzy convex programming problems remarking some limitations in their formulations.

The paper is organized as follows: Section 2 shows how is formulated a convex programming problem and what problems belong to it; Section 3 presents the formulation of a convex programming problem under fuzzy environment and describes some known flexible and possibilistic methods that were developed to solve fuzzy convex programming problems with uncertainties in the relationships or coefficients. In Section 4 is described a approach that solves convex programming problems with uncertainties in the relationships and an numerical example is solved by using this appraoch. Finally, conclusions are presented in Section 5.

2 Convex Programming

Some real-world problems can be formulated as mathematical programming problems that find to obtain the best solution according to the situation to be solved. This problems can have one or several objectives over a set of constraints or not, but only the problems with one objective will be described in this work. Then, these mathematical programming problems can be written as

$$\begin{aligned} &\min\ f_0(\mathbf{x}) \\ &\text{s.t.}\ \ f_i(\mathbf{x}) \leq \mathbf{0}\ i = 1,\ldots,m \\ &\qquad \mathbf{x} \in \Omega. \end{aligned} \tag{1}$$

where $\mathbf{x}$ is a decision variables vector, Ω is the feasible solutions convex set in $\mathbb{R}^n$, $f_0 : \mathbb{R}^n \to \mathbb{R}$ is the objective or cost function, and $f_i : \mathbb{R}^n \to \mathbb{R}$, for each $i = 1,\ldots,m$, are constraint functions.

Nevertheless, a convex optimization problem is one in which the objective function, f_0, is convex, and the feasible solution set formed by the constraints, $f_i (i = 1,\ldots,m)$, if any, form a convex set [5, 4]; i.e. the function satisfies the following equation:

$$f(\alpha\mathbf{x} + (1-\alpha)\mathbf{y}) \leq \alpha f(\mathbf{x}) + (1-\alpha) f(\mathbf{y})$$

for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ and all $\alpha \in [0,1]$.

According to the definition of convex optimization, it is easy to see that the linear programming problems belong to this kind of problems. There are many other problems that belong to the set of convex programming problems but we can highlight the quadratic programming problems that are most important in a great field of real-world problems and defined how a quadratic problem

if the objective function is quadratic and constraint functions are linear. Thus, Problem (1) can be rewritten as a quadratic programming problem in the following way:

$$\begin{aligned} \min\ & z = f_0(\mathbf{c},\mathbf{x}) = \mathbf{c}^{\mathbf{t}}\mathbf{x} + \tfrac{1}{2}\mathbf{x}^{\mathbf{t}}\mathbf{Q}\mathbf{x} \\ \text{s.t.}\ & \mathbf{f}(\mathbf{A},\mathbf{x}) = \mathbf{A}\mathbf{x} \leq \mathbf{b} \\ & \mathbf{x} \leq 0 \end{aligned} \tag{2}$$

where $\mathbf{c}$ is the n-dimensional vector and $\mathbf{Q}$ is the $n \times n$-dimensional matrix of profit coefficients of the objective function, $\mathbf{b}$ is the m-dimensional vector of total resources available, $\mathbf{A}$ is the matrix of technical coefficients, and $\mathbf{x}$ is the n-dimensional vector of decision variables (or alternatives).

If an optimization problem can be formulated as a convex optimization problem, then it is solved efficiently.

3 Fuzzy Convex Programming

Similar as the case of fuzzy linear programming problems [15], a large number of fuzzy convex programming problems can be defined. The uncertainties can be found in the relationships, constants, decision variables or in all parameters of the problem. In contrast to fuzzy linear programming problems where much research has been conducted, unfortunately, little has been done with this important class of problems.

A conventional programming problem (1) can be transformed into a fuzzy convex formulation with vagueness in all their parameters as:

$$\begin{aligned} & \widetilde{\min}\ f_0(\tilde{\mathbf{c}};\tilde{\mathbf{x}}) \\ & \text{s.t.}\ \ f_i(\tilde{\mathbf{a}};\tilde{\mathbf{x}}) \lesssim \tilde{\mathbf{0}}\ i = 1,\ldots,m \\ & \qquad \tilde{\mathbf{x}} \in \tilde{\Omega}. \end{aligned} \tag{3}$$

where $\tilde{\mathbf{x}}$ is the fuzzy variables decision vector, $\tilde{\Omega}$ is the fuzzy feasible solutions set, $\tilde{\mathbf{c}}$ is the fuzzy cost vector, $\tilde{\mathbf{a}}$ is the fuzzy coefficients vector of the constraint functions, and $\lesssim$ is the fuzzy relation.

Each fuzzy parameter in Problem (3) can be defined by membership functions such as: $\mu_i : \mathbb{R} \to [0,1],\ i = 1,\ldots,m$.

According to a general classification of fuzzy mathematical programming into *flexible programming* [12, 37] and *possibilistic programming* [10, 11, 22] several methods to solve convex programming will be presented in next subsections. The distinction between flexible (fuzzy) programming and possibilistic programming is developed in . Flexible programming uses fuzzy sets to represent the vagueness of the decision maker's aspirations and constraints and possibilistic programming models imprecise or ambiguous data by possibility distributions.

3.1 Flexible Convex Programming

In this subsection we will describe some approaches that use flexible optimization to solve fuzzy convex programming problems. The uncertainties can be found in the costs of the objective function and/or constants of the constraint functions of the problem.

3.1.1 Tang and Wang Approach

Tang and Wang [26, 27] proposed two methods to solve the quadratic programming problems. They study quadratic programming problems with a type of fuzzy objective and resource constraints and its solution method: an interactive approach. Also, they focus on a non-symmetric model for fuzzy nonlinear programming problems with penalty coefficients and attempts to develop a systematic approach to solve them. It uses a kind of nonlinear membership function to describe the fuzzy available resources and fuzzy constraints.

Hence, quadratic programming problems with uncertainty in the vector of the independent coefficients $\tilde{\mathbf{b}}$ is defined by Tang and Wang in the following way:

$$\begin{aligned} &\widetilde{\min}\ \mathbf{c^t x} + \tfrac{1}{2}\mathbf{x^t Q x} \\ &\text{s.t. } \mathbf{A x} \leq \tilde{\mathbf{b}} \\ &\qquad \mathbf{x} \in \Omega. \end{aligned} \tag{4}$$

A membership function for each fuzzy component of vector $\tilde{\mathbf{b}}$ is needed to solve this problem:

$$\mu_i(y) = \begin{cases} 1 & y \leq b_i \\ 1 - \dfrac{y - b_i}{d_i} & b_i \leq y \leq b_i + d_i \\ 0 & y > b_i + d_i \end{cases}$$

where $d_i, (i = 1, \ldots, m)$ is the allowed maximum tolerance in each restriction. Similar as extended Zimmermann's approach [2] the membership function can be transformed as:

$$\begin{aligned} &\max\ \alpha \\ &\text{s.t. } \mu_0(\mathbf{x}) \leq \alpha \\ &\qquad \mu_i(\mathbf{x}) \leq \alpha,\ i = 1, \ldots, m \\ &\qquad \mathbf{x} \in \Omega,\ \alpha \in [0, 1]. \end{aligned} \tag{5}$$

or

$$\begin{aligned} &\max \ \textit{Target} \\ &\text{s.t. } \mu_0(\mathbf{x}) \leq \alpha_0 \\ &\quad \mu_i(\mathbf{x}) \leq \alpha_0, \ i = 1, \ldots, m \\ &\quad \mathbf{x} \in \Omega, \ \alpha_0 \in [0,1]. \end{aligned} \tag{6}$$

where α_0 is a satisfaction degree that is in the interval [0,1]. This degree is an acceptable value chosen by the decision maker. *Target* can be an objective function, restrictions or another goal given by the decision maker.

According to described idea in [26, 27] an optimal solution can be defined as:

Definition 1. *A fuzzy optimal solution of Problem (4) is a fuzzy set $\tilde{S}$ defined by*

$$\tilde{S} = \{(\mathbf{x}, \mu_{\tilde{S}}(\mathbf{x})) | \mathbf{x} \in \Omega\}$$

with

$$\mu_{\tilde{S}}(\mathbf{x}) = \min\{\mu_0(\mathbf{x}), \mu_1(\mathbf{x}), \ldots, \mu_m(\mathbf{x})\}.$$

If

$$S_\alpha = \{\mathbf{x} \in \Omega | \mu_{\tilde{S}}(\mathbf{x}) \geq \alpha\},$$

where $\alpha \in [0,1]$. Then S_α is a set with cutting level $\geq \alpha$ of $\tilde{S}$.

Definition 2. *α^* is the best satisfaction level if there exists an $\alpha \in [0,1]$, such that $\forall\ 0 \leq \alpha < \alpha^*$, S_α is non-empty and, $\forall\ \alpha \geq \alpha^*$, S_α is empty.*

This approach uses a classic inequality concept to compare kind of different numbers. Thus, it is limited because this approach formulates a comparison of fuzzy numbers, vector of right-hand sides, with a crisp matrix of constraint coefficients.

3.1.2 Liu Approach

An approach to solve quadratic programming problems with fuzzy costs and fuzzy coefficients in the restrictions set was proposed by Liu [17, 18]. Vector constants $\mathbf{c}$ and $\mathbf{b}$, and the matrix $\mathbf{A}$ are uncertainties. Thus, Problem (2) can be transformed into fuzzy programming problem in the following form:

$$\begin{aligned} &\min \ \tilde{\mathbf{c}}\mathbf{x} + \tfrac{1}{2}\mathbf{x}^{\mathbf{t}}\mathbf{Q}\mathbf{x} \\ &\text{s.t. } \tilde{\mathbf{A}}\mathbf{x} \leq \tilde{\mathbf{b}} \\ &\quad \mathbf{x} \in \Omega. \end{aligned} \tag{7}$$

where $\tilde{\mathbf{c}} = \{(c_j, \mu_{\tilde{c}_j}(c_j)),\ j = 1,\ldots,n | c_j \in supp(\tilde{c}_j)\}$, $\tilde{\mathbf{A}} = \{(a_{ij}, \mu_{\tilde{a}_{ij}}(a_{ij})),\ i = 1,\ldots,m$ y $j = 1,\ldots,n | a_{ij} \in supp(\tilde{a}_{ij})\}$ and $\tilde{\mathbf{b}} = \{(b_i, \mu_{\tilde{b}_i}(b_i)),\ i = 1,\ldots,m | b_i \in supp(\tilde{b}_i)\}$.

The authors derive the membership function of the fuzzy goal, and then they apply Zadeh's extension principle to transform the fuzzy quadratic problem into a pair of two-level mathematical programs to calculate the upper and lower bounds of the objective value at possibility level. These programs can be solved by conventional optimization techniques. Thus, the membership function of the objective function can be defined as

$$\mu_{\tilde{Z}}(z) = \sup_{\mathbf{c},\mathbf{A},\mathbf{b}} \min\{\mu_{\tilde{c}_j}(c_j), \mu_{\tilde{a}_{ij}}(a_{ij}), \mu_{\tilde{b}_i}(b_i),\ \forall i, j | z = Z(\mathbf{c},\mathbf{A},\mathbf{b})\} \tag{8}$$

where $Z(\mathbf{c},\mathbf{A},\mathbf{b})$ is the function of the conventional quadratic problem (3). Membership function $\mu_{\tilde{Z}}$ can be computed by finding the functions that describe the shape of the left and right sides of the fuzzy numbers. Then, it is possible to obtain the upper bound of the objective value Z_α^U and the lower bound Z_α^L to each value α. Thus, Z_α^U is the maximum and Z_α^L is the minimum of $Z(\mathbf{c},\mathbf{A},\mathbf{b})$, respectively, that can be described as:

$$Z_\alpha^U = \max\{Z(\mathbf{c},\mathbf{A},\mathbf{b})\} \tag{9}$$

$$Z_\alpha^L = \min\{Z(\mathbf{c},\mathbf{A},\mathbf{b})\} \tag{10}$$

where each component j of the vector $\mathbf{c}$ belong to the interval $[(c_j)_\alpha^L, (c_j)_\alpha^U]$, each components ij of the matrix $\mathbf{A}$ belong to the interval $[(a_{ij})_\alpha^L, (a_{ij})_\alpha^U]$, and each component i of the vector $\mathbf{b}$ belong to the interval $[(b_i)_\alpha^L, (b_i)_\alpha^U]$, for all $i = 1,\ldots,m$ and $j = 1,\ldots,n$.

Different values of fuzzy parameters produce different objective values, then equations (9) and (10) can transform the fuzzy quadratic problem into two levels.

Using equation (9) the fuzzy problem can be transformed as:

$$Z_\alpha^U = \max_{\substack{(c_j)_\alpha^L \le c_j \le (c_j)_\alpha^U \\ (a_{ij})_\alpha^L \le a_{ij} \le (a_{ij})_\alpha^U \\ (b_i)_\alpha^L \le b_i \le (b_i)_\alpha}} \left\{ \begin{array}{l} \min_{\mathbf{x}} \sum_{j=1}^n c_j x_j + \frac{1}{2}\sum_{j=1}^n \sum_{l=1}^n q_{jl} x_j x_l \\ \text{s.t.} \quad \sum_{j=1}^n a_{ij} x_j \le b_i, \quad i = 1,\ldots,m \\ \qquad \mathbf{x} \in \Omega \end{array} \right. \tag{11}$$

where goal value Z_α^U is the upper bound of the classical quadratic programming problem.

Using equation (10) the fuzzy problem can be transformed as:

$$Z_\alpha^L = \min_{\substack{(c_j)_\alpha^L \le c_j \le (c_j)_\alpha^U \\ (a_{ij})_\alpha^L \le a_{ij} \le (a_{ij})_\alpha^U \\ (b_i)_\alpha^L \le b_i \le (b_i)_\alpha}} \left\{ \begin{array}{l} \min_{\mathbf{x}} \sum_{j=1}^n c_j x_j + \frac{1}{2}\sum_{j=1}^n \sum_{k=1}^n q_{jk} x_j x_k \\ \text{s.t.} \quad \sum_{j=1}^n a_{ij} x_j \le b_i, \quad i = 1,\ldots,m \\ \qquad \mathbf{x} \in \Omega \end{array} \right. \tag{12}$$

where goal value Z_{α}^{L} is the lower bound of the classical quadratic programming problem.

These two formulations above need two programs to solve them called outer-level and inner-level programs. Outer-level program obtains the values c_j, a_{ij} and b_i that are used with parameters by inner-level program. Inner-level program solves a classical quadratic programming problem with the data obtained by outer-level program. The authors state that the formulation of two-level quadratic problems is a generalization of the conventional parametric quadratic programming problem.

Thus, firstly, the two-level mathematical program is transformed into the following quadratic problem by dual formulation:

$$\begin{array}{l} \max \ -\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} h_{ij}x_i x_j - \sum_{i=1}^{m}(b_i)_{\alpha}^{U}\lambda_i \\ \text{s.t. } \sum_{i=1}^{n} h_{ij}x_i + \sum_{i=1}^{m} a_{ij}\lambda_i - \delta_j = -c_j, \quad j = 1,2,\ldots,n \\ \quad (c_j)_{\alpha}^{L} \le c_j \le (c_j)_{\alpha}^{U}, \quad j = 1,2,\ldots,n \\ \quad (a_{ij})_{\alpha}^{L} \le a_{ij} \le (a_{ij})_{\alpha}^{U}, \quad i = 1,2,\ldots,m, \quad j = 1,2,\ldots,n \\ \quad \lambda_i, \delta_j \ge 0, \quad i = 1,2,\ldots,m \text{ and } j = 1,2,\ldots,n. \end{array} \tag{13}$$

Since both the inner-level program and outer-level of the second program have the same minimization operation, they can be combined into a conventional one-level program with the constraints of the two programs considered simultaneously. Consequently, some points must be analyzed which are shown in [17, 18]. The second program can be described as:

$$\begin{array}{l} \max \ \sum_{j=1}^{n}(c_j)_{\alpha}^{L} + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} h_{ij}x_i x_j \\ \text{s.t. } \sum_{i=1}^{n}(a_{ij})_{\alpha}^{L}x_i \le (b_i)_{\alpha}^{U}, \quad i = 1,2,\ldots,m \\ \quad x_j \ge 0, \quad j = 1,2,\ldots,n \end{array} \tag{14}$$

This approach is limited because it uses in their formulation a crisp inequality to compare fuzzy numbers, vector of right-hand sides, with a fuzzy matrix of constraint coefficients.

3.1.3 Ammar and Khalifa Approach

An approach to solve quadratic programming problem with fuzzy costs $\tilde{Q}$, fuzzy matrix coefficients $\tilde{A}$ and the restrictions set vector $\tilde{b}$ was proposed in [1]. Thus, Problem (2) can be defined in the following form:

$$\begin{array}{l} \min \ \mathbf{x^t}\tilde{\mathbf{Q}}\mathbf{x} \\ \text{s.t. } \tilde{\mathbf{A}}\mathbf{x} \le \tilde{\mathbf{b}} \\ \quad \mathbf{x} \in \Omega. \end{array} \tag{15}$$

where all decision variables are non-negative

Problem (15), can be defined in α-cut intervals as:

$$\begin{aligned}(P_\alpha): \min\ & \mathbf{x^t}[\mathbf{Q}^{\mathbf{L}}_\alpha, \mathbf{Q}^{\mathbf{U}}_\alpha]\mathbf{x} \\ \text{s.t. } & [\mathbf{A}^{\mathbf{L}}_\alpha, \mathbf{A}^{\mathbf{U}}_\alpha]\mathbf{x} \leq [\mathbf{b}^{\mathbf{L}}_\alpha, \mathbf{b}^{\mathbf{U}}_\alpha] \\ & \mathbf{x} \in \Omega.\end{aligned} \tag{16}$$

where all decision variables are non-negative and $\alpha \in (0,1]$.

Problem (16) can be divided into two classic problems. Using the lower bound of the interval of α-cuts for the first problem and using the upper bound for the second we can transform this problem as:

$$\begin{aligned}(P^L_\alpha): \min\ & \mathbf{x^t}\mathbf{Q}^{\mathbf{L}}_\alpha\mathbf{x} \\ \text{s.t. } & \mathbf{A}^{\mathbf{L}}_\alpha\mathbf{x} \leq \mathbf{b}^{\mathbf{L}}_\alpha \\ & \mathbf{x} \in \Omega.\end{aligned} \tag{17}$$

$$\begin{aligned}(P^U_\alpha): \min\ & \mathbf{x^t}\mathbf{Q}^{\mathbf{U}}_\alpha\mathbf{x} \\ \text{s.t. } & \mathbf{A}^{\mathbf{U}}_\alpha\mathbf{x} \leq \mathbf{b}^{\mathbf{U}}_\alpha \\ & \mathbf{x} \in \Omega.\end{aligned} \tag{18}$$

where all decision variables are non-negative and $\alpha \in (0,1]$.

These problems can be solved by Karush-Kuhn-Tucker´s conditions. The optimal solution of original problem (16), (P_α), is inside the solution interval formed by each value of $\alpha \in [0,1]$, and where bounds are obtained by some convex optimization technique in the problems (P^L_α) and (P^U_α).

This approach, similar as above approach, is limited because it formulates a crisp inequality to compare fuzzy numbers, vector of right-hand sides, with a fuzzy matrix of constraint coefficients. In addition, there is other limitation because it divides the cost matrix with fuzzy elements in two parts which form two quadratic programming problems. One of this problems is formed with the inferior boundaries while the other is formed with the superior ones. However, there not exists a proof that these problems obtain the inferior and superior boundaries of the optimal solution, respectively.

3.2 Possibilistic Convex Programming

Some approaches that use possibilistic optimization to solve uncertain convex programming problems will be showed in this subsection. The uncertainties of problems will be described by the possibility theory. Possibilistic programming uses real-valued entities that exist, but the evidence associated with whether or not a particular element belongs to the set is incomplete or hard to obtain.

3.2.1 Canestrelli, Giove and Fullér Approach

Possibilistic quadratic problems with classic variables and imprecise coefficients can be well-posed with small change on possibilistic distribution of objective function when a small change of membership function is provoked. This happens if all α-level sets of two vagueness numbers are close to each other then there can be only a small difference between their membership degrees. The application of this approach described in [6] is used in a possibilistic quadratic problem that is defined as:

$$\begin{aligned} \min\ & Z = \mathbf{c}^{\mathbf{t}}\mathbf{x} + \mathbf{x}^{\mathbf{t}}\mathbf{Q}\mathbf{x} \\ \text{s.t. } & \mathbf{A}\mathbf{x} \leq \mathbf{b} \\ & \mathbf{x} \in \Omega. \end{aligned} \tag{19}$$

where $\tilde{\mathbf{Q}}$ y $\tilde{\mathbf{A}}$ are matrixes with imprecise numbers, $\tilde{\mathbf{c}}$ and $\tilde{\mathbf{b}}$ are vectors with imprecise numbers, and Ω is a decision variables set in $\mathbb{R}^n$. Each imprecise value is associated to a possibilistic distribution.

$Poss[Z = z]$ is defined as possibilistic distribution of the objective function Z as developed in the following paragraphs. However, first it is necessary to determine the possibility that $\mathbf{x}$ satisfy the i-th constraint which is described as:

$$Poss[\mathbf{x} \in \mathscr{F}_i] = \sup_{(\mathbf{A})_i, b_i} \{\Pi((\mathbf{A})_i, b_i) | (\mathbf{A}\mathbf{x})_i \leq b_i\}$$

where $\Pi((\mathbf{A})_i, b_i) = \min\{\tilde{a}_{i1}(a_{i1}), \ldots, \tilde{a}_{in}(a_{in}), \tilde{b}_i(b_i)\}$, $(\mathbf{A})_i$ is the marginal possibilistic distribution of $(\tilde{\mathbf{A}})_i$, and b_i is the marginal possibilistic distribution of $(\tilde{b})_i$, for all $i = 1, \ldots, m$. Then, for $\mathbf{x} \in \Omega$,

$$Poss[\mathbf{x} \in \mathscr{F}] = \min_{i=1,\ldots,m} Poss[\mathbf{x} \in \mathscr{F}_i].$$

A conditional possibility, $Poss[Z = z|\mathbf{x}]$, is defined in the second phase. Then, the degree of possibility of objective function is formulated as:

$$Poss[Z = z|\mathbf{x}] = \sup_{\mathbf{c},\mathbf{Q}} \{\Pi(\mathbf{c}, \mathbf{Q}) | \mathbf{c}^{\mathbf{t}}\mathbf{x} + \mathbf{x}^{\mathbf{t}}\mathbf{Q}\mathbf{x} = z\}$$

where $\Pi(\mathbf{c}, \mathbf{Q}) = \min_{i,j}\{\tilde{c}_j(c_j), \tilde{q}_{ij}(q_{ij})\}$.

Therefore, applying the decision method of Bellman and Zadeh [3], the problem distribution possibilistic is defined as:

$$Poss[Z = z] = \sup_{\mathbf{x} \in \Omega} \min\{Poss[Z = z|\mathbf{x}], Poss[\mathbf{x} \in \mathscr{F}]\}.$$

Also, this approach formulates a crisp inequality to compare fuzzy numbers, vector of right-hand sides, with a fuzzy matrix of constraint coefficients.

3.2.2 Tonon and Bernardini Approach

An approach to solve fuzzy convex programming with fuzzy costs and fuzzy coefficients in restrictions set is described in [30]. Problem (1) can be formulated as a fuzzy convex problem in the following form:

$$\begin{aligned} &\min\ f_0(\mathbf{u}_0;\mathbf{x}) \\ &\text{s.t.}\ \ f_i(\mathbf{u}_i;\mathbf{x}) \leq \mathbf{0}\ i = 1,\ldots,m \\ &\qquad \mathbf{x} \in \Omega. \end{aligned} \tag{20}$$

where $\mathbf{x}$ is a decision variables vector in the feasible solutions set Ω, f is the objective function, g_i are restrictions set functions for each $i = 1,\ldots,m$, and $\mathbf{u}_p$ are fuzzy parameters vectors for each $p = 0,1,\ldots,m$ in the objective function and restrictions.

Vectors $\mathbf{u}_j$ can be ordered as:

$$\mathbf{u}_j = \{u_{i,1},\ldots,u_{i,k_i},u_{i,k_i+1,\ldots,u_{i,k_i+l_i}}\}.$$

Non-interactive parameters are in the first k_i positions and interactive parameters are in the next l_i vector positions [30]. The allowed values of $u_{i,j},\ j = 1,\ldots,k_i$ are restricted by the possibility distribution function $F_{i,j}$. Each dependent parameters vector, $u_{i,j},\ j = k_i+1,\ldots,k_i+l_i$, is restricted by a fuzzy relation $F_{i,j}$.

A level $\alpha_{i,j} \in [0,1]$, which is chosen by decision maker, for each vector $\mathbf{u}_i$ is selected for $i = 0,\ldots,m$ and $j = 1,\ldots,k_i+l_i$, and each vector varies in a Ψ_i set:

$$\Psi_i = \{F_{i,1}(\alpha_{i,1}) \times \ldots \times F_{i,k_i+l_i}(\alpha_{i,k_i+l_i})\}$$

where $F_{i,j}(\alpha_{i,j})$ marks a cut in the level $\alpha_{i,j}$ of fuzzy set $F_{i,j}$. Thus, Problem (20) can be transformed as:

$$\begin{aligned} &\min_{\mathbf{x},y}\ y \\ &\text{s.t.}\quad \begin{cases} f_i(\mathbf{u}_i;\mathbf{x}) \leq \mathbf{0} & \forall\ \mathbf{u}_i \in \Psi - i,\ i = 1,\ldots,m \\ f_0(\mathbf{u}_0;\mathbf{x}) - y \leq 0 & \forall\ \mathbf{u}_0 \in \Psi_0 \end{cases} \qquad \mathbf{x} \in \Omega. \end{aligned} \tag{21}$$

4 Extended Verdegay's Linear Approach ([32])

As in [9], the constraints of a problem are defined as having a fuzzy nature, that is, some violations in the accomplishment of such restrictions are permitted. In this way, this approach tries to solve the limitations of formulations of the almost last approaches. Therefore if we denote each constraint $\sum_{j\in J} a_{ij}x_j$, by $(Ax)_i$, Problem (2) can be addressed as follows:

$$\begin{aligned} &\min \mathbf{c}^{\mathbf{t}}\mathbf{x} + \tfrac{1}{2}\mathbf{x}^{\mathbf{t}}\mathbf{Q}\mathbf{x} \\ &\text{s.t. } (\mathbf{A}\mathbf{x})_i \lesssim b_i, i \in I \\ &\quad \mathbf{x} \in \Omega \end{aligned} \tag{22}$$

where the membership functions:

$$\mu_i : \mathbb{R}^n \to (0,1], \quad i \in I$$

of the fuzzy constraints are to be determined by decision maker. It is clear that each membership function will give the membership (satisfaction) degree such that any $\mathbf{x} \in \mathbb{R}^n$ accomplishes the corresponding fuzzy constraint upon which it is defined. This degree is equal to 1 when the constraint is perfectly accomplished (no violation), and decreases to zero for greater violations. For non-admissible violations the accomplishment degree will equal zero in all cases. In the linear case, these membership functions can be formulated as follows:

$$\mu_i(x) = \begin{cases} 1 & (\mathbf{A}\mathbf{x})_i \leq b_i \\ 1 - \dfrac{(\mathbf{A}\mathbf{x})_i - b_i}{d_i} & b_i \leq (\mathbf{A}\mathbf{x})_i \leq b_i + d_i \\ 0 & (\mathbf{A}\mathbf{x})_i > b_i + d_i \end{cases}$$

In order to solve this problem in a two-phase method, as it was shown in [23], first let us define for each fuzzy constraint, $i \in I$

$$X_i = \{\mathbf{x} \in \mathbb{R}^n / (\mathbf{A}\mathbf{x})_i \lesssim b_i, \mathbf{x} \in \Omega\}.$$

If $\mathbf{X} = \bigcap_{i \in I} X_i$ then last fuzzy quadratic problem can be described as:

$$\min \left\{ \mathbf{c}^{\mathbf{t}}\mathbf{x} + \frac{1}{2}\mathbf{x}^{\mathbf{t}}\mathbf{Q}\mathbf{x} / \mathbf{x} \in \mathbf{X} \right\}$$

It is clear that $\forall \alpha \in (0,1]$ an α-cut of the fuzzy constraint set will be the classical set

$$X(\alpha) = \{\mathbf{x} \in \mathbb{R}^n / \mu_X(\mathbf{x}) \geq \alpha\}$$

where $\forall x \in \mathbb{R}^n$,

$$\mu_X(\mathbf{x}) = \min \mu_i(\mathbf{x}), i \in I$$

in which the **inf** function is used because the fuzzy number can be non-closed set but the I set is finite.

Hence an α-cut of the i-th constraint will be denoted by $X_i(\alpha)$. Therefore, if $\forall \alpha \in (0,1]$,

$$S(\alpha) = \left\{ \mathbf{x} \in \mathbb{R}^n / \mathbf{c}^{\mathbf{t}}\mathbf{x} + \frac{1}{2}\mathbf{x}^{\mathbf{t}}\mathbf{Q}\mathbf{x} = \min \ \mathbf{c}^{\mathbf{t}}\mathbf{y} + \frac{1}{2}\mathbf{y}^{\mathbf{t}}\mathbf{Q}\mathbf{y}, \mathbf{y} \in X(\alpha) \right\}$$

where the change of the variable is used to find the best solution in the set $X(\alpha)$ formed by $x \in \mathbb{R}^n$. The fuzzy solution to the problem will be the fuzzy set defined by the following membership function

$$S(\alpha) = \begin{cases} \sup\{\alpha : \mathbf{x} \in S(\alpha)\}, & \mathbf{x} \in \bigcup_\alpha S(\alpha) \\ 0, & \text{otherwise.} \end{cases}$$

Provided that $\forall \alpha \in (0,1]$,

$$X(\alpha) = \bigcap_{i \in I} \{x \in \mathbb{R}^n / (\mathbf{Ax})_i \leq r_i(\alpha), \mathbf{x} \in \Omega\}$$

with $r_i(\alpha) = b_i + d_i(1-\alpha)$. The operative solution to the former problem can be found, α-cut by α-cut, by means of the following auxiliary parametric classic programming model,

$$\begin{aligned} &\min \ \mathbf{c^t x} + \tfrac{1}{2}\mathbf{x^t Q x} \\ &\text{s.t.} \ \ (\mathbf{Ax})_i \leq b_i + d_i(1-\alpha), i \in I \\ &\qquad \mathbf{x} \in \Omega, \alpha \in (0,1]. \end{aligned} \tag{23}$$

It is easy to see that the first phase ends when the fuzzy convex programming problem is transformed into several classic convex programming problems. Each one of this problems depends on a parameter which represents the satisfaction level defined by decision maker.

In the second phase the parametric quadratic programming problem is solved for each of the different α values using conventional quadratic programming techniques. We must find solutions to Problem (23) for each α that satisfies Karush-Kuhn-Tucker´s necessary and sufficient optimality conditions. One of the conventional techniques is to decide the Lagrange function that is a transformation of Problem (23) in a unconstrained mathematical problem:

$$L(\mathbf{x}, \mu, \nu) = \mathbf{c^t x} + \tfrac{1}{2}\mathbf{x^t Q x} + \mu^t(\mathbf{Ax} - \mathbf{b} + \mathbf{d}(1-\alpha)) + \nu^t(\mathbf{0} - \mathbf{x}) \tag{24}$$

where μ and ν are the Lagrange multipliers for the inequality and non-negativity constraints respectively.

Each α is associated to a optimal solution of the parametric convex programming problem and this solutions are called satisfactory solution which generate a set of solutions. Then the Representation Theorem can be used to integrate all these specific α-solutions.

Example 1. In order to show the performance of our method, we used the set of historical data shown in Table 1 introduced by Markowitz. The columns 2-10 represent American Tobacco, A.T.&T., United States Steel, General Motors, Atcheson&Topeka&Santa Fe, Coca-Cola, Borden, Firestone and Sharon Steel securities data, respectively. The returns on the nine securities, during the years 1937-54, are presented in Table 1.

Table 1 Fuzzy portfolio selection problem

	#1	#2	#3	#4	#5	#6	#7	#8	#9
Year	Am.T	A.T&T.	U.S.S.	G.M.	A.T.&S.	C.C.	Bdm.	Frstn.	S.S.
1937	-0.305	-0.173	-0.318	-0.477	-0.457	-0.065	-0.319	-0.4	-0.435
1938	0.513	0.098	0.285	0.714	0.107	0.238	0.076	0.336	0.238
1939	0.055	0.2	-0.047	0.165	-0.424	-0.078	0.381	-0.093	-0.295
1940	-0.126	0.03	0.104	-0.043	-0.189	-0.077	-0.051	-0.09	-0.036
1941	-0.28	-0.183	-0.171	-0.277	0.637	-0.187	0.087	-0.194	-0.24
1942	-0.003	0.067	-0.039	0.476	0.865	0.156	0.262	1.113	0.126
1943	0.428	0.300	0.149	0.255	0.313	0.351	0.341	0.580	0.639
1944	0.192	0.103	0.260	0.290	0.637	0.233	0.227	0.473	0.282
1945	0.446	0.216	0.419	0.216	0.373	0.349	0.352	0.229	0.578
1946	-0.088	-0.046	-0.078	-0.272	-0.037	-0.209	0.153	-0.126	0.289
1947	-0.127	-0.071	0.169	0.144	0.026	0.355	-0.099	0.009	0.184
1948	-0.015	0.056	-0.035	0.107	0.153	-0.231	0.038	0	0.114
1949	0.305	0.030	0.133	0.321	0.067	0.246	0.273	0.223	-0.222
1950	-0.096	0.089	0.732	0.305	0.579	-0.248	0.091	0.650	0.327
1951	0.016	0.090	0.021	0.195	0.040	-0.064	0.054	-0.131	0.333
1952	0.128	0.083	0.131	0.390	0.434	0.079	0.109	0.175	0.062
1953	-0.010	0.035	0.006	-0.072	-0.027	0.067	0.21	-0.084	-0.048
1954	0.154	0.176	0.908	0.715	0.469	0.077	0.112	0.756	0.185

This example will consider performances of portfolios with respect to "return" thus defined. This assumes that a dollar of realized or unrealized capital gains is exactly equivalent to a dollar of dividends, no better and no worse. This assumption is appropriate for certain investors, for example, some types of tax-free institutions. Other ways of handling capital gains and dividends, which are appropriate for other investors, can be viewed in [20].

Here we show the results obtained for the porfolio selection problem, described by Table 1, by the fuzzy quadratic programming methods introduced in this section and its solution is shown in Table 2. By computing the average value of all the years of each column of random variables of Table 1, we obtained the expected values of each return of the securities.

Table 2 Results of the first phase of the portfolio selection problem

α	x	Solution	Time
0.0	[-0.0000 ; 0.1236 ; 0.1374 ; 0.0000 ; 0.0910 ; 0.0641 ; 0.5838 ; 0.0000 ; 0.0000]	0.0235	1.0313
0.1	[0.0000 ; 0.0698 ; 0.1509 ; 0.0000 ; 0.0924 ; 0.0610 ; 0.6259 ; 0.0000 ; -0.0000]	0.0247	0.1563
0.2	[0.0000 ; 0.0303 ; 0.1713 ; 0.0000 ; 0.0936 ; 0.0465 ; 0.6584 ; 0.0000 ; -0.0000]	0.0259	0.1250
0.3	[0.0000 ; -0.0000 ; 0.1768 ; 0.0056 ; 0.0967 ; 0.0251 ; 0.6958 ; 0.0000 ; -0.0000]	0.0272	0.1563
0.4	[0.0000 ; -0.0000 ; 0.1627 ; 0.0306 ; 0.1168 ; -0.0000 ; 0.6899 ; -0.0000 ; 0.0000]	0.0287	0.1406
0.5	[0.0000 ; -0.0000 ; 0.1323 ; 0.0736 ; 0.1553 ; -0.0000 ; 0.6388 ; -0.0000 ; 0.0000]	0.0307	0.1563
0.6	[-0.0000 ; -0.0000 ; 0.0975 ; 0.1224 ; 0.1912 ; 0.0000 ; 0.5889 ; 0.0000 ; 0.0000]	0.0332	0.1250
0.7	[-0.0000 ; -0.0000 ; 0.0703 ; 0.1554 ; 0.2354 ; 0.0000 ; 0.5389 ; -0.0000 ; 0.0000]	0.0364	0.1250
0.8	[0.0000 ; 0.0000 ; 0.0368 ; 0.2002 ; 0.2735 ; 0.0000 ; 0.4894 ; 0.0000 ; -0.0000]	0.0402	0.1250
0.9	[0.0000 ; -0.0000 ; 0.0048 ; 0.2422 ; 0.3131 ; 0.0000 ; 0.4399 ; 0.0000 ; -0.0000]	0.0445	0.1250
1.0	[-0.0000 ; -0.0000 ; -0.0000 ; 0.2717 ; 0.3537 ; 0.0000 ; 0.3746 ; -0.0000 ; 0.0000]	0.0495	0.1250

5 Conclusions

Fuzzy Convex Programming problems are of utmost importance in an increasing variety of practical fields, but unfortunately little study has been done with this important class of problems.

This paper shows a general view about fuzzy convex mathematical programming and some known methods that were developed to solve convex problems with vagueness in different parts are described. Tang and Wang's approach can transform a flexible programming problem into two classical problems. The transformed problems have different constraints where are limited by α-levels which may be defined by decision maker or used as objective function of the classical problem. However this approach is limited because it formulates a comparison of fuzzy numbers, vector right-hand sides, with a crisp matrix of constraint coefficients. Liu's and, Ammar and Khalifa's approaches transform flexible programming problem into two classical programming problems that determine the superior and inferior boundaries to each α-level chosen by decision maker. Thus, this interval is a solution set to the original flexible programming problem but there is not a proof that these problems obtain the inferior and superior boundaries of the optimal solution, respectively.

The limitations of formulations of the presented approaches are solved with the extension of Verdegay's linear approach because our approach is not necessary to choose α-level. Therefore, this approach can be used as a general method to solve convex programming problems with uncertainties in the set of constraints.

In addition, there are other approaches that deal with the uncertain data in optimization problems. For example, we showed two approaches that use the possibilistic programming which represents imprecise data by possibility distributions. In this case, the real-valued exist, but it may be or not to belong to the set is incomplete or hard to obtain. However, we cannot take a match of these approaches because they model the uncertainties of different way.

Acknowledgements. The authors want to thank the support provided by the Brazilian agency CAPES and the Spanish projects TIN2008-06872-C04-04, TIN2008-01948, P07-TIC-02970 and 04552/GERM/06.

References

1. Ammar, E., Khalifa, H.A.: Fuzzy portfolio optimization a quadratic programming approach. Chaos, Solitons and Fractals 18, 1045–1054 (2003)
2. Bector, C.R., Chandra, S.: Fuzzy mathematical programming and fuzzy matrix games. Studies in Fuzziness and Soft Computing, vol. 169. Springer, Berlin (2005)
3. Bellman, R.E., Zadeh, L.A.: Decision-marking in a fuzzy environment. Management Science 17(4), B141–B164 (1970)
4. Bertsekas, D.P., Nedié, A., Azdaglar, A.E.: Convex analysis and optimization. Athena Scientific, Belmont (2003)
5. Boyd, S., Vandenberghe, L.: Convex optimization. Cambridge University Press, Cambridge (2004)
6. Canestrelli, E., Giove, S., Fullér, R.: Stability in possibilistic quadratic programming. Fuzzy Sets and Systems 82, 51–56 (1996)
7. Chen, Y.S.: Fuzzy ranking and quadratic fuzzy regression. Computers and Mathematics with Applications 38, 265–279 (1999)
8. Donoso, S., Marín, N., Vila, M.A.: Fuzzy Regression with Quadratic Programming: An Application to Financial Data. In: Corchado, E., Yin, H., Botti, V., Fyfe, C. (eds.) IDEAL 2006. LNCS, vol. 4224, pp. 1304–1311. Springer, Heidelberg (2006)
9. Delgado, M., Verdegay, J.L., Vila, M.A.: A general model for fuzzy linear programming. Fuzzy Sets and Systems 29, 21–29 (1989)
10. Dubois, D., Prade, H.: Fuzzy sets and systems: theory and applications. Academic Press, San Diego (1980)
11. Dubois, D., Prade, H.: Possibility theory: an approach to computerized processing of uncertain. Plenum Press, New York (1988)
12. Inuiguchi, M., Ramik, J.: Possibilistic linear programming: A brief review of fuzzy mathematical programming and a comparison with stochastic programming in portfolio selection problem. Fuzzy Sets and Systems 111(1), 3–28 (2000)
13. Kim, E., Kang, H.J., Park, M.: Numerical stability analysis of fuzzy control systems via quadratic programming and linear matrix inequalities. IEEE Transactions on Systems, Man and Cybernetics - Part A: Systems and Humans 29(4), 333–346 (1999)
14. Klir, G.: Where do we stand on measures of uncertainty, ambiguity, fuzziness, and the like. Fuzzy Sets and Systems 24, 141–160 (1987)
15. Lai, Y.J., Hwang, C.L.: Fuzzy mathematical programming: methods and applications. Lecture notes in economics and mathematical systems, vol. 394. Springer, Berlin (1992)
16. León, T., Liern, V., Vercher, E.: Viability of infeasible portfolio selection problems: A fuzzy approach. European Journal of Operational Research 139, 178–189 (2002)
17. Liu, S.T.: Quadratic programming with fuzzy parameters: A membership function approach. Chaos, Solitons and Fractals (2007) (Article in press)

18. Liu, S.T.: Solving quadratic programming with fuzzy parameters based on extension principle. In: IEEE International Conference Fuzzy Systems, FUZZ-IEEE 2007, London, UK, pp. 1–5 (2007)
19. Lodwick, W.A., Jamison, K.D.: Theoretical and semantic distinctions of fuzzy, possibilistic, and mixed fuzzy/possibilistic optimization. Fuzzy Sets and Systems 158, 1861–1872 (2007)
20. Markowitz, H.M.: Portfolio Selection: Efficient Diversification of Investments, 2nd edn. Blackwell Publisher, Massachusetts (1991)
21. Mollov, S., Babuska, R., Abonyi, J., Verbruggen, H.B.: Effective optimization for fuzzy model predictive control. IEEE Transactions on Fuzzy Systems 12(5), 661–675 (2004)
22. Negoita, C.V., Ralescu, D.A.: Applications of fuzzy sets to systems analysis. Birkhauser Verlag, Stuttgard (1975)
23. Silva, R.C., Verdegay, J.L., Yamakami, A.: Two-phase method to solve fuzzy quadratic programming problems. In: IEEE International Conference Fuzzy Systems, FUZZ-IEEE 2007, London, UK, pp. 1–6 (2007)
24. Tanaka, H., Guo, P., Türksen, B.: Portfolio selection based on fuzzy probabilities and possibility distributions. Fuzzy Sets and Systems 111, 387–397 (2000)
25. Tanaka, H., Lee, H.: Interval regression analysis by quadratic programming approach. IEEE Transactions on Fuzzy Systems 6(4), 473–481 (1998)
26. Tang, J., Wang, D.: An interactive approach based on a genetic algorithm for a type of quadratic programming problems with fuzzy objective and resources. Computers & Operations Research 24(5), 413–422 (1997)
27. Tang, J., Wang, D.: A nonsymmetric model for fuzzy nonlinear programming problms with penalty coefficients. Computers & Operations Research 24(8), 717–725 (1997)
28. Tang, J., Wang, D., Fung, R.Y.K.: Fuzzy formulation for multi-product aggregate production planning. Production, Planning & Control 11(7), 670–676 (2000)
29. Tankeh, A., Mandani, H.: Convex fuzzy controller: neuro-fuzzy and convex optimisation. In: Sixth IEEE International Conference on Fuzzy Systems, pp. 1133–1139 (1997)
30. Tonon, F., Bernardini, A.: A random set approach to the optimizatino of uncertain structures. Computers and Structures 68, 583–600 (1998)
31. Vercher, E., Bermúdez, J., Segura, J.: Portfolio selection based on on fuzzy probability and possibility distributions. Fuzzy Sets and Systems 158(7), 769–782 (2007)
32. Verdegay, J.L.: Fuzzy mathematical programming. In: Gupta, M.M., Sanches, E. (eds.) Fuzzy Information and Decision Procsses, pp. 231–237. North-Holland Publishing Company, Amsterdam (1982)
33. Verdegay, J.L., Yager, R.R., Bonissone, P.P.: On heuristics as a fundamental constituent of soft computing. Fuzzy Sets and Systems 159(7), 846–855 (2008)
34. Wang, S., Zhu, S.: On fuzzy portfolio selection problems. Fuzzy Optimization and Decision Making 1, 361–377 (2002)
35. Watada, J.: Fuzzy portfolio selection and its applications to decision making. Tatra Mountains Mathematics Publication 13, 219–248 (1997)
36. Xu, Z., Da, Q., Chen, Q.: Priority approach based on quadratic programming model to fuzzy preference relation. Journal of Southeast University (English Edition) 21(1), 108–110 (2005)
37. Zimmermann, H.: Applications of fuzzy sets theory to mathematical programming. Information Sciences 36, 29–58 (1985)

Approaches to Linear Programming Problems with Interactive Fuzzy Numbers

Masahiro Inuiguchi

Abstract. Fuzzy programming has been developed mainly under the assumption of non-interaction among fuzzy coefficients. However, this assumption is not always suitable in the treatment of real world problems. Several approaches have been proposed to treat the interaction among fuzzy coefficients. In this paper, we review treatments of interaction among fuzzy coefficients in fuzzy linear programming problems. Using a necessity fractile model of a simple linear programming with fuzzy coefficients, we will see the differences among non-interactive case and five approaches to the treatment of interaction by showing the reduced problems. The five approaches are weak independent fuzzy numbers, a fuzzy vector with a quadratic membership function, scenario decomposed fuzzy numbers, an oblique fuzzy vector and a fuzzy polytope.

1 Introduction

Fuzzy programming approach [4, 11, 14] is useful and efficient to treat a programming problem under uncertainty. While classical and stochastic programming approaches may require a lot of cost to obtain the exact coefficient value or distribution, fuzzy programming approach does not (see Rommelfanger [12]). From this fact, fuzzy programming approach will be very advantageous when the coefficients are not known exactly but vaguely by human expertise.

Fuzzy programming [4, 11, 14] has been developed under an implicit assumption that all uncertain coefficients are non-interactive one another, with few exceptions. The non-interaction is a similar concept to the independence

Masahiro Inuiguchi
Graduate School of Engineering Science, Osaka University, 1-3 Machikaneyama, Osaka 560-8531, Japan
e-mail: inuiguti@sys.es.osaka-u.ac.jp

W.A. Lodwick & J. Kacprzyk (Eds.): Fuzzy Optimization, STUDFUZZ 254, pp. 145–161.
springerlink.com

which means that the possible range of an uncertain coefficient is unchanged even if we know the exact value of any other uncertain coefficient. The non-interaction is a little weaker mathematical concept than the independence (see Hisdal [1]) but, in practice, the independence rather than non-interaction would be assumed because of its simple meaning.

The assumption of non-interaction makes the reduced problem very tractable. The tractability can be seen as one of advantages of fuzzy programming approaches (see Inuiguchi et al. [4]). However, it is observed that in a simple problem, such as a portfolio selection problem, solutions of models are often intuitively unacceptable because of the implicit assumption (see Inuiguchi and Tanino [7]). As is known in the stock market, share prices of companies are not totally independent but somehow dependent (interactive). Such interaction can occur in parameters of fuzzy mathematical programming problems. Therefore, we may conclude that the non-interaction assumption is not sufficient to model all real world problems.

While dependencies among uncertain coefficients can be treated rather easily by covariances, or almost equivalently, correlations in stochastic programming, generally speaking, interaction among uncertain coefficients cannot be treated easily in fuzzy programming because the reduced problems often become non-convex problems. However, recently, it is shown that some special types of interaction among uncertain coefficients can be treated without great loss of tractability in fuzzy linear programming. In this paper, we review several approaches to treat interaction among uncertain coefficients.

2 Problem Statement

In order to show the difference among treatments of the interaction, we consider a linear programming with uncertain parameters. In this paper, we consider the following linear programming with a single fuzzy objective function:

$$\begin{array}{ll} \text{minimize} & \boldsymbol{\gamma}^{\mathrm{T}}\boldsymbol{x}, \\ \text{subject to} & A\boldsymbol{x} \leq \boldsymbol{b}, \end{array} \tag{1}$$

where A is a constant $m \times n$ matrix and $\boldsymbol{b} = (b_1, b_2, \ldots, b_m)^{\mathrm{T}}$ is a constant m-dimensional vector. $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)^{\mathrm{T}}$ is a decision variable vector. $\boldsymbol{\gamma} = (\gamma_1, \gamma_2, \ldots, \gamma_n)^{\mathrm{T}}$ is an uncertain vector which is often called a fuzzy vector or a possibilistic vector. We assume that $\boldsymbol{\gamma}$ is not totally uncertain but its possible range is known as a fuzzy set $\boldsymbol{C}$ by some way, e.g., by the knowledge/experience of the experts/decision maker or by data showing the partial information. $\boldsymbol{C}$ is an n-dimensional fuzzy set and its membership function is denoted by $\mu_{\boldsymbol{C}}$. Note that Problem (1) does not always have the non-negativity condition for $\boldsymbol{x}$.

In order to treat the uncertainty in Problem (1), we apply a necessity fractile model [4]. Given fuzzy information about conceivable values as a

fuzzy set D, we define a necessity measure as a function of a fuzzy event E by

$$N_D(E) = \inf_r \max\left(1 - \mu_D(r), \mu_E(r)\right), \tag{2}$$

where μ_D and μ_E are membership functions of D and E, respectively. Value $N_D(E)$ shows the necessity (certainty) degree of a fuzzy event E. We have the following property for $N_D(E)$:

$$N_D(E) \geq h \quad \text{if and only if} \quad (D)_{1-h} \subseteq [E]_h, \tag{3}$$

where $(D)_{1-h}$ and $[E]_h$ are strong $(1-h)$-level set of D and h-level set of E, i.e., $(D)_{1-h} = \{r \mid \mu_D(r) > 1-h\}$ and $[E]_h = \{r \mid \mu_E(r) \geq h\}$. This property implies that $N_D(E) \geq h$ means for all r having a membership grade to fuzzy set D more than $1-h$ the condition of fuzzy event E is satisfied to not less than degree h. Therefore, when the variation of r is specified by fuzzy set D, $N_D(E)$ may indicate to what extent fuzzy event E occurs.

Then, based on a necessity fractile model, Problem (1) is formulated as

$$\begin{array}{ll} \text{minimize} & z, \\ \text{subject to} & N_{\boldsymbol{C}}(\{\boldsymbol{\gamma}^{\mathrm{T}}\boldsymbol{x} \leq z\}) \geq h^0, \\ & A\boldsymbol{x} \leq \boldsymbol{b}, \end{array} \tag{4}$$

where $h^0 \in [0,1]$ is a constant necessity level specified by the decision maker.

Define a fuzzy set $Y(\boldsymbol{x})$ by a membership function

$$\mu_{Y(\boldsymbol{x})}(y) = \sup\left\{\mu_{\boldsymbol{C}}(\boldsymbol{c}) \mid \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} = y\right\}, \tag{5}$$

based on the extension principle [17]. We note that we have

$$N_{\boldsymbol{C}}(\{\boldsymbol{\gamma}^{\mathrm{T}}\boldsymbol{x} \leq z\}) = N_{Y(\boldsymbol{x})}((-\infty, z]). \tag{6}$$

In what follows, we review the reduced problems under different assumptions about fuzzy set $\boldsymbol{C}$.

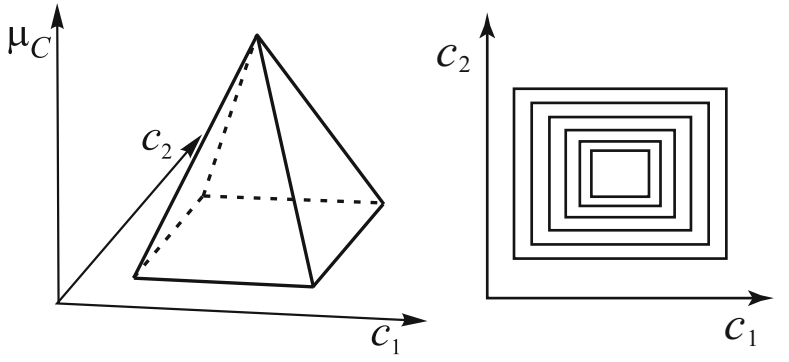
For reference, we describe the necessity measure optimization model corresponding to Problem 1. The model is formulated as

$$\begin{array}{ll} \text{minimize} & N_{\boldsymbol{C}}(\{\boldsymbol{\gamma}^{\mathrm{T}}\boldsymbol{x} \leq z^0\}), \\ \text{subject to} & A\boldsymbol{x} \leq \boldsymbol{b}, \end{array} \tag{7}$$

or equivalently, by using an auxiliary variable $h \in [0,1]$,

$$\begin{array}{ll} \text{minimize} & h, \\ \text{subject to} & N_{\boldsymbol{C}}(\{\boldsymbol{\gamma}^{\mathrm{T}}\boldsymbol{x} \leq z^0\}) \geq h, \\ & A\boldsymbol{x} \leq \boldsymbol{b}, \end{array} \tag{8}$$

where z^0 is a constant target value specified by the decision maker. Comparing Problems (4) and (8), we observe that the constant and variable are



(a) 3D image (b) Level curves

Fig. 1 The membership function of a vector of non-interactive fuzzy numbers

replaced. While the necessity level is a constant but the target value is a variable in Problem (4), the necessity level is a variable but the target value is a constant in Problem (8).

From property (3), necessity measure constraint $N_{\boldsymbol{C}}(\{\boldsymbol{\gamma}^{\mathrm{T}}\boldsymbol{x} \leq z\}) \geq h^0$ represents the safety aspect of the uncertain objective function value. It guarantees that the objective function value is at least z for all possible coefficient vectors $\boldsymbol{c}$ having the membership degree more than $1 - h^0$.

3 Non-interactive Fuzzy Numbers

The information about uncertain vector $\boldsymbol{\gamma}$ is often given in a component-wise way. Namely, we may know a possible range of γ_j as a fuzzy number C_j. In the conventional fuzzy mathematical programming, we assume the non-interaction among uncertain coefficients and model the membership function of $\boldsymbol{C}$ by

$$\mu_{\boldsymbol{C}}(\boldsymbol{c}) = \min\left(\mu_{C_1}(c_1), \mu_{C_2}(c_2), \ldots, \mu_{C_n}(c_n)\right), \tag{9}$$

where $\boldsymbol{c} = (c_1, c_2, \ldots, c_n)^{\mathrm{T}}$ and μ_{C_j} is a membership function of C_j. An example of the membership function of $\boldsymbol{C}$ with $n = 2$ is depicted in Figure 1. In Figure 1, C_j, $j = 1, 2$ are triangular fuzzy numbers. As shown in Figure 1, the level curves are rectangles parallel to coordinate axes. This property does not depend on the fact C_j, $j = 1, 2$ are triangular fuzzy numbers but on the non-interaction between uncertain parameters γ_1 and γ_2.

We further assume that C_j satisfies the following condition:

(A1) $[C_j]_h = \{r \mid \mu_{C_j}(r) \geq h\}$ is a bounded closed interval.

In other words, assumption (A1) means that μ_{C_i} satisfies (a1) upper semi-continuity, (a2) quasi-concavity and (a3) $\lim_{r\to+\infty}\mu_{C_i}(r) = \lim_{r\to-\infty}\mu_{C_i}(r) = 0$.

Under the assumption above, we obtain

$$[Y(\boldsymbol{x})]_h = \sum_{j=1}^{n}[C_j]_h x_j = \sum_{j=1}^{n}\left[C_j^{\mathrm{L}}(h), C_j^{\mathrm{R}}(h)\right] x_j, \tag{10}$$

where we write a bounded closed interval $[C_j]_h$ as $[C_j^{\mathrm{L}}(h), C_j^{\mathrm{R}}(h)]$.

Then Problem (4) is reduced to the following linear programming problem:

$$\begin{aligned} \text{minimize} \quad & \sum_{j=1}^{n} y_j, \\ \text{subject to} \quad & \bar{C}_j^{\mathrm{L}}(1-h^0)x_j \le y_j,\ j=1,2,\ldots,n, \\ & \bar{C}_j^{\mathrm{R}}(1-h^0)x_j \le y_j,\ j=1,2,\ldots,n, \\ & A\boldsymbol{x} \le \boldsymbol{b}, \end{aligned} \tag{11}$$

where we define $\mathrm{cl}(C_j)_h = [\bar{C}_j^{\mathrm{L}}(h), \bar{C}_j^{\mathrm{R}}(h)]$ ("cl" stands for "closure").

When $A\boldsymbol{x} \le \boldsymbol{b}$ includes non-negativity condition $\boldsymbol{x} \ge \boldsymbol{0}$, the problem becomes the following simpler problem:

$$\begin{aligned} \text{minimize} \quad & \sum_{j=1}^{n} \bar{C}_j^{\mathrm{R}}(1-h^0)x_j, \\ \text{subject to} \quad & A\boldsymbol{x} \le \boldsymbol{b}. \end{aligned} \tag{12}$$

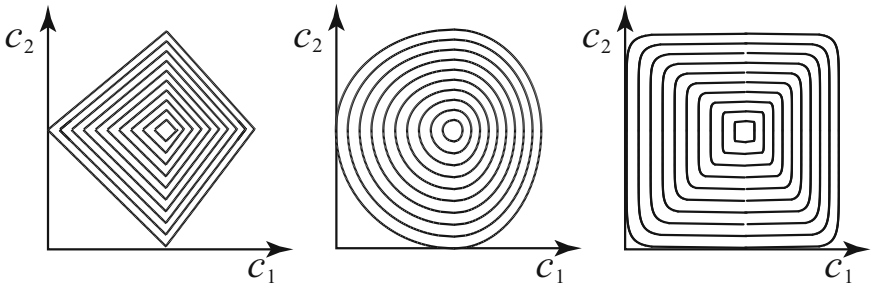
Therefore, the necessity fractile model can be solved easily. It is known that the necessity measure optimization model can be solved by a linear fractional programming technique (see Inuiguchi and Ramík [4]).

4 Weak Independent Fuzzy Numbers

In fuzzy sets and systems, the minimum operation is often replaced with a t-norm [10]. Then we may consider

$$\mu_{\boldsymbol{C}}(\boldsymbol{c}) = T\left(\mu_{C_1}(c_1), \mu_{C_2}(c_2), \ldots, \mu_{C_n}(c_n)\right), \tag{13}$$

where $T : [0,1]\times[0,1] \to [0,1]$ is a t-norm, a two place function satisfying (T1) $T(a,1) = T(1,a) = a$ for all $a \in [0,1]$ (boundary condition), (T2) $T(a,b) = T(b,a)$ for all $a,b \in [0,1]$ (commutativity), (T3) $T(a,T(b,c)) = T(T(a,b),c)$ for all $a,b,c \in [0,1]$ (associativity), and (T4) $T(a,b) \le T(c,d)$ for all $a,b,c,d \in [0,1]$ such that $a \le c$ and $b \le d$. When the joint fuzzy set C of marginal fuzzy sets C_j, $j = 1,2,\ldots,n$ has a membership function



(a) Level curves ($p = 1$) (b) Level curves ($p = 2$) (c) Level curves ($p = 10$)

Fig. 2 The membership functions of weak independent fuzzy numbers

represented by (13), C_j, $j = 1, 2, \ldots, n$ are called weak independent fuzzy numbers.

Rommelfanger and Kresztfalvi [13] proposed to use Yager's parameterized t-norm in order to control the spreads of fuzzy linear function values. The interaction among uncertain parameters is treated indirectly in this approach.

Rommelfanger and Kresztfalvi [13] treated a fuzzy vector $\boldsymbol{C}$ whose membership function $\mu_{\boldsymbol{C}}$ is defined by

$$\mu_{\boldsymbol{C}}(\boldsymbol{c}) = T_p\left(\mu_{C_1}(c_1), \mu_{C_2}(c_2), \ldots, \mu_{C_n}(c_n)\right), \tag{14}$$

where μ_{C_j} is a membership function of a fuzzy number C_j representing a possible range of the coefficient of x_j. T_p is defined by

$$T_p(r_1, r_2, \ldots, r_n) = \max\left[0, 1 - \left(\sum_{j=1}^{n}(1 - r_1)^p\right)^{1/p}\right], \tag{15}$$

and $p \geq 1$. T_p is an extension of Yager's t-norm. The level curves of $\mu_{\boldsymbol{C}}$ defined by (14) with different p values when $n = 2$ are shown in Figure 2. In other words, Rommelfanger and Kresztfalvi treated a kind of weak independent fuzzy numbers in the setting of fuzzy programming problems.

Let us assume C_j is a trapezoidal fuzzy number $(c_j^{\mathrm{L}}, c_j^{\mathrm{R}}, \alpha_j^{\mathrm{L}}, \alpha_j^{\mathrm{R}})$, i.e.,

$$\mu_{C_j}(r) = \begin{cases} \dfrac{c_j^{\mathrm{L}} - r}{\alpha_j^{\mathrm{L}}}, & \text{if } c_j^{\mathrm{L}} - \alpha_j^{\mathrm{L}} \leq r < c_j^{\mathrm{L}}, \\ 1 & \text{if } c_j^{\mathrm{L}} \leq r \leq c_j^{\mathrm{R}}, \\ \dfrac{r - c_j^{\mathrm{R}}}{\alpha_j^{\mathrm{R}}}, & \text{if } c_j^{\mathrm{R}} < r \leq c_j^{\mathrm{R}} - \alpha_j^{\mathrm{R}}, \\ 0, & \text{otherwise}, \end{cases} \tag{16}$$

where we have $c_j^{\mathrm{L}} \leq c_j^{\mathrm{R}}$ and $\alpha_j^{\mathrm{L}}, \alpha_j^{\mathrm{R}} > 0$. Then, Rommelfanger and Kresztfalvi [13] proved that $Y(\boldsymbol{x})$ defined by (5) becomes a trapezoidal fuzzy number $(y^{\mathrm{L}}(\boldsymbol{x}), y^{\mathrm{R}}(\boldsymbol{x}),\ \alpha^{\mathrm{L}}(p, \boldsymbol{x}), \alpha^{\mathrm{R}}(p, \boldsymbol{x}))$ when $A\boldsymbol{x} \leq \boldsymbol{b}$ includes the non-negativity condition $\boldsymbol{x} \geq \boldsymbol{0}$. Here we define

$$y^{\mathrm{L}}(\boldsymbol{x}) = \sum_{j=1}^{n} c_j^{\mathrm{L}} x_j, \tag{17}$$

$$y^{\mathrm{R}}(\boldsymbol{x}) = \sum_{j=1}^{n} c_j^{\mathrm{R}} x_j, \tag{18}$$

$$\alpha^{\mathrm{L}}(p, \boldsymbol{x}) = \left(\sum_{j=1}^{n} (\alpha_j^{\mathrm{L}} x_j)^q \right)^{1/q}, \tag{19}$$

$$\alpha^{\mathrm{R}}(p, \boldsymbol{x}) = \left(\sum_{j=1}^{n} (\alpha_j^{\mathrm{R}} x_j)^q \right)^{1/q}, \tag{20}$$

where $q \geq 1$ is defined from p so as to fulfill

$$\frac{1}{p} + \frac{1}{q} = 1. \tag{21}$$

(21) shows that the larger p is, the larger $1/q$ is. From (19) and (20), this fact implies that the spreads of $Y(\boldsymbol{x})$ enlarges as p increases. Therefore, we can control the uncertainty propagation from $\boldsymbol{C}$ to $Y(\boldsymbol{x})$ by selecting a suitable p.

Applying this result, Problem (4) is reduced to the following programming problem:

$$\begin{aligned} &\text{minimize } y^{\mathrm{R}}(\boldsymbol{x}) + (1-h)\alpha^{\mathrm{R}}(p, \boldsymbol{x}), \\ &\text{subject to } A\boldsymbol{x} \leq \boldsymbol{b}. \end{aligned} \tag{22}$$

This problem is non-linear except $q = 1$ ($p = \infty$, i.e., p is sufficiently large) and $p = 1$ ($q = \infty$, i.e., q is sufficiently large and we can approximate $\alpha^{\mathrm{R}}(p, \boldsymbol{x}) = \max\{\alpha_j^{\mathrm{R}} x_j \mid 1 \leq j \leq n\}$). Because $q \geq 1$, this problem is a convex programming problem so that it can be solved by a gradient method.

As shown in Figure 2, p changes the interaction among uncertain coefficients γ_j, $j = 1, 2, \ldots, n$. It is difficult to understand the meaning of p other than we can control the spread of $Y(\boldsymbol{x})$ by p. Therefore, the selection of the parameter p would be a difficult task.

The necessity measure optimization model can be formulated as a fractional programming problem which is not always linear due to $\alpha^{\mathrm{R}}(p, \boldsymbol{x})$.

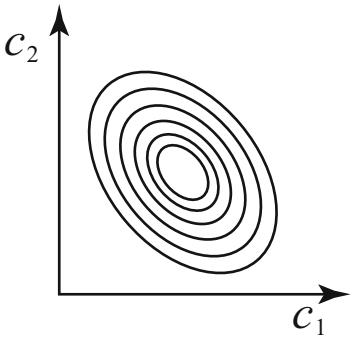


Fig. 3 The level curves of a quadratic membership function

5 Fuzzy Vector with a Quadratic Membership Function

Inuiguchi and Sakawa [6] treated a fuzzy linear programming with a quadratic membership function. A quadratic membership function can be considered as a counterpart of a multivariate normal distribution. It is defined by center values (corresponding to mean values) and a symmetrical positive definite matrix (corresponding to variance matrix). Therefore, by a quadratic membership function, we may express correlations of all pairs of uncertain parameters γ_j, $j = 1, 2, \ldots, n$. Because of the similarity to a multivariate normal distribution, Inuiguchi and Sakawa [6] succeeded to show the equivalence between special models of stochastic linear programming problem and fuzzy linear programming problem.

In this approach, fuzzy set $\boldsymbol{C}$ is defined by the following membership function:

$$\mu_C(\boldsymbol{c}) = L((\boldsymbol{c} - \boldsymbol{d})^{\mathrm{T}} U^{-1} (\boldsymbol{c} - \boldsymbol{d})), \tag{23}$$

where $\boldsymbol{d} = (d_1, \ldots, d_n)^{\mathrm{T}}$ is a constant vector, U is an $n \times n$ symmetrical positive definite matrix representing the interactions among objective coefficients. U^{-1} is the inverse matrix of U. $L : [0, +\infty) \mapsto [0, 1]$ is a reference function which is an upper semi-continuous and non-increasing function such that $L(0) = 1$ and $\lim_{r \to +\infty} L(r) = 0$.

The level curves of a quadratic membership function with $n = 2$ are depicted in Figure 3. The level curves of quadratic membership functions are ellipsoids as those of multivariate normal distributions are.

Extending Tanaka and Ishibuchi's result [16], we obtain the membership function of $Y(\boldsymbol{x})$ as

$$\mu_{Y(\boldsymbol{x})}(y) = L\left(\frac{(y - \boldsymbol{d}^t \boldsymbol{x})^2}{\boldsymbol{x}^{\mathrm{T}} U \boldsymbol{x}}\right). \tag{24}$$

From (24), we obtain

$$N_{Y(\boldsymbol{x})}((-\infty, z]) \geq h^0 \Leftrightarrow \boldsymbol{d}^{\mathrm{T}}\boldsymbol{x} + \sqrt{L^*(1-h^0)\boldsymbol{x}^t U\boldsymbol{x}} \leq z, \tag{25}$$

where $L^* : [0,1] \to [0,+\infty) \cup \{-\infty\}$ is defined by

$$L^*(h) = \begin{cases} \sup\{r \mid L(r) > h,\ r \geq 0\}, & \text{if } h < 1, \\ -\infty, & \text{if } h = 1, \end{cases} \tag{26}$$

Thus, when $h^0 > 0$, Problem (4) is reduced to the following nonlinear programming problem:

$$\begin{aligned} &\text{minimize } \boldsymbol{d}^{\mathrm{T}}\boldsymbol{x} + \sqrt{L^*(1-h^0)\boldsymbol{x}^t U\boldsymbol{x}}, \\ &\text{subject to } A\boldsymbol{x} \leq \boldsymbol{b}. \end{aligned} \tag{27}$$

This problem is a convex programming problem and can be solved by a method developed in stochastic programming problem [15].

The necessity measure optimization models for problems with quadratic membership functions are reduced to fractional programming problems as minimum-risk models in stochastic programming problems with multivariate normal distributions are (see Stancu-Minasian [15]). The solution procedures for the minimum-risk models can be applied to the necessity measure optimization models.

6 Scenario Decomposed Fuzzy Numbers

Inuiguchi and Tanino [8] proposed scenario decomposed fuzzy numbers. In their approach, the interaction between uncertain parameters are expressed by fuzzy if-then rules.

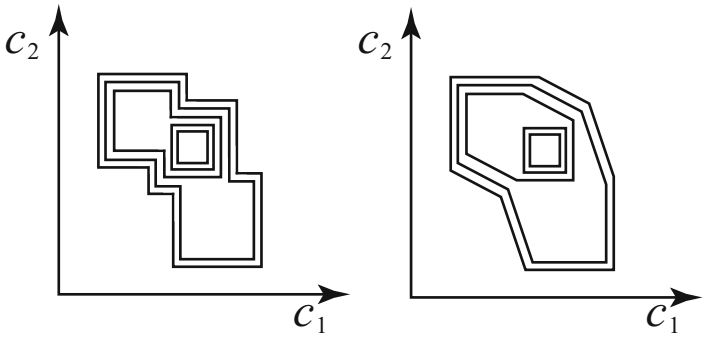
We may have a vague knowledge about the range of $\boldsymbol{\gamma}$ as the following k fuzzy if-then rules:

$$\text{if } s = s_k \text{ then } \boldsymbol{\gamma} \in \boldsymbol{C}^k,\ k = 1, 2, \ldots, u, \tag{28}$$

where s is a variable taking a value from $\{s_1, s_2, \ldots, s_u\}$. s is called a *scenario variable*. $\boldsymbol{C}^k = (C_1^k, C_2^k, \ldots, C_n^k)^{\mathrm{T}}$ is a vector of non-interactive fuzzy numbers. Namely, $\boldsymbol{C}^k$ has a membership function,

$$\mu_{\boldsymbol{C}^k}(\boldsymbol{c}) = \min\left(\mu_{C_1^k}(c_1), \mu_{C_2^k}(c_2), \ldots, \mu_{C_n^k}(c_n)\right), \tag{29}$$

and C_j^k is a fuzzy number such that $[C_j^k]_h = \{r \mid \mu_{C_j^k}(r) \geq h\}$ is a bounded closed interval, where $\mu_{C_j^k}$ is a membership function of a fuzzy number C_j^k.



(a) discrete scenario variable (b) continuous scenario variable

Fig. 4 Lever curves of the membership functions of scenario decomposed fuzzy numbers

For example, we may have knowledge,

– if economy s is good then share prices $\boldsymbol{\gamma} \in \boldsymbol{C}^1$,
– if economy s is normal then share prices $\boldsymbol{\gamma} \in \boldsymbol{C}^2$,
– if economy s is bad then share prices $\boldsymbol{\gamma} \in \boldsymbol{C}^3$.

The body of rules (28) shows a fuzzy relation between scenario s_k and possible range of uncertain vector $\boldsymbol{\gamma}$.

When we obtain our estimation or information about scenario variable by a fuzzy set S showing a possible realizations of s under if-then knowledge (28), based on the fuzzy set induction from S through R [17], the estimated fuzzy set $\boldsymbol{C}$ is obtained as

$$\mu_{\boldsymbol{C}}(\boldsymbol{c}) = \max_{k=1,2,\ldots,u} \min\left(\mu_S(s_k), \mu_{\boldsymbol{C}^k}(\boldsymbol{c}),\right) \tag{30}$$

where μ_S is a membership function of S.

Level curves of the membership function of a scenario decomposed fuzzy numbers $\boldsymbol{C}$ defined by (30) with $n = 2$ is depicted in Figure 4(a). In (28), we consider a discrete scenario variable s but Inuiguchi and Tanino [8] considered a continuous scenario variable. In the continuous scenario variable case, the knowledge can be represented by a set of fuzzy rules. Level curves of the membership function of scenario decomposed fuzzy numbers $\boldsymbol{C}$ with a continuous scenario variable when $n = 2$ is illustrated in Figure 4(b). However, in this paper, we concentrate on the discrete scenario variable case because of its simplicity. However, even in the continuous scenario variable case, we obtained similar results as shown in Inuiguchi and Tanino [8].

The pair $(S, \text{knowledge}(28))$ can be called a scenario decomposed fuzzy numbers.

Let $Y^k(\boldsymbol{x})$ be a fuzzy number defined by the following membership function:

$$\mu_{Y^k(\boldsymbol{x})}(y) = \sup\left\{\mu_{\boldsymbol{C}^k}(\boldsymbol{c}) \mid \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} = y\right\}. \tag{31}$$

Then we obtain $Y(\boldsymbol{x})$ by the following membership function:

$$\mu_{Y(\boldsymbol{x})}(y) = \max_{k=1,2,\ldots,u} \min\left(\mu_S(s_k), \mu_{Y^k(\boldsymbol{x})}(y).\right) \tag{32}$$

Since we have

$$\begin{aligned} &N_{Y(\boldsymbol{x})}((-\infty, z]) \geq h^0 \Leftrightarrow \mathrm{cl}(Y(\boldsymbol{x}))_{1-h^0} \subseteq (-\infty, z] \\ &\Leftrightarrow \mathrm{cl}(Y^k(\boldsymbol{x}))_{1-h^0} \subseteq (-\infty, z],\ \forall k \text{ such that } \mu_S(s_k) > 1 - h^0. \end{aligned} \tag{33}$$

Problem (4) is reduced to the following linear programming problem:

$$\begin{aligned} \text{minimize} \quad & z, \\ \text{subject to} \quad & \bar{C}^{\mathrm{L}}_{jk}(1-h^0)x_j \leq y^k_j,\ j = 1, 2, \ldots, n, \\ & \bar{C}^{\mathrm{R}}_{jk}(1-h^0)x_j \leq y^k_j,\ j = 1, 2, \ldots, n, \\ & \sum_{j=1}^{n} y^k_j \leq z,\ \forall k \text{ such that } \mu_S(s_k) > 1 - h^0, \\ & A\boldsymbol{x} \leq \boldsymbol{b}, \end{aligned} \tag{34}$$

where we define $\mathrm{cl}(C^k_j)_h = [\bar{C}^{\mathrm{L}}_{jk}(h), \bar{C}^{\mathrm{R}}_{jk}(h)]$.

When $A\boldsymbol{x} \leq \boldsymbol{b}$ includes non-negativity condition $\boldsymbol{x} \geq \boldsymbol{0}$, the problem becomes the following simpler problem:

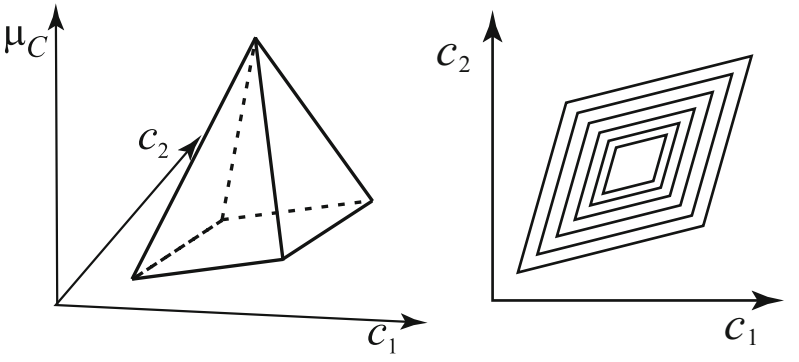
$$\begin{aligned} \text{minimize} \quad & z, \\ \text{subject to} \quad & \sum_{j=1}^{n} \bar{C}^{\mathrm{R}}_{jk}(1-h^0)x_j \leq z,\ \forall k \text{ such that } \mu_S(s_k) > 1 - h^0, \\ & A\boldsymbol{x} \leq \boldsymbol{b}. \end{aligned} \tag{35}$$

The necessity measure optimization model for problems with scenario decomposed fuzzy numbers can be formulated similarly and solved by a bisection method together with a simplex method.

7 Oblique Fuzzy Vector

In the real world, we may have vague knowledge about the sums of uncertain values and/or the differences of two uncertain values, e.g., the sum of γ_1, γ_2 and γ_3 is about 5, the difference between γ_4 and γ_5 is approximately 3, and so on.

Inuiguchi, Ramík and Tanino [5] proposed oblique fuzzy vectors. An oblique fuzzy vector can express n independent pieces of vague knowledge about the linear function values of uncertain values. A non-singular matrix shows the interaction among uncertain parameters in an oblique fuzzy vector as a



(a) 3D image (b) Level curves

Fig. 5 The membership function of an oblique fuzzy vector

covariance matrix shows in a multivariate normal distribution. It is shown that linear function values of oblique fuzzy vectors can be calculated easily.

An oblique fuzzy vector $\boldsymbol{C}$ is defined by the following membership function,

$$\mu_{\boldsymbol{C}}(\boldsymbol{c}) = \min_{j=1,2,\ldots,n} \mu_{B_j}(\boldsymbol{d}_j^{\mathrm{T}}\boldsymbol{c}), \tag{36}$$

where μ_{B_j} is a membership function of an L-L fuzzy number $B_j = (b_j^{\mathrm{L}}, b_j^{\mathrm{R}}, \beta_j^{\mathrm{L}}, \beta_j^{\mathrm{R}})_{LL}$ and $\boldsymbol{d}_j$, $j = 1, 2, \ldots, n$ are vectors such that $D = (\boldsymbol{d}_1, \boldsymbol{d}_2, \ldots, \boldsymbol{d}_n)^{\mathrm{T}}$ be a non-singular real-valued $n \times n$ matrix. An L-L fuzzy number $B_j = (b_j^{\mathrm{L}}, b_j^{\mathrm{R}}, \beta_j^{\mathrm{L}}, \beta_j^{\mathrm{R}})_{LL}$ can be characterized by the following membership function:

$$\mu_{B_j}(r) = \begin{cases} L\left(\dfrac{b_j^{\mathrm{L}} - r}{\beta_j^{\mathrm{L}}}\right) & \text{if } r < b_j^{\mathrm{L}}, \\ 1 & \text{if } b_j^{\mathrm{L}} \le r \le b_j^{\mathrm{R}}, \\ L\left(\dfrac{r - b_j^{\mathrm{R}}}{\beta_j^{\mathrm{R}}}\right) & \text{if } r > b_j^{\mathrm{R}}, \end{cases} \tag{37}$$

where we assume $b_j^{\mathrm{L}} \le b_j^{\mathrm{R}}$, $\beta_j^{\mathrm{L}} > 0$ and $\beta_j^{\mathrm{R}} > 0$. $L : [0, +\infty) \to [0, 1]$ is a reference function which is an upper semi-continuous and non-increasing function such that $L(0) = 1$ and $\lim_{r \to +\infty} L(r) = 0$.

Namely, an oblique fuzzy vector can be obtained from n pieces of knowledge '$\boldsymbol{d}_j^{\mathrm{T}}\boldsymbol{c}$ takes a value in a fuzzy number B_j', $j = 1, 2, \ldots, n$, where $\boldsymbol{d}_j$, $j = 1, 2, \ldots, n$ should be linearly independent.

An example of an oblique fuzzy vector when $n = 2$ is given in Figure 5. Unlike non-interactive fuzzy numbers, the level curves of an oblique fuzzy vector are neither always rectangle nor parallel to coordinate axes.

Utilizing the result by Inuiguchi et al. [5], we obtain

$$\mathrm{cl}(Y(\boldsymbol{x}))_h = \Bigg[\sum_{j:k_j(\boldsymbol{x})\geq 0} \bar{b}_j^{\mathrm{L}}(h)k_j(\boldsymbol{x}) + \sum_{j:k_j(\boldsymbol{x})<0} \bar{b}_j^{\mathrm{R}}(h)k_j(\boldsymbol{x}), \\ \sum_{j:k_j(\boldsymbol{x})\geq 0} \bar{b}_j^{\mathrm{R}}(h)k_j(\boldsymbol{x}) + \sum_{j:k_j(\boldsymbol{x})<0} \bar{b}_j^{\mathrm{L}}(h)k_j(\boldsymbol{x}) \Bigg], \quad (38)$$

where $k_j(\boldsymbol{x})$, $j = 1, 2, \ldots, n$ are defined as follows with d_{ij}^*, the (i, j) component of D^{-1};

$$k_j(\boldsymbol{x}) = \sum_{i=1}^{n} d_{ij}^* x_i. \quad (39)$$

$\bar{b}_j^{\mathrm{L}}(h)$ and $\bar{b}_j^{\mathrm{R}}(h)$ are defined by

$$\bar{b}_j^{\mathrm{L}}(h) = b_j^{\mathrm{L}} - \beta_j^{\mathrm{L}} L^*(h), \quad (40)$$

$$\bar{b}_j^{\mathrm{R}}(h) = b_j^{\mathrm{R}} - \beta_j^{\mathrm{R}} L^*(h), \quad (41)$$

where L^* is defined by (26).

This result implies that the linear function values of an oblique fuzzy vector is an L-L fuzzy number (see Inuiguchi et al [5]).

Applying the result above and the first line of (33), Problem (4) can be reduced to the following problem:

$$\begin{aligned} \text{minimize} \quad & \sum_{j:k_j(\boldsymbol{x})\geq 0} \bar{b}_j^{\mathrm{R}}(1-h^0)k_j(\boldsymbol{x}) + \sum_{j:k_j(\boldsymbol{x})<0} \bar{b}_j^{\mathrm{L}}(1-h^0)k_j(\boldsymbol{x}), \\ \text{subject to} \quad & A\boldsymbol{x} \leq \boldsymbol{b}, \\ & k_j(\boldsymbol{x}) = \sum_{i=1}^{n} d_{ij}^* x_i. \end{aligned} \quad (42)$$

We have

$$\boldsymbol{k}(\boldsymbol{x}) = (k_1(\boldsymbol{x}), k_2(\boldsymbol{x}), \ldots, k_n(\boldsymbol{x}))^{\mathrm{T}} = D^{-\mathrm{T}}\boldsymbol{x}, \quad (43)$$

where $D^{-\mathrm{T}} = D^{-1^{\mathrm{T}}} = D^{\mathrm{T}^{-1}}$. From this fact, we introduce variable vectors $\boldsymbol{y}^+ = (y_1^+, y_2^+, \ldots, y_n^+)^{\mathrm{T}}$ and $\boldsymbol{y}^- = (y_1^-, y_2^-, \ldots, y_n^-)^{\mathrm{T}}$ such that

$$D^{-\mathrm{T}}\boldsymbol{x} = \boldsymbol{y}^+ - \boldsymbol{y}^-, \ \boldsymbol{y}^{+\mathrm{T}}\boldsymbol{y}^- = 0, \ \boldsymbol{y}^+ \geq \boldsymbol{0}, \ \boldsymbol{y}^- \geq \boldsymbol{0}. \quad (44)$$

From (43) and (44), we have $k_j(\boldsymbol{x}) = y_j^+$ if $k_j(\boldsymbol{x}) \geq 0$ and $k_j(\boldsymbol{x}) = -y_j^-$ if $k_j(\boldsymbol{x}) < 0$. Moreover, from the first equation of (44), we have $\boldsymbol{x} = D^{\mathrm{T}}(\boldsymbol{y}^+ - \boldsymbol{y}^-)$. Introducing those, we can prove Problem (42) is reduced to the following linear programming problem:

$$\begin{aligned} \text{minimize} \quad & \sum_{j=1}^{n} \bar{b}_j^{\mathrm{R}}(1-h^0)y_j^+ - \sum_{j=1}^{n} \bar{b}_j^{\mathrm{L}}(1-h^0)y_j^-, \\ \text{subject to} \quad & A\boldsymbol{x} \le \boldsymbol{b}, \\ & \boldsymbol{x} = D^{\mathrm{T}}(\boldsymbol{y}^+ - \boldsymbol{y}^-),\ \boldsymbol{y}^+ \ge \boldsymbol{0},\ \boldsymbol{y}^- \ge \boldsymbol{0}. \end{aligned} \tag{45}$$

Note that a complementary condition $(\boldsymbol{y}^+)^{\mathrm{T}}\boldsymbol{y}^- = 0$ can be omitted in Problem (45) because we obtain a solution satisfying this condition easily from any optimal solution of Problem (45) without change of $\boldsymbol{x}$. It is shown that Bender's decomposition method can be applied to Problem (45) (see Inuiguchi et al. [5]).

The necessity measure optimization model for a problem with an oblique fuzzy vector is studied by Inuiguchi [2]. The model can be also reduced to a linear fractional programming problem and solved by Bender's decomposition method.

8 Fuzzy Polytope

By oblique fuzzy vector, we can express n independent pieces of vague knowledge about linear function values of uncertain parameters. However, in the real world, we may have more than n pieces of vague knowledge including vague knowledge about the ratio between two uncertain parameters. The ratio between two uncertain parameters cannot be expressed as a linear function of uncertain parameters. Therefore such a body of vague knowledge cannot be expressed well by an oblique fuzzy vector.

Inuiguchi and Tanino [9] introduced a fuzzy polytope to fuzzy linear programming problems. A fuzzy polytope can express more than n pieces of vague knowledge about linear fractional function values of uncertain parameters. Then the oblique fuzzy vector is a special case of the fuzzy polytope.

When $\boldsymbol{C}$ is a fuzzy polytope, its membership function is expressed as

$$\mu_{\boldsymbol{C}}(\boldsymbol{c}) = \min_{k=1,2,\ldots,v} L_k\left(\frac{\dfrac{\boldsymbol{w}_k^{\mathrm{T}}\boldsymbol{c}+w_{0k}}{\boldsymbol{d}_k^{\mathrm{T}}\boldsymbol{c}+d_{0k}} - \bar{q}_k}{\alpha_k}\right), \tag{46}$$

where $L_k : \mathbf{R} \to [0,1]$, $k = 1,2,\ldots,v$ are reference functions, i.e., upper semi-continuous non-increasing functions such that $L_k(0) = 1$ and $\lim_{r\to+\infty} L_k(r) = 0$. $\bar{q}_k$ are the most plausible value for the k-th linear fractional function value $(\boldsymbol{w}_k^{\mathrm{T}}\boldsymbol{\gamma} + w_{0k})/(\boldsymbol{d}_k^{\mathrm{T}}\boldsymbol{\gamma} + d_{0k})$. α_k shows the spread, i.e., to what extent the linear fractional function value $(\boldsymbol{w}_k^{\mathrm{T}}\boldsymbol{\gamma} + w_{0k})/(\boldsymbol{d}_k^{\mathrm{T}}\boldsymbol{\gamma} + d_{0k})$ possibly exceeds $\bar{q}_k$. When we know the maximum possible shortage of $(\boldsymbol{w}_k^{\mathrm{T}}\boldsymbol{\gamma}+w_{0k})/(\boldsymbol{d}_k^{\mathrm{T}}\boldsymbol{\gamma}+d_{0k})$ from $\bar{q}_k$, by multiplying $\boldsymbol{w}_k$, w_{0k} and $\bar{q}_k$ by (-1), we

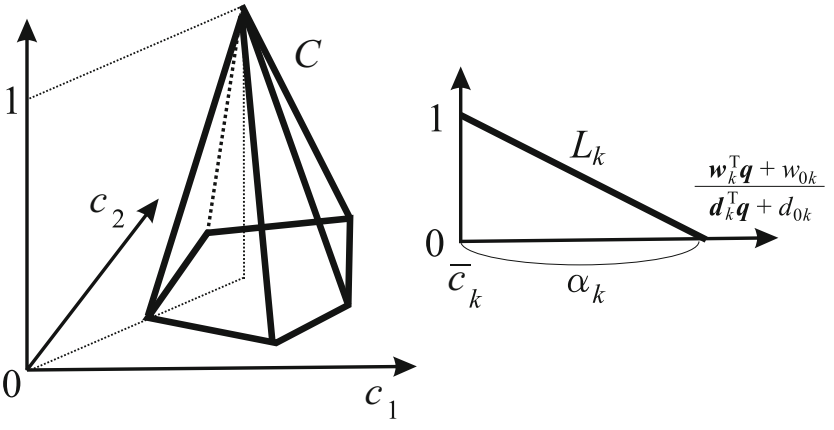


Fig. 6 The membership function of a fuzzy polytope

can treat it as we know the maximum possible excess of $(-\boldsymbol{w}_k^{\mathrm{T}}\boldsymbol{\gamma}-w_{0k})/(\boldsymbol{d}_k^{\mathrm{T}}\boldsymbol{\gamma}+d_{0k})$ from $-\bar{q}_k$. The fuzzy set $\boldsymbol{C}$ is assumed to be bounded, i.e., h-level sets $[\boldsymbol{C}]_h = \{\boldsymbol{c} \mid \mu_{\boldsymbol{C}}(\boldsymbol{c}) \geq h\}$ for all $h \in (0,1]$ are bounded. Moreover, without loss of generality, we assume that $\boldsymbol{d}_k^{\mathrm{T}}\boldsymbol{c} + d_{k0} > 0$ for all possible $\boldsymbol{c}$. Let $L_k^*(h) = \sup\{r \mid L_k(r) > h\}$ for $h \in [0,1)$ and $L_k^*(h) = -\infty$ for $h = 1$.

Since a linear fractional function includes a sum, a difference a linear function, a ratio, a fuzzy polytope is useful when we know possible ranges of a sum of uncertain variables, a difference between two uncertain variables, a linear function values of uncertain variables and a ratio between two uncertain variables. The membership function of a fuzzy polytope when $n = 2$ is depicted in Figure 6.

Because of (46), we have

$$\begin{aligned}\mathrm{cl}(\boldsymbol{C})_h &= \{\boldsymbol{c} \mid \boldsymbol{w}_k^{\mathrm{T}}\boldsymbol{c} + w_{0k} \leq (\bar{q}_k + \alpha_k L_k^*(h))(\boldsymbol{d}_k^{\mathrm{T}}\boldsymbol{c} + d_{0k}),\ k = 1,2,\ldots,v\} \\ &= \{\boldsymbol{c} \mid \boldsymbol{w}\boldsymbol{d}_k^*(h)^{\mathrm{T}}\boldsymbol{c} \leq -wd_{0k}^*(h),\ k = 1,2,\ldots,v\},\end{aligned} \tag{47}$$

where $\boldsymbol{w}\boldsymbol{d}_k^*(h) = \boldsymbol{w}_k - (\bar{q}_k + \alpha_k L_k^*(h))\boldsymbol{d}_k$ and $wd_{0k}^*(h) = w_{0k} - (\bar{q}_k + \alpha_k L_k^*(h))d_{0k}$. Since $[\boldsymbol{C}]_h \subseteq \mathbf{R}^n$ is bounded, from (47), we know that $v > n$ and that $[\boldsymbol{C}]_h$ and $\mathrm{cl}(\boldsymbol{C})_h$ are polytopes for all $h \in [0,1)$.

From the first line of (33), Problem (4) is reduced to the following semi-infinite programming problem:

$$\begin{array}{ll}\text{minimize} & z, \\ \text{subject to} & \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} \leq z,\ \forall \boldsymbol{c} \in \mathrm{cl}(\boldsymbol{C})_{1-h^0}, \\ & A\boldsymbol{x} \leq \boldsymbol{b}.\end{array} \tag{48}$$

Together with (47), Problem (48) can be solved by the following relaxation procedure.

Solution Algorithm for Problem (48)

Step 1. Select $\boldsymbol{x}^0$ satisfying $A\boldsymbol{x}^0 \leq \boldsymbol{b}$. Let $z^0 = -\infty$ and $l = 0$.
Step 2. Solve a linear programming problem

$$\begin{array}{ll} \text{maximize} & {\boldsymbol{x}^0}^{\mathrm{T}}\boldsymbol{c}, \\ \text{subject to} & \boldsymbol{w}\boldsymbol{d}_k^*(h)^{\mathrm{T}}\boldsymbol{c} \leq -wd_{0k}^*(h),\ k = 1, 2, \ldots, v. \end{array} \tag{49}$$

Let $\hat{\boldsymbol{c}}$ be an obtained optimal solution to Problem (49).
Step 3. If $\hat{\boldsymbol{c}}^{\mathrm{T}}\boldsymbol{x}^0 > z^0$ then update $l = l + 1$ and let $\boldsymbol{c}_l = \hat{\boldsymbol{c}}$. Otherwise, we terminate the algorithm and obtain an optimal solution $\boldsymbol{x}^0$ to Problem (48).
Step 4. Solve a linear programming problem,

$$\begin{array}{ll} \text{minimize} & z, \\ \text{subject to} & \boldsymbol{c}_w^{\mathrm{T}}\boldsymbol{x} \leq z,\ w = 1, 2, \ldots, l, \\ & A\boldsymbol{x} \leq \boldsymbol{b}. \end{array} \tag{50}$$

Let $({\boldsymbol{x}^0}^{\mathrm{T}}, z^0)^{\mathrm{T}}$ be an obtained optimal solution. Return to Step 2.

In the algorithm described above, we solve two kinds of linear programming problems. Therefore, we can solve Problem (48) using linear programming techniques only.

The necessity measure optimization model for a problem with a fuzzy polytope is investigated by Inuiguchi [3]. Different from the necessity fractile model, the necessity measure optimization model cannot be solved only by a relaxation procedure. A solution algorithm based on a bisection method and a relaxation procedure is proposed by Inuiguchi [3].

9 Concluding Remarks

We have described five kinds of treatments of interaction among fuzzy parameters. The first two approaches cannot preserve the linearity. However, the reduced problems are convex programming problems so that they can be solved rather easily by nonlinear programming techniques. On the other hand, the last three approaches preserve the linearity at least to some extent so that reduced problems can be solved by linear programming techniques.

Those five approaches treat special cases of general interaction. However, general interaction would not be easily treated. This is the same in stochastic programming.

While many of the five approaches have been applied to more general linear programming problems with fuzzy coefficients, the interaction among fuzzy parameters has not yet investigated thoroughly. Other treatments of interaction, identification of tractable interaction from given data and applications of the tractable interactions described in this paper would be future topics.

Acknowledgements. The author acknowledges that this work has been partially supported by the Grant-in-Aid for Scientific Research (B) NO.17310098.

References

1. Hisdal, E.: Conditional possibilities independence and noninteraction. Fuzzy Sets and Systems 1(4), 283–297 (1978)
2. Inuiguchi, M.: A necessity measure optimization approach to linear programming problems with oblique fuzzy vectors. Kybernetika 42(4), 441–452 (2006)
3. Inuiguchi, M.: Necessity measure optimization in linear programming problems with fuzzy polytopes. Fuzzy Sets and Systems 158, 1882–1891 (2007)
4. Inuiguchi, M., Ramík, J.: Possibilistic linear programming: A brief review of fuzzy mathematical programming and a comparison with stochastic programming in portfolio selection problem. Fuzzy Sets and Systems 111(1), 3–28 (2000)
5. Inuiguchi, M., Ramík, J., Tanino, T.: Oblique fuzzy vectors and their use in possibilistic linear programming. Fuzzy Sets and Systems 137(1), 123–150 (2003)
6. Inuiguchi, M., Sakawa, M.: A possibilistic linear program is equivalent to a stochastic linear program in a special case. Fuzzy Sets and Systems 76, 309–318 (1995)
7. Inuiguchi, M., Tanino, T.: Portfolio selection under independent possibilistic information. Fuzzy Sets and Systems 115(1), 83–92 (2000)
8. Inuiguchi, M., Tanino, T.: Possibilistic linear programming with fuzzy if-then rule coefficients. Fuzzy Optimization and Decision Making 1(1), 65–91 (2002)
9. Inuiguchi, M., Tanino, T.: Fuzzy linear programming with interactive uncertain parameters. Reliable Computing 10(5), 357–367 (2004)
10. Klement, E.P., Mesiar, R., Pap, E.: Triangular Norms. Kluwer Academic Publishers, Dordrecht (2000)
11. Rommelfanger, H.: Fuzzy linear programming and applications. European Journal of Operational Research 92(3), 512–527 (1996)
12. Rommelfanger, H.: The advantages of fuzzy optimization models in practical use. Fuzzy Optimization and Decision Making 3(4), 295–309 (2004)
13. Rommelfanger, H., Kresztfalvi, T.: Multicriteria fuzzy optimization based on Yager's parameterized t-norm. Foundation of Computing and Decision Sciences 16(2), 99–110 (1991)
14. Słowiński, R., Teghem, J. (eds.): Stochastic versus Fuzzy Approaches to Multiobjective Mathematical Programming under Uncertainty. Kluwer Academic Publishers, Dordrecht (1990)
15. Stancu-Minasian, I.M.: Stochastic Programming with Multiple Objective Functions. D. Reidel Publishing Company, Dordrecht (1984)
16. Tanaka, H., Ishibuchi, H.: Identification of possibilistic linear systems by quadratic membership function. Fuzzy Sets and Systems 41, 145–160 (1991)
17. Zadeh, L.A.: The concept of a linguistic variable and its application to approximate reasoning, Part 1, 2, and 3. Information Sciences 8, 199–249; 8, 301–357; 9, 43–80 (1975)

Possibilistic Optimization Tasks with Mutually T-Related Parameters: Solution Methods and Comparative Analysis

Alexander Yazenin and Ilia Soldatenko

Abstract. The problems of possibilistic linear programming are studied in the article. Unlike in other known related publications, t-norms are used to describe the interaction (relatedness) of fuzzy parameters. Solution methods are proposed, models of possibilistic optimization are compared for different t-norms.

Introduction

It seems to us that systematic research of fuzzy programming problems in the context of possibility theory axiomatics was initiated in [5, 6, 7, 24]. The article by Stephen Nahmias [21] which offered the most common at that time possibilistic model of uncertainty acted like an incitement for the authors of these researches. In particular the above mentioned article contained the concept of fuzzy (possibilistic) variable. Later using fuzzy variables and corresponding calculus of possibilities, as it seems to us now, became the most appropriate instrument for modelling fuzzy parameters of optimization tasks.

One should mention that relatedness of fuzzy parameters in possibilistic optimization problems in those works was based generally on standard conjunction operation that was widely-spread in fuzzy logic. However this way of fuzzy information aggregation and modelling of fuzzy parameters relatedness is not quite appropriate in the series of cases. For example, performing of additive operations with this way of information aggregation leads to linear growth of result's fuzziness, which is not always reasonable.

Alexander Yazenin
Tver State University, Department of Information Technologies, Tver, Russia
e-mail: Alexander.Yazenin@tversu.ru

Ilia Soldatenko
Tver State University, Department of Information Technologies, Tver, Russia
e-mail: soldis@tversu.ru

W.A. Lodwick & J. Kacprzyk (Eds.): Fuzzy Optimization, STUDFUZZ 254, pp. 163–192.
springerlink.com

At present, methods of aggregating possibilistic information based on the technique of triangular norms (t-norms) are being developed. They provide more flexibility in controlling fuzziness in decision-making. Such approach seems to be presently most general. In particular, the operation of conjunction (minimum) acts as one of the types of t-norms.

In this work, we develop this scientific line of investigation with regard to possibilistic linear programming tasks. For the case of T_W-norm, we study two models of possibilistic linear programming problems, which were first introduced in [5, 6, 7, 24]. We propose methods, which are combinations of the indirect method and genetic algorithm of optimization, to solve it. Afterwards we compare models of possibilistic optimization for T_W- and T_M-norms.

The article has the following structure.

In the first section necessary definitions and mathematical apparatus from the possibility theory are introduced. In the second section notion of mutual T-relatedness of possibilistic parameters based on t-norms is defined. It generalizes the notion of fuzzy parameters unrelatedness introduced by Stephen Nahmias in [21] and defined more accurately as min-relatedness by Rao [23].

In the third section behaviour of weighted sum of possibilistic variables in which summands' interaction is based in weak t-norm T_W is investigated. Possibility distribution of weighted sum in case when fuzzy operands are modelled with the help of (L,R)-type distributions is proposed.

In the fourth section of the article two models of possibilistic linear optimization problems are considered: maximax models, known as optimistic model of decision making, and model of mamimization of possibility of fuzzy goal achievement [5, 6, 7, 24]. In this section theorems that allow to build equivalent deterministic analogs of posiibilistic programming tasks with mutually T_W-related parameters are proved. The acquired equivalent deterministic analogues in the case of possibility measure correspond to tasks of non-convex mathematical programming, as opposed to equivalent deterministic analogues in the case of necessity measure. In order to solve these deterministic analogs in the fifth section genetic optimization algorithm is specified.

In the sixth section of the article comparative analysis of fuzzy optimization tasks subject to different trinagular norms is made. It is shown that feasible region of T_W model is a subset of feasible region of T_M model. In this section numerical implementation of modelled problems with different t-norms, that describe relatedness of the tasks' parameters, is also given.

Discussion of acquired results and directions of further investigations form the conclusion of the article.

1 Necessary Mathematical Apparatus

Let us introduce some necessary definitions and concepts from the theory of possibility [1, 4, 21, 22, 23, 26, 27].

Suppose Γ is the model space, $\gamma \in \Gamma$ are its elements, $\mathbb{P}(\Gamma)$ is the set of all subsets of the set Γ, and $\mathbb{E}^1$ is the number scale.

Definition 1. *A possibilistic measure $\pi : \Gamma \to \mathbb{E}^1$ is a set function such that:*

1. $\pi\{\varnothing\} = 0,\ \pi\{\Gamma\} = 1,$
2. $\pi\left\{\bigcup\limits_{i\in I} A_i\right\} = \sup\limits_{i\in I} \pi\{A_i\},\ \forall A_i \in \mathbb{P}(\Gamma),\ \forall I.$

Definition 2. *The triple $(\Gamma, \mathbb{P}(\Gamma), \pi)$ forms the possibilistic space.*

Definition 3. *A necessity measure $\nu : \Gamma \to \mathbb{E}^1$ is a set function such that:*

1. $\nu\{\varnothing\} = 0,\ \nu\{\Gamma\} = 1,$
2. $\nu\left\{\bigcap\limits_{i\in I} A_i\right\} = \inf\limits_{i\in I} \nu\{A_i\},\ \forall A_i \in \mathbb{P}(\Gamma),\ \forall I.$

Let us note some important properties of possibility and necessity measures.

1. π and ν are special cases of uncertainty measure and thus possess all its characteristics (boundedness, monotonicity).
2. $\nu\{A\} = 1 - \pi\{A^C\}$, where A^C — complement of set A. This property shows duality of the measures.
3. $\max\{\pi\{A\}, \pi\{A^C\}\} = 1,\ \min\{\nu\{A\}, \nu\{A^C\}\} = 0.$
4. $\pi\{A\} + \pi\{A^C\} \geq 1,\ \nu\{A\} + \nu\{A^C\} \leq 1.$
5. $\pi\{A\} \geq \nu\{A\} : \nu\{A\} > 0 \Rightarrow \pi\{A\} = 1, \pi\{A\} < 1 \Rightarrow \nu\{A\} = 0.$
6. Possibility measure is lower continuous, necessity measure is upper continuous.

Definition 4. *A possibilistic (fuzzy) variable is a real function*

$$A(\cdot) : \Gamma \to \mathbb{E}^1,$$

with the values characterized by a distribution of possibilities $\mu_A(x)$:

$$\mu_A(x) = \pi\{\gamma \in \Gamma : A(\gamma) = x\},\ \forall x \in \mathbb{E}^1.$$

$\mu_A(x)$ is the possibility that A takes the value x.

We can define "necessity" variable right the same way [26].

Definition 5. *A necessity variable is a real function*

$$N(\cdot) : \Gamma \to \mathbb{E}^1,$$

with the values characterized by a distribution of possibilities $\mu_N^\nu(x)$:

$$\mu_N^\nu(x) = \nu\{\gamma \in \Gamma : N(\gamma) = x\} =$$

$$= 1 - \pi\{\gamma \in \Gamma : X(\gamma) \neq x\} = 1 - \sup_{u\neq x} \mu_X(u),\ \forall x \in \mathbb{E}^1.$$

It is quite obviuos that "necessity" variable defined above will be nearly always equal to zero. Its straight application is insensible. But necessity measure itself can be used in modelling possibilistic optimization problems in dual context of possibility-necessity.

Definition 6. *The crisp subset*

$$supp(A) = \{x \,|\, \mu_A(x) > 0\},\ x \in \mathbb{E}^1.$$

is the support of the possibilistic variable A .

Definition 7. *For any possibilistic variable A and any* $\alpha \in [0,1]$, A^α *is called the* α*-level set if*

- $A^\alpha = \{x \in \mathbb{E}^1 \,|\, \mu_A(x) \geq \alpha\}$, *for* $\alpha \in (0,1]$,
- $A^\alpha = cl(supp(A))$, *for* $\alpha = 0$,

where $cl(supp(A))$ *is the closure of the support of the possibilistic variable A.*

Definition 8. *The possibilistic variable A is convex if its distribution function is quasiconcave:*

$$\mu_A(\lambda x_1 + (1-\lambda)x_2) \geq \min\{\mu_A(x_1), \mu_A(x_2)\},\ \lambda \in [0,1],\ x_1, x_2 \in \mathbb{E}^1.$$

Generally, possibilistic variables that take values in $\mathbb{E}^1$ and are characterized by strictly unimodular, quasiconcave and upper semicontinuous distribution functions and bounded supports are called fuzzy numbers. If, in this case, the membership function is not strictly unimodular, the possibilistic variable is called a fuzzy interval.

In order to model fuzzy numbers and fuzzy intervals, distributions of (L,R)-type are often used [1, 9, 10, 11].

Definition 9. *We call* (L,R)*-functions, or shape representation functions (or just shapes), the non-increasing and upper semicontinuous functions given on the non-negative part of the number scale such that*

1. $L(0) = R(0) = 1$,
2. $L(t), R(t) < 1, \quad \forall t > 0$,
3. $\lim\limits_{t\to\infty} L(t) = \lim\limits_{t\to\infty} R(t) = 0$.

Definition 10. *A possibilistic variable A is the possibilistic variable of* (L,R)*-type if its distribution has the form*

$$\mu_A(x) = \begin{cases} L\left(\dfrac{a-x}{\alpha}\right), & x < a, \\ 1, & a \leq x \leq b, \\ R\left(\dfrac{x-b}{\beta}\right), & x > b. \end{cases}$$

Here, $[a,b]$ is the tolerance interval of A, a and b are the lower and upper modal values respectively, α and β are the fuzziness coefficients that allow controlling fuzziness of the possibilistic variable. We denote fuzzy intervals by the tuple $A = (a,b,\alpha,\beta)_{LR}$. If $a = b$, the fuzzy interval is reduced to the $A = (a,\alpha,\beta)_{LR}$.

Example 1. If we take a piecewise linear functions $L(x) = R(x) = \max\{1-x,0\}$ as left and right shape functions, then in case of fuzzy numbers we get a so-called triangular fuzzy variables, and in case of fuzzy intervals — trapezoidal fuzzy variables.

Possibilistic variable of (L,R)-type is a convenient mathematical object in the sense that its distribution is parametrized and for calculus of possibilities based on appropriate t-norms (for example, on a well-known $T_M(x,y) = \min\{x,y\}$) when performing arithmetic operations on these possibilistic variables we actually need to perform these operations on their parameters. And in case when corresponding left (right) shape functions of operands are identical then result of an operation has the same left (right) shape — this property belongs to a series of t-norms and is called shape preserving property [19].

2 Aggregation of Fuzzy Information and Definition of Fuzzy Variables Relatedness Based on t-Norms

Aggregation of fuzzy information is based on t-norms and t-conorms, that are extensions of min and max functions, which are used in operations on fuzzy sets and possibilistic variables (calculus of possibilities) [13, 14, 15].

Definition 11. *A triangular norm (briefly t-norm) is a binary operation T on the unit interval $[0,1]$ which is commutative, associative, monotone and has 1 as neutral element, i.e., it is a function $T : [0,1] \times [0,1] \to [0,1]$ such that for all $x,y,z \in [0,1]$:*

1. $T(1,x) = x$, *boundedness,*
2. $T(x,y) = T(y,x)$, *symmetry,*
3. $T(x,T(y,z)) = T(T(x,y),z)$, *associativity,*
4. $T(w,y) \le T(x,z)$, $w \le x, y \le z$, *monotonicity.*

Example 2. Here are some examples of well-known t-norms.

1. $T_W(x,y) = \begin{cases} min\{x,y\}, & \text{if } max\{x,y\} = 1, \\ 0, & \text{otherwise,} \end{cases}$
2. the Lukasiewicz t-norm $T_L(x,y) = \max\{x+y-1,0\}$,
3. the algebraic product $T_P(x,y) = xy$,
4. minimization operation $T_M(x,y) = \min\{x,y\}$.

It is easy to prove the following theorem [13].

Theorem 1. *If T is a t-, then $\forall x,y \in [0,1]$:*

$$T_W(x,y) \le T(x,y) \le T_M(x,y).$$

Triangular norms T_W and T_M are extreme ones, where T_M is called the largest (strongest) and T_W — the smallest (weakest) t-norm respectively (largest and smallest with respect to the pointwise order).

Following after [12], we introduce the notion of mutual t-relatedness of fuzzy sets and possibilistic variables, that generalizes the corresponding notion of unrelatedness given in [21].

Definition 12. *Let $(\Gamma, \mathbb{P}(\Gamma), \pi)$ is a possibilistic space and T is an arbitrary t-norm. Sets $X_1, \ldots, X_n \in \mathbb{P}(\Gamma)$ are called mutually T-related if for any subset $\{i_1, \ldots, i_k\}$ of the set $\{1, \ldots, n\}$, $1 \le k \le n$ the following equation holds:*

$$\pi\Big(X_{i_1} \cap \ldots \cap X_{i_k}\Big) = T\Big(\pi(X_{i_1}), \ldots, \pi(X_{i_k})\Big),$$

where

$$T\Big(\pi(X_{i_1}), \ldots, \pi(X_{i_k})\Big) = T(T(\ldots T(T(\pi(X_{i_1}), \pi(X_{i_2})), \pi(X_{i_3})), \ldots \pi(X_{i_{k-1}})), \pi(X_{i_k})).$$

Let $A_1, \ldots, A_n$ be possibilistic variables defined on possibilistic space $(\Gamma, \mathbb{P}(\Gamma), \pi)$.

Definition 13. *Possibilistic variables $A_1, \ldots, A_n$ are called mutually T-related if for any subset $\{i_1, \ldots, i_k\}$ of the set $\{1, \ldots, n\}$, $1 \le k \le n$:*

$$\mu_{A_{i_1}, \ldots, A_{i_k}}(x_{i_1}, \ldots, x_{i_k}) = \pi(\gamma \in \Gamma \mid A_{i_1}(\gamma) = x_{i_1}, \ldots, A_{i_k}(\gamma) = x_{i_k}) = \pi(A_{i_1}^{-1}\{x_{i_1}\} \cap \ldots \cap A_{i_k}^{-1}\{x_{i_k}\}) = T(\pi(A_{i_1}^{-1}\{x_{i_1}\}), \ldots, \pi(A_{i_k}^{-1}\{x_{i_k}\})), x_{i_j} \in \mathbb{E}^1.$$

One of the main properties of t-norms is their ability to control uncertainty ("fuzziness") growth, which is very likely to appear when performing arithmetic operations on fuzzy numbers. For example, when adding two fuzzy numbers of (L,R)-type using the strongest t-norm T_M corresponding coefficients of fuzziness are summed:

$$(a, \alpha_1, \beta_1)_{LR} \oplus_M (b, \alpha_2, \beta_2)_{LR} = (a+b, \alpha_1 + \alpha_2, \beta_1 + \beta_2)_{LR},$$

i.e. result's coefficient of fuzziness is a sum of operands' coefficients of fuzziness. Therefore uncertainty is growing. Intensive computations under this conditions can lead to uncontrollable growth of fuzziness and thus practical interpretation of final results can be dubious.

With the help of t-norms other than T_M we can have slower growth of fuzziness.

3 Weighted Sum of Possibilistic Variables Based on the Weak t-Norm T_W

If we want to solve a possibilistic optimization problem we must be able to identify possibility distribution functions of those fuzzy variables which form the task's criterion and constraints models. Depending on t-norm that lies in the basis of opti-

mization model we need corresponding calculus of possibilities, which will be used in the process of construction of equivalent crisp analogue.

Elements of such calculus based on the weak t-norm T_W were proposed in [8, 9, 10, 17, 18, 20]. Main result that we will need later on is the following.

Consider n fuzzy intervals: $A_i = (a_i, b_i, \alpha_i, \beta_i)_{LR}$, $i = 1 \ldots n$. Their T_W-sum $\bigoplus_{T_{Wi=1}}^{n} A_i$ is defined on the level of their distribution parameters as

$$\bigoplus_{T_{Wi=1}}^{n} A_i = \left(\sum_{i=1}^{n} a_i, \sum_{i=1}^{n} b_i, \max_{i=1}^{n} \alpha_i, \max_{i=1}^{n} \beta_i \right)_{LR} \tag{8}$$

Thus when summation process is based on the weak t-norm T_W sum's coefficients of fuzziness are calculated as maximums of corresponding coefficients of fuzziness of operands. Note that left and right shapes of operands are identical.

In [3] this result was extended on the case of weighted t-sums. According to [3] we have the following proposition.

Proposition 1. *Let we have n fuzzy intervals: $A_i = (a_i, b_i, \alpha_i, \beta_i)_{LR}$, $i = 1 \ldots n$. Then their weighted T_W-sum $\bigoplus_{T_{Wi=1}}^{n} \lambda_i A_i$, where $\lambda_i \geq 0$ is defined as*

$$\bigoplus_{T_{Wi=1}}^{n} \lambda_i A_i = \left(\sum_{i=1}^{n} \lambda_i a_i, \sum_{i=1}^{n} \lambda_i b_i, \max_{i=1}^{n} \lambda_i \alpha_i, \max_{i=1}^{n} \lambda_i \beta_i \right)_{LR} \tag{9}$$

4 Statements of the Possibilistic Optimization Problems and Their Solution Methods in Case of Mutually T_W-Related Parameters

In this section we investigate two of the most important models [4, 5, 6, 7, 24, 25, 26] of possibilistic programming.

$$k \to \max, \tag{1}$$

$$\tau \{ f_0(x, \gamma) \, \mathfrak{R}_0 \, k \} \geq \alpha_0, \tag{2}$$

$$\begin{cases} \tau \{ f_i(x, \gamma) \, \mathfrak{R}_i \, 0 \} \geq \alpha_i, \; i = 1, \ldots, m, \\ x \in \mathbb{E}_+^N. \end{cases} \tag{3}$$

and

$$\tau \{ f_0(x, \gamma) \, \mathfrak{R}_0 \, 0 \} \to \max, \tag{4}$$

$$\begin{cases} \tau \{ f_i(x, \gamma) \, \mathfrak{R}_i \, 0 \} \geq \alpha_i, \; i = \overline{1, m}, \\ x \in \mathbb{E}_+^n. \end{cases} \tag{5}$$

Here $f_0(x, \gamma) = \sum_{j=1}^{n} a_{0j}(\gamma) x_j$ in case of model (1)-(3), and $f_0(x, \gamma) = \sum_{j=1}^{n} a_{0j}(\gamma) x_j - b_0(\gamma)$ in model (4)-(5), $f_i(x, \gamma) = \sum_{j=1}^{n} a_{ij}(\gamma) x_j - b_i(\gamma), i = 1, \ldots, m$, $\mathbb{E}_+^N$ is non-negative subset of n-dimensional Euclidean space, $\tau \in \{\pi, \nu\}$, $\mathfrak{R}_0, \mathfrak{R}_i \in \{=, \leq, \geq\}$.

These tasks, first of which is known in the context of interval analysis as a so-called optimistic model of decision making and the second one known as fuzzy goal achievement optimization problem, were well studied in the case of the strongest triangular norm T_M that models min-relatedness of its parameters and $\tau = \pi$ in [5, 7, 24] in the class of normal (in Nahmias sense) possibilistic variables. And in general case of possibility and necessity measures these tasks were investigated in [4, 25, 26].

As it was merntioned all these problems were considered in the corresponding works in the case of the strongest triangular norm T_M. We study them for the case of the weak triangular norm T_W.

4.1 Level Optimization Model in the Possibility Context

First, we analyze model (1)-(3) for $\tau = \pi$ and $\Re_0, \Re_i = " = "$. Consider at first the system (3). Let us construct an equivalent crisp analogue. There holds the following theorem.

Theorem 2. *Let in the system of possibilistic inequalities (3) $a_{ij}(\gamma)$ and $b_i(\gamma)$ be mutually T_W-related fuzzy variables of (L,R)-type:*

$$a_{ij}(\gamma) = (a'_{ij}, a''_{ij}, \eta_{ij}, \beta_{ij})_{LR},\ i = 1,\dots,m,\ j = 1,\dots,n,$$
$$b_i(\gamma) = (b'_i, b''_i, \eta_i, \beta_i)_{LR},\ i = 1,\dots,m,$$

with identical left and right shapes L and R. If L and R have inverse functions than equivalent crisp analogue of (3) is

$$\begin{cases} \sum\limits_{j=1}^{n} a'_{ij}x_j - \max\limits_{j=1,\dots,n}\{x_j\eta_{ij}\}L^{-1}(\alpha_i) \le b''_i + \beta_i R^{-1}(\alpha_i), \quad i = 1,\dots,m, \\ \sum\limits_{j=1}^{n} a''_{ij}x_j + \max\limits_{j=1,\dots,n}\{x_j\beta_{ij}\}R^{-1}(\alpha_i) \ge b'_i - \eta_i L^{-1}(\alpha_i), \quad i = 1,\dots,m, \\ x \in \mathbb{E}^N_+. \end{cases} \tag{6}$$

Proof. The function $f_i(x,\gamma)$ is described by

$$f_i(x,\gamma) = \sum_{j=1}^{n} a_{ij}(\gamma)x_j - b_i(\gamma).$$

We move the absolute term $b_i(\gamma)$ into the right-hand side of the equation to obtain the constraint of the following form

$$\begin{cases} \pi\{f'_i(x,\gamma) = b_i(\gamma)\} \ge \alpha_i,\ i = 1,\dots,m, \\ x \in \mathbb{E}^N_+. \end{cases}$$

The function $f_i'(x,\gamma)$ is the weighted T_W-sum of fuzzy variables of the (L,R)-type with the same left and right shapes. By proposition 1 on the weighted T_W-sum of fuzzy variables of (L,R)-type, the possibility distribution $f_i'(x,\gamma)$ has the form

$$\mu_{f_i'(x,\gamma)} = (l_i^*, r_i^*, \eta_i^*, \beta_i^*)_{LR},$$

where:

$$l_i^* = \sum_{j=1}^{n} a'_{ij} x_j,$$
$$r_i^* = \sum_{j=1}^{n} a''_{ij} x_j,$$

$$\eta_i^* = \max\{x_1\eta_{i1}, \ldots, x_n\eta_{in}\} = \max_{j=1\ldots n}\{x_j\eta_{ij}\},$$
$$\beta_i^* = \max\{x_1\beta_{i1}, \ldots, x_n\beta_{in}\} = \max_{j=1\ldots n}\{x_j\beta_{ij}\}.$$

Thus, for $f_i'(x,\gamma)$, the left and right shapes $L^{f'}(x)$ and $R^{f'}(x)$, respectively, are expressed by

$$L^{f'}(x) = L\left(\frac{l_i^* - x}{\eta_i^*}\right), \quad R^{f'}(x) = R\left(\frac{x - r_i^*}{\beta_i^*}\right).$$

For the absolute term, the possibility distribution function has the form given in the hypothesis of the theorem – $b_i(\gamma) = (b_i', b_i'', \eta_i, \beta_i)_{LR}$, with the left and right shapes expressed by

$$L^b(x) = L\left(\frac{b_i' - x}{\eta_i}\right), \quad R^b(x) = R\left(\frac{x - b_i''}{\beta_i}\right).$$

It is obvious that the possibilistic inequality $\pi\{f_i'(x,\gamma) = b_i(\gamma)\} \geq \alpha_i$ is equivalent to the following system of deterministic inequalities

$$\begin{cases} l_i^{f'} \leq r_i^b, \\ r_i^{f'} \geq l_i^b, \end{cases} \tag{7}$$

where $l_i^{f'}$ and $r_i^{f'}$ — are values of arguments of functions $L^{f'}$ and $R^{f'}$, respectively, such that these functions take values α_i: $L^{f'}(l_i^{f'}) = R^{f'}(r_i^{f'}) = \alpha_i$, and l_i^b r_i^b are the values of arguments of functions L^b and R^b, for which they are α_i: $L^b(l_i^b) = R^b(r_i^b) = \alpha_i$.

Let us find l and r

$$L\left(\frac{l_i^* - l_i^{f'}}{\eta_i^*}\right) = \alpha_i \Longrightarrow \frac{l_i^* - l_i^{f'}}{\eta_i^*} = L^{-1}(\alpha_i) \Longrightarrow l_i^{f'} = l_i^* - \eta_i^* L^{-1}(\alpha_i),$$

$$R\left(\frac{r_i^{f'} - r_i^*}{\beta_i^*}\right) = \alpha_i \Longrightarrow \frac{r_i^{f'} - r_i^*}{\beta_i^*} = R^{-1}(\alpha_i) \Longrightarrow r_i^{f'} = r_i^* + \beta_i^* R^{-1}(\alpha_i).$$

And now we find l_i^b and r_i^b:

$$L\left(\frac{b_i' - l_i^b}{\eta_i}\right) = \alpha_i \Longrightarrow \frac{b_i' - l_i^b}{\eta_i} = L^{-1}(\alpha_i) \Longrightarrow l_i^b = b_i' - \eta_i L^{-1}(\alpha_i),$$

$$R\left(\frac{r_i^b - b_i''}{\beta_i}\right) = \alpha_i \Longrightarrow \frac{r_i^b - b_i''}{\beta_i} = R^{-1}(\alpha_i) \Longrightarrow r_i^b = b_i'' + \beta_i R^{-1}(\alpha_i).$$

We rewrite (7) in the form

$$\begin{cases} l_i^* - \eta_i^* L^{-1}(\alpha_i) \le b_i'' + \beta_i R^{-1}(\alpha_i), \\ r_i^* + \beta_i^* R_i^{-1}(\alpha_i) \ge b_i' - \eta_i L^{-1}(\alpha_i). \end{cases}$$

Expanding l_i^*, r_i^*, η_i^* and β_i^*, we have the theorem hypothesis. The theorem is proved.

Equivalent deterministic analogue of criterion model (1)-(2) is given by the following theorem.

Theorem 3. *Let $a_{0j}(\gamma)$ — be mutually T_W-related fuzzy variables of (L,R)-type:*

$$a_{0j}(\gamma) = (a_{0j}', a_{0j}'', \eta_{0j}, \beta_{0j})_{LR},\ j = 1, \ldots, n,$$

with identical left and right shapes L and R. If L and R have inverse functions than equivalent crisp analogue of (1)-(2) is

$$k \to \max, \tag{8}$$

$$\begin{cases} \sum\limits_{j=1}^{n} a_{0j}' x_j - \max\limits_{j=1,\ldots,n} \{x_j \eta_{0j}\} L^{-1}(\alpha_0) \le k, \\ \sum\limits_{j=1}^{n} a_{0j}'' x_j + \max\limits_{j=1,\ldots,n} \{x_j \beta_{0j}\} R^{-1}(\alpha_0) \ge k. \end{cases} \tag{9}$$

Proof. The function $f_0(x,\gamma)$ is described by the expression

$$f_0(x,\gamma) = \sum_{j=1}^{n} a_{0j}(\gamma) x_j.$$

We have the following form of the criterion

$$\begin{cases} \pi\{f_0(x,\gamma) = k\} \ge \alpha_0, \\ x \in \mathbb{E}_+^n. \end{cases}$$

The function $f_0(x,\gamma)$ is a weighted T_W-sum of fuzzy variables of the (L,R)-type with the same left and right shapes. Similar to the proof of Theorem 2, we find the possibility distribution $f_0(x,\gamma)$ and the left and right boundaries of its α_0-level set

$$l_0^f = l_0^* - \eta_0^* L^{-1}(\alpha_0),$$

$$r_0^f = r_0^* + \beta_0^* R^{-1}(\alpha_0),$$

where:

$$l_0^* = \sum_{j=1}^{n} a'_{0j} x_j, \qquad r_0^* = \sum_{j=1}^{n} a''_{0j} x_j,$$

$$\eta_0^* = \max_{j=1\ldots n} \{x_j \eta_{0j}\}, \qquad \beta_0^* = \max_{j=1\ldots n} \{x_j \beta_{0j}\}.$$

The possibilistic inequality $\pi\{f_0(x,\gamma) = k\} \geq \alpha_0$ is equivalent to the following system of deterministic inequalities

$$\begin{cases} l_0^f \leq k, \\ r_0^f \geq k. \end{cases} \tag{10}$$

We rewrite (10) in the form

$$\begin{cases} l_0^* - \eta_0^* L^{-1}(\alpha_0) \leq k, \\ r_0^* + \beta_0^* R^{-1}(\alpha_0) \geq k. \end{cases} \tag{11}$$

Expanding l_0^*, r_0^*, η_0^* and β_0^* (11), we have the theorem hypothesis. The theorem is proved.

Let us simplify deterministic analogue (8)-(9) of the model of the criterion. We denote

$$f'(x) = \sum_{j=1}^{n} a'_{0j} x_j - \max_{j=1,\ldots,n} \{x_j \eta_{0j}\} L^{-1}(\alpha_0),$$

$$f''(x) = \sum_{j=1}^{n} a''_{0j} x_j + \max_{j=1,\ldots,n} \{x_j \beta_{0j}\} R^{-1}(\alpha_0).$$

Then, we can rewrite model (8)-(9) as

$$\begin{gathered} k \to \max, \\ f'(x) \leq k \leq f''(x). \end{gathered} \tag{12}$$

In turn, model (12) is reduced to the equivalent model with no additional (level) variable k

$$\begin{gathered} f''(x) \to \max, \\ f''(x) - f'(x) \geq 0. \end{gathered}$$

As a result, we can reduce problem (8),(9),(6) to the form

$$f''(x) \to \max, \tag{13}$$

$$\begin{cases} f''(x) - f'(x) \geq 0, \\ \sum\limits_{j=1}^{n} a'_{ij} x_j - \max\limits_{j=1,\ldots,n} \{x_j \eta_{ij}\} L^{-1}(\alpha_i) \leq b''_i + \beta_i R^{-1}(\alpha_i), \quad i = 1,\ldots,m, \\ \sum\limits_{j=1}^{n} a''_{ij} x_j + \max\limits_{j=1,\ldots,n} \{x_j \beta_{ij}\} R^{-1}(\alpha_i) \geq b'_i - \eta_i L^{-1}(\alpha_i), \quad i = 1,\ldots,m, \\ x \in \mathbb{E}^N_+. \end{cases} \tag{14}$$

In the general case, (13)-(14) is the problem of nonconvex and nonsmooth optimization. Therefore we use genetic algorithm method to solve it.

4.2 Level Optimization Model in the Necessity Context

Let us consider now t model (1)-(3) in the case of $\tau = \nu$, $R_0 =$ " $\geq$ " and $R_i =$ " $\leq$ ". The corresponding model takes the following form

$$k \to \max, \tag{15}$$

$$\nu\{f_0(x,\gamma) \geq k\} \geq \alpha_0, \tag{16}$$

$$\begin{cases} \nu\{f_i(x,\gamma) \leq 0\} \geq \alpha_i, \; i = \overline{1,m}, \\ x \in \mathbb{E}^n_+. \end{cases} \tag{17}$$

At first we consider constraints model (17). We construct an equivalent deterministic system for it.

Theorem 4. *Let in the constraints model (17) $a_{ij}(\gamma)$ and $b_i(\gamma)$ are mutually T_W-related fuzzy variables of (L,R)-type:*

$$a_{ij}(\gamma) = (a'_{ij}, a''_{ij}, \eta_{ij}, \beta_{ij})_{LR}, \; i = \overline{1,m}, \; j = \overline{1,n},$$
$$b_i(\gamma) = (b'_i, b''_i, \eta_i, \beta_i)_{LR}, \; i = \overline{1,m}$$

with identical left and right shapes L and R. If L and R have inverse functions than equivalent deterministic analogue of (17) is

$$\begin{cases} \sum\limits_{j=1}^{n} a''_{ij} x_j + \max\limits_{j=1,\ldots,n} \{x_j \beta_{ij}\} R^{-1}(1-\alpha_i) \leq b'_i - \eta_i L^{-1}(1-\alpha_i), \; i = \overline{1,m}, \\ x \in \mathbb{E}^n_+. \end{cases} \tag{18}$$

Proof. The function $f_i(x,\gamma)$ is described by

$$f_i(x,\gamma) = \sum_{j=1}^{n} a_{ij}(\gamma) x_j - b_i(\gamma)$$

therefore the constraints model (17) can be rewritten in the form

$$\begin{cases} v\{f_i'(x,\gamma) \le b_i(\gamma)\} \ge \alpha_i,\ i = \overline{1,m}, \\ x \in \mathbb{E}_+^N. \end{cases} \tag{19}$$

The function $f_i'(x,\gamma) = \sum_{j=1}^n a_{ij}(\gamma)x_j$ in the constraints model (19) is weighted T_W-sum of fuzzy variables of the (L,R)-type with the same left and right shapes. By proposition 1 on the weighted T_W-sum of fuzzy variables of (L,R)-type, the possibility distribution $f_i'(x,\gamma)$ has the form

$$\mu_{f_i'(x,\gamma)} = (l_i^*, r_i^*, \eta_i^*, \beta_i^*)_{LR},$$

where

$$l_i^* = \sum_{j=1}^n a_{ij}' x_j,\ r_i^* = \sum_{j=1}^n a_{ij}'' x_j,$$

$$\eta_i^* = \max\{x_1\eta_{i1}, \ldots, x_n\eta_{in}\} = \max_{j=1,\ldots,n}\{x_j\eta_{ij}\},$$
$$\beta_i^* = \max\{x_1\beta_{i1}, \ldots, x_n\beta_{in}\} = \max_{j=1,\ldots,n}\{x_j\beta_{ij}\}.$$

As a result we have that $f_i'(x,\gamma)$ has the following shape function $R^{f_i'}(x)$:

$$R^{f_i'}(x) = R\left(\frac{x - r_i^*}{\beta_i^*}\right).$$

It is obvious that for the absolute term, the possibility distribution function has the form given in the hypothesis of the theorem – $b_i(\gamma) = (b_i', b_i'', \eta_i, \beta_i)_{LR}$, with the left and right shapes expressed by

$$L^{b_i}(x) = L\left(\frac{b_i' - x}{\eta_i}\right).$$

Because of duality of possibility and necessity measures we have

$$v\{f_i'(x,\gamma) \le b_i(\gamma)\} \ge \alpha_i \Rightarrow \pi\{f_i'(x,\gamma) > b_i(\gamma)\} \le 1 - \alpha_i.$$

According to [26] possibilistic inequality $\pi\{f_i'(x,\gamma) > b_i(\gamma)\} \le 1 - \alpha_i$ is equivalent to the following system of deterministic inequalities

$$r_i^{f_i'} \le l_i^{b_i}, \tag{20}$$

where $r_i^{f_i'}$ is the value of argument of function $R^{f_i'}$, such that this function reaches extremum $1 - \alpha_i$: $R^{f_i'}(r_i^{f_i'}) = \alpha_i$, and l^{b_i} is the value of argument of function L^{b_i}, such that this function also reaches the same extremum $1 - \alpha_i$ $L^{b_i}(l_i^{b_i}) = \alpha_i$. Let us find $r_i^{f_i'}$:

$$R\left(\frac{r_i^{f_i'} - r_i^*}{\beta_i^*}\right) = 1 - \alpha_i \Longrightarrow \frac{r_i^{f_i'} - r_i^*}{\beta_i^*} = R^{-1}(1 - \alpha_i) \Longrightarrow r_i^{f_i'} = r_i^* + \beta_i^* R^{-1}(1 - \alpha_i).$$

And now let us find $l_i^{b_i}$. We have:

$$L\left(\frac{b_i' - l_i^{b_i}}{\eta_i}\right) = 1 - \alpha_i \Longrightarrow \frac{b_i' - l_i^{b_i}}{\eta_i} = L^{-1}(1-\alpha_i) \Longrightarrow l_i^{b_i} = b_i' - \eta_i L^{-1}(1-\alpha_i).$$

Than inequality (20) can be rewritten in the form

$$r_i^* + \beta_i^* R_i^{-1}(1-\alpha_i) \le b_i' - \eta_i L^{-1}(1-\alpha_i).$$

Putting l_i^*, r_i^*, η_i^* and β_i^* in their explicit form, we have the theorem hypothesis. The theorem is proved.

Remark 1. In the hypotheses of Theorem 4 constraints model (18) defines convex set of optimization task (1)-(3) feasible region.

Indeed, it is easy to see that for any $i = \overline{1,n}$ function

$$\sum_{j=1}^{n} a_{ij}'' x_j + \max_{j=1,\ldots,n} \{x_j \beta_{ij}\} R^{-1}(1-\alpha_i)$$

is convex because it is a sum of linear and convex functions. Thus i-th constraint is a convex set. Intersection of convex sets is also a convex set.

4.3 Fuzzy Goal Achievement Maximization Model in the Possibility Context

We now move to investigating the model (4)-(5) with $\tau = \pi$, $R_0, R_i = " = "$.

Let us consider the following possibilistion programming problem

$$\pi\{f_0(x,\gamma) = 0\} \to \max \tag{21}$$

$$\begin{cases} \pi\{f_i(x,\gamma) = 0\} \ge \alpha_i, i = 1,\ldots,m \\ x \in X \end{cases} \tag{22}$$

Here $f_0(x,\gamma) = \sum_{j=1}^{n} a_{0j}(\gamma)x_j - b_0(\gamma)$, and $f_i(x,\gamma) = \sum_{j=1}^{n} a_{ij}(\gamma)x_j - b_i(\gamma), i = 1,\ldots,m$.

At first, we study the system (22). We construct an equivalent deterministic system for it. We have the following theorem.

Theorem 5. *Let in the system of possibilistic inequalities (22) $a_{ij}(\gamma)$ and $b_i(\gamma)$ are mutually T_W-related fuzzy variables of (L,R)-type:*

$$a_{ij}(\gamma) = (a_{ij}', a_{ij}'', \eta_{ij}, \beta_{ij})_{LR}, i = \overline{1,m}, j = \overline{1,n},$$
$$b_i(\gamma) = (b_i', b_i'', \eta_i, \beta_i)_{LR}, i = \overline{1,m},$$

with identical left and right shapes L and R. If L and R have inverse functions than equivalent crisp analogue of (22) is

$$\begin{cases} \sum\limits_{j=1}^{n} a'_{ij}x_j - \max\limits_{j=1,\dots,n}\{x_j\eta_{ij}\}L^{-1}(\alpha_i) \le b''_i + \beta_i R^{-1}(\alpha_i), \quad i=\overline{1,m}, \\ \sum\limits_{j=1}^{n} a''_{ij}x_j + \max\limits_{j=1,\dots,n}\{x_j\beta_{ij}\}R^{-1}(\alpha_i) \ge b'_i - \eta_i L^{-1}(\alpha_i), \quad i=\overline{1,m}, \\ x \in E^n_+. \end{cases} \tag{23}$$

Proof. Proof of this theorem is the same as in the Theorem 2.

The following theorem gives us an equivalent deterministic analogue of criterion model (21).

Theorem 6. *Let in the criterion model (21) $a_{0j}(\gamma)$ — be mutually T_W-related fuzzy variables of (L,R)-type:*

$$a_{0j}(\gamma) = (a'_{0j}, a''_{0j}, \eta_{0j}, \beta_{0j})_{LR}, j=\overline{1,n},$$
$$b_0(\gamma) = (b'_0, b''_0, \eta_0, \beta_0)_{LR}$$

with identical left and right shapes L and R. If L and R have inverse functions than equivalent crisp analogue of (21) is

$$\begin{cases} \alpha \to \max, \\ \sum\limits_{j=1}^{n} a'_{0j}x_j - \max\limits_{j=1,\dots,n}\{x_j\eta_{0j}\}L^{-1}(\alpha) \le b''_0 + \beta_0 R^{-1}(\alpha), \\ \sum\limits_{j=1}^{n} a''_{0j}x_j + \max\limits_{j=1,\dots,n}\{x_j\beta_{0j}\}R^{-1}(\alpha) \ge b'_0 - \eta_0 L^{-1}(\alpha), \\ \alpha \in [0,1], \\ x \in E^n_+. \end{cases} \tag{24}$$

Proof. The function $f_0(x,\gamma)$ is described by the following expression

$$f_0(x,\gamma) = \sum_{j=1}^{n} a_{0j}(\gamma)x_j - b_0(\gamma).$$

We move the absolute term $b_0(\gamma)$ into the right-hand side of the equation to obtain the constraint of the following form

$$\pi\{f'_0(x,\gamma) = b_0(\gamma)\} \to \max,$$
$$x \in \mathbb{E}^N_+.$$

The function $f'_0(x,\gamma)$ is the weighted T_W of fuzzy variables of the (L,R)-type with the same representation functions of the form. By proposition 1 on the weighted T_W-sum of fuzzy variables of (L,R)-type, the possibility distribution of $f'_0(x,\gamma)$ has the form

$$\mu_{f'_0(x,\gamma)} = (l^*_0, r^*_0, \eta^*_0, \beta^*_0)_{LR},$$

where

$$l_0^* = \sum_{j=1}^{n} a'_{0j} x_j,$$
$$r_0^* = \sum_{j=1}^{n} a''_{0j} x_j,$$

$$\eta_0^* = \max\{x_1\eta_{01}, \ldots, x_n\eta_{0n}\} = \max_{j=1\ldots n}\{x_j\eta_{0j}\},$$
$$\beta_0^* = \max\{x_1\beta_{01}, \ldots, x_n\beta_{0n}\} = \max_{j=1\ldots n}\{x_j\beta_{0j}\}.$$

Thus, for $f'_0(x,\gamma)$, the left and right shapes $L^{f'_0}(x)$ and $R^{f'_0}(x)$, respectively, are expressed by

$$L^{f'_0}(x) = L\left(\frac{l_0^* - x}{\eta_0^*}\right), \quad R^{f'_0}(x) = R\left(\frac{x - r_0^*}{\beta_0^*}\right).$$

For the absolute term, the possibility distribution function has the form given in the hypothesis of the theorem – $b_0(\gamma) = (b'_0, b''_0, \eta_0, \beta_0)_{LR}$, with the left and right shapes expressed by

$$L^{b_0}(x) = L\left(\frac{b'_0 - x}{\eta_0}\right), \quad R^{b_0}(x) = R\left(\frac{x - b''_0}{\beta_0}\right).$$

Let us consider some arbitrary $\alpha \in [0,1]$ and find deterministic analog of the following possibilistic equality

$$\pi\{f'_0(x,\gamma) = b_0(\gamma)\} \geq \alpha. \tag{25}$$

Putting $\alpha \to \max$ after that, we obtain necessary deterministic analog of the criterion model.

As in the proof of the Theorem 3 we have that possibilistic inequality (25) is equivalent to the following system of deterministic inequalitities

$$\begin{cases} l_0^{f'_0} \leq r_0^{b_0}, \\ r_0^{f'_0} \geq l_0^{b_0}, \end{cases} \tag{26}$$

where $l_0^{f'_0}$ and $r_0^{f'_0}$ are values of arguments of functions L and R, respectively, such that these functions take values α: $L(l_0^{f'_0}) = R(r_0^{f'_0}) = \alpha$, and $l_0^{b_0}$ and $r_0^{b_0}$ are the values of arguments of functions L and R, for which they are α: $L(l_0^{b_0}) = R(r_i^{b_0}) = \alpha$.

Let us find $l_0^{f'_0}$ and $r_0^{f'_0}$:

$$L\left(\frac{l_0^* - l_0^{f'_0}}{\eta_0^*}\right) = \alpha \Longrightarrow \frac{l_0^* - l_0^{f'_0}}{\eta_0^*} = L^{-1}(\alpha) \Longrightarrow l_0^{f'_0} = l_0^* - \eta_0^* L^{-1}(\alpha_0),$$

$$R\left(\frac{r_0^{f'_0} - r_0^*}{\beta_0^*}\right) = \alpha \Longrightarrow \frac{r_0^{f'_0} - r_0^*}{\beta_0^*} = R^{-1}(\alpha) \Longrightarrow r_0^{f'_0} = r_0^* + \beta_0^* R^{-1}(\alpha).$$

And now we find $l_0^{b_0}$ and $r_0^{b_0}$:

$$L\left(\frac{b_0' - l_0^{b_0}}{\eta_0}\right) = \alpha \Longrightarrow \frac{b_0' - l_0^{b_0}}{\eta_0} = L^{-1}(\alpha) \Longrightarrow l_0^{b_0} = b_0' - \eta_0 L^{-1}(\alpha),$$

$$R\left(\frac{r_0^{b_0} - b_0''}{\beta_0}\right) = \alpha \Longrightarrow \frac{r_0^{b_0} - b_0''}{\beta_0} = R^{-1}(\alpha) \Longrightarrow r_0^{b_0} = b_0'' + \beta_0 R^{-1}(\alpha).$$

Expanding variables in the system (26) and putting $\alpha \to \max$, we obtain the theorem hypothesis. The theorem is proved.

5 Genetic Algorithm for Solving the Problems

We can represent the general scheme of a genetic algorithm as the following sequence of steps [16].

1. Initializing the algorithm parameters, in particular, *popsize*, a, P_c, and P_m, where
   - *popsize* is the number of chromosomes in each population;
   - a is the constant that influences the algorithm of chromosome selection;
   - P_c is the crossover probability (about $P_c \times popsize$ chromosomes generate new chromosomes);
   - P_m is the mutation probability (about $P_m \times popsize$ chromosomes mutate).
2. Initializing the initial population by *popsize* chromosomes.
3. Modifying the current set of chromosomes by crossover and mutation operations.
4. Calculating the criterial function for each chromosome.
5. Calculating the fitness function for each chromosome, using the criterial function.
6. Selecting chromosomes by the roulette algorithm and forming a new population.
7. Repeating steps 3–6 necessary number of times.
8. Output the best chromosome as the optimal solution to the problem.

Since the best chromosome of the i-th population does not necessarily move to the $(i+1)$-th population, we store the value of the best chromosome obtained at the previous stages of the algorithm. If we find a better chromosome, we store it rather than the previous one. We consider all steps of the algorithm in more detail.

5.1 Chromosomes

In genetic algorithms, chromosomes are used to encode solutions to the stated optimization problem. There are at least two ways to represent chromosomes, binary and real. In the first case, all components of the vector of unknowns $x = (x_1, x_2, \ldots, x_n)$

are transformed into the binary form and concatenated. Then, all crossover and mutation operation are performed over sequences of zeros and ones. In the second case, the genetic algorithm operates directly with the very vector x that is not anyhow encoded. Both methods have their own advantages and drawbacks.

We consider the second method for the algorithm not to become more complicated with encoding and decoding binary rows and for the sake of visualization of the genetic process.

5.2 Initializing the Algorithm Parameters

Using parameters, we can control the way the genetic algorithm operates and, in general case, determine its efficiency. The software implementation of the algorithm allows us experimentally choose the values of parameters that provide the most efficient way to solve the stated problem. Initially, we choose parameters depending on the estimates of experts in the subject field of the problem to be solved or at random.

The size of the population is one of the most important parameters. If it is too small, steps of the algorithm are performed rapidly; however, the general algorithm convergence to the optimal solution becomes slow and the risk of sticking in the area of the local optimum increases. When the size of population is too big, performing one step of the algorithm to construct a new population takes much time; however, operation spectrum of such algorithm is very wide and the probability of finding the global extremum for a smaller number of iterations is greater than in the first case. However, since it takes long to perform each iteration, the very algorithm operates for a long time, similar to the first case.

5.3 Initializing the Initial Population

To initialize the initial population, in the general case, we assign arbitrary values, which nevertheless stay within the region of admissible solutions to the stated problem, to all $popsize$ chromosomes.

5.4 Performing the Crossover Operations

To perform the crossover operation, we select $P_c \times popsize$ chromosomes at random. For all $i = 1, \ldots, popsize$ chromosomes, we generate a random number $r \in [0,1]$ uniformly distributed on this intervall. If $r \leq P_c$, we select the ith chromosome for the crossover operation.

As a result, we have the set of $P_c \times popsize$ chromosomes – $V_1, V_2, V_3, \ldots, V_c$, $c \approx P_c \times popsize$. Since pairs of chromosomes take part in the crossover operation, we can eliminate one of the chromosomes if their number is odd. We form pairs of chromosomes $(V_1', V_2'), (V_3', V_4'), \ldots$, etc. at random. Consider the pair (V_1', V_2'). We generate the random number $c \in (0,1)$. Then, we perform the crossover operation for (V_1', V_2')

$$X = c \times V_1' + (1-c) \times V_2', \qquad Y = (1-c) \times V_1' + c \times V_2'.$$

Values X and Y are descendants of two initial chromosomes. If the region of admissible solutions were convex, this crossover operation would ensure that X and Y belong to the region of admissible solutions as well. However, this is not true in our case. Therefore, having obtained X and Y, we need to check if they still stay within this region. If one of new chromosomes is outside the region, we save the second chromosome and perform another crossover operation. We repeat until we either get two admissible but different chromosomes or exceed the admissible number of iterations. Then, we replace the initial chromosomes V_1' and V_2' by the new X Y (if any).

5.5 Performing the Mutation Operation

Similar to Section 6.4, we select, at random, about $P_m \times popsize$ chromosomes to be acted upon by the mutation operation. For all $i = 1, \ldots, popsize$, we generate the number $r \in [0,1]$ at random. If $r \leq P_m$, we select the ith chromosome for the mutation operation.

Then, for each selected chromosome V_i'', we perform the mutation operation as follows. At random, we assign a sufficiently big number M (the range is also determined at step 1 of the algorithm). Then, at random, we generate the vector $d \in \mathbb{R}^n$. We form the mutated chromosome by the formula

$$Z = V_i'' + M \times d.$$

If Z is outside the region of admissible solutions to the problem, we obtain the number $M' \in (0, M)$ at random and repeat the operation once again. We repeat until we either get a new admissible chromosome or exceed the given number of iterations of the mutation operation. Then, we replace the initial chromosome V_i'' by the found one.

5.6 The Criterial Function and the Fitness Function

In our case, the criterion of the optimization problem to be solved is the criterial function. We use the suitability function to range chromosomes of the population,

assigning the survivability probabilities to them so that the more suitable chromosome has the greater probability to move to the next population. We give one of the ways to construct this function.

We arrange all chromosomes in the decreasing order of their "fitness" (for our maximization problem, in the decreasing order of the value of the criterial function). Then, we choose the parameter $a \in (0,1)$ at random and write the fitness formula

$$eval(V_i) = a \times (1-a)^{i-1},\ i = 1, \ldots, popsize. \tag{27}$$

As can be seen from (27), the real values of the criterial function are of no importance here, what counts is the order of chromosomes. In what follows, this function is used in the chromosome selection algorithm.

5.7 Chromosome Selection Algorithm

Chromosome selection is based on the so-called roulette algorithm, which can be visually imagined as follows. Suppose we have a round disk that can revolve on its axis and is put on a rod. An meter pointing to the circumference of the disk is attached to the rod rigidly. We divide the disk into *popsize* sectors and put one chromosome from the current population into each sector. The area of the sector with the i chromosome is

$$\frac{eval(V_i)}{\sum_{j=1}^{popsize} eval(V_j)}$$

which is some portion of the total disk area. We spin the roulette, and when it stops, we choose the chromosome in the sector pointed to by the meter. We duplicate this chromosome and place the duplicate to the resulting set. We perform this operation *popsize* times. Formally, we can describe this algorithm as follows.

- Calculate the cumulative "probability" of selection q_i for each chromosome V_i

$$\begin{cases} q_0 = 0, \\ q_i = \sum\limits_{j=1}^{i} eval(V_j),\ i = 1, \ldots, popsize. \end{cases}$$

- Repeat the next two steps *popsize* times to obtain *popsize* duplicates of chromosomes. To do this, generate the number $r \in (0, q_{popsize}]$, choose the chromosome V_i, where i is determined by the rule $q_{i-1} < r \leq q_i$, and place it to the resulting set.

It is worth noting that the selection results in the population with a number of duplicating chromosomes. The more suitable the chromosome was in the initial population, the more duplicates it will have in the new population.

6 Model Example and Comparative Study

6.1 Comparative Study

By the model example, we study the behavior of the problem of possibilistic optimization, depending on the t-norm that describes the interaction of fuzzy parameters.

6.1.1 Example of Solving the Problem of Level Optimization with Mutually T_W-Related Parameters

We consider problem of level optimization with mutually T_W-related parameters for $n = m = 2$ and $\alpha_0 = 0.5$, $\alpha_1 = 0.5$ and $\alpha_2 = 0.5$

$$k \to \max, \tag{28}$$

$$\pi\{a_{01}(\gamma)x_1 + a_{02}(\gamma)x_2 = k\} \geq 0.5, \tag{29}$$

$$\begin{cases} \pi\{a_{11}(\gamma)x_1 + a_{12}(\gamma)x_2 - b_1(\gamma) = 0\} \geq 0.5, \\ \pi\{a_{21}(\gamma)x_1 + a_{22}(\gamma)x_2 - b_2(\gamma) = 0\} \geq 0.5, \\ x \in \mathbb{E}^2_+. \end{cases} \tag{30}$$

Here $a_{01}(\gamma)$, $a_{02}(\gamma)$, $a_{11}(\gamma)$, $a_{12}(\gamma)$, $a_{21}(\gamma)$, $a_{22}(\gamma)$, $b_1(\gamma)$ $b_2(\gamma)$ are triangular mutually T_W-related fuzzy variables with (L,R)-type distributions

$$a_{01}(\gamma) = (3,3,4.5,4)_{LR}, \quad a_{02}(\gamma) = (-2,-2,5,7)_{LR},$$

$$a_{11}(\gamma) = (-6,-6,3,3)_{LR}, \; a_{12}(\gamma) = (8,8,0.5,0.6)_{LR},$$

$$a_{21}(\gamma) = (4,4,3,3)_{LR}, \qquad a_{22}(\gamma) = (10,10,4.3,4.2)_{LR},$$

$$b_1(\gamma) = (3,3,1.5,1.5)_{LR}, \; b_2(\gamma) = (10,10,6,6)_{LR},$$

where

$$L(t) = R(t) = \max\{0, 1-t\}, \; t \in \mathbb{E}^1_+.$$

We construct the deterministic equivalent analogue of the problem (28)-(30). By theorems 2 and 3, we have

$$3x_1 - 2x_2 + 0.5\max\{4x_1, 7x_2\} \to \max, \tag{31}$$

$$\begin{cases} \max\{4x_1, 7x_2\} + \max\{4.5x1, 5x2\} \geq 0, \\ -6x_1 + 8x_2 - 0.5\max\left\{3x_1, \frac{1}{2}x_2\right\} \leq 3.75, \\ -6x_1 + 8x_2 + 0.5\max\{3x_1, 0.6x_2\} \geq 2.25, \\ 4x_1 + 10x_2 - 0.5\max\{3x_1, 4.3x_2\} \leq 13, \\ 4x_1 + 10x_2 + 0.5\max\{3x_1, 4.2x_2\} \geq 7, \\ x \in \mathbb{E}^2_+. \end{cases} \tag{32}$$

We use the following constants to initialize the genetic algorithm of solving the obtained problem

1) the population size $popsize = 50$;
2) the probability of mutation $P_m = 0.2$;
3) the probability of crossover $P_c = 0.8$;
4) the constant taking part in selection of chromosomes $a = 0.05$.

Figure 1 illustrates the operation of the genetic algorithm. We can see that the genetic algorithm converges to the optimum as early as at the 80th iteration up to the accuracy of three decimal digits. The ordinate axis shows the values of the best chromosome up to the accuracy of three decimal digits, while the abscissa axis shows the number of iterations of the genetic algorithm.

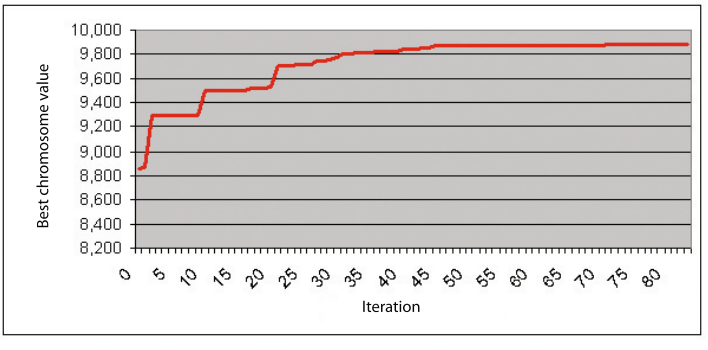


Fig. 1 Operation of the genetic algorithm

The solution to the problem (31)-(32) obtained by the genetic algorithm is

$$x_1 \approx 1.2824, \quad x_2 \approx 1.0026, \quad F_W(x_1, x_2) \approx 5.3511.$$

6.1.2 An Example of Solving the Problem of Level Optimization with Minrelated (Unrelated) Parameters and Crisp Parameters

For the sake of comparison, we solve the stated problem (28)-(30) with mutually T_M-related (minrelated) fuzzy parameters and when there is no fuzziness. The corresponding crisp problem has the form

$$3x_1 - 2x_2 \to \max,$$
$$\begin{cases} -6x_1 + 8x_2 - 3 = 0, \\ 4x_1 + 10x_2 - 10 = 0, \\ x \in \mathbb{E}^2_+, \end{cases}$$

There exists a unique solution

$$x_1 \approx 0.5435,\ x_2 \approx 0.7826, \quad F_C(x_1, x_2) \approx 0.0653.$$

According to [25], the equivalent deterministic analogue of the model example (28)-(30) with mutually T_M-related fuzzy parameters is represented by the linear programming problem

$$5x_1 + 1.5x_2 \to \max,$$
$$\begin{cases} -7.5x_1 + 7.75x_2 \le 3.75, \\ -4.5x_1 + 8.3x_2 \ge 2.25, \\ 2.5x_1 + 7.85x_2 \le 13, \\ 5.5x_1 + 12.1x_2 \ge 7, \\ x \in \mathbb{E}^2_+. \end{cases}$$

We write its solution as

$$x_1 \approx 1.6092, \quad x_2 \approx 1.1436, \quad F_M(x_1, x_2) \approx 9.7615.$$

The table represents solutions of the model problem for all three cases. Figure 2 illustrates feasible regions of the tasks from the corresponding examples.

Table 1

parameters	x_1	x_2	$F_T(x_1, x_2)$
crisp	0.5435	0.7826	0.0653
mutually T_W-related	1.2824	1.0026	5.3511
mutually minrelated	1.6092	1.1436	9.7615

6.2 Theorem of Feasible Regions Subsethood

As one can see from the comparative study feasible region in the case of mutually T_W-related parameters turns out to be more compact and corresponding solutions are less "fuzzy" than in the case of minrelated parameters of initial task. We can prove this strictly mathematically.

Consider the task of maximization of fuzzy goal achievement level in possibility context.

$$k \to \max, \tag{33}$$

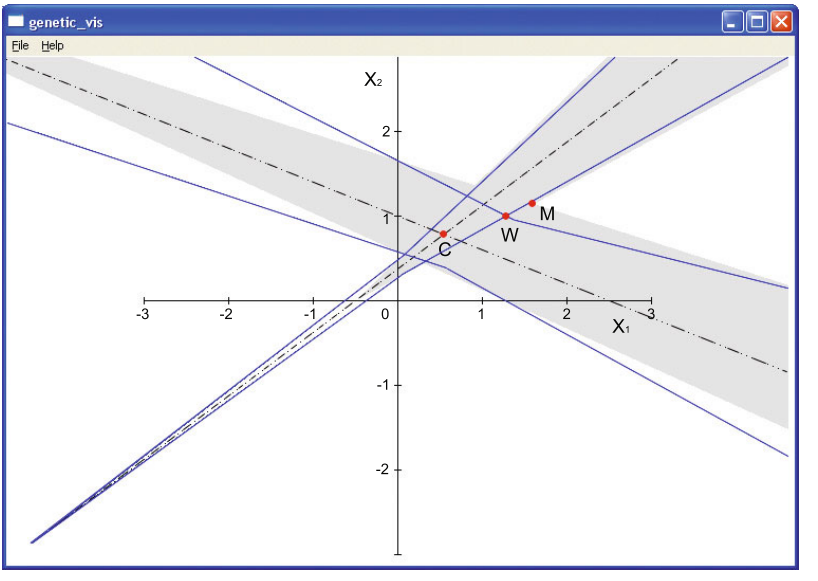


Fig. 2 Feasible regions of equivalent deterministic analogues of the problem (28)-(30) for the case of mutually T_W-related (the intersection of highlighted regions), minrelated (the region bounded by four solid lines) and crisp (the point of intersection of two dash-and-dot lines) parameters; W, M, and C are the points of optimum for the corresponding cases. For the sake of obviousness region $\mathbb{E}^2_+$ and region spcified by the first inequality of the system (32) are omitted.

$$\pi\{f_0(x,\gamma) = k\} \geq \alpha_0, \tag{34}$$

$$\begin{cases} \pi\{f_i(x,\gamma) = 0\} \geq \alpha_i,\ i = \overline{1,m}, \\ x \in \mathbb{E}^n_+. \end{cases} \tag{35}$$

We use $X_M^{\alpha_i}$ to denote the set of solutions x corresponding to the ith constraint of model (35) that satisfies

$$X_M^{\alpha_i} = \left\{x : \pi\{f_i(x,\gamma) = 0\} \geq \alpha_i\right\},$$

$a_{ij}(\gamma),\ j = \overline{1,n}$ and $b_i(\gamma)$ are minrelated,

and $X_W^{\alpha_i}$ denote the set given by conditions

$$X_W^{\alpha_i} = \left\{x : \pi\{f_i(x,\gamma) = 0\} \geq \alpha_i\right\},$$

$a_{ij}(\gamma),\ j = \overline{1,n}$ and $b_i(\gamma)$ are mutually T_W-related.

The following theorem holds.

Theorem 7

$$\bigcap_{i=1}^{m} X_W^{\alpha_i} \subseteq \bigcap_{i=1}^{m} X_M^{\alpha_i}.$$

Proof. Consider the ith constraint of model (35) $\pi\{f_i(x,\gamma)=0\} \geq \alpha_i$. We construct its corresponding sets $X_W^{\alpha_i}$ and $X_M^{\alpha_i}$. Let all $a_{ij}(\gamma),\ j=\overline{1,n}$ and $b_i(\gamma)$ be minrelated. Then, by [25], we have the equivalent deterministic analogue of the given constraint

$$\begin{cases} \sum\limits_{j=1}^{n} a_{ij}^{-}(\alpha_i)x_j \leq b_i^{+}(\alpha_i), \\ \sum\limits_{j=1}^{n} a_{ij}^{+}(\alpha_i)x_j \geq b_i^{-}(\alpha_i), \\ x \in \mathbb{E}_+^n, \end{cases} \tag{36}$$

where $a_{ij}^{-}(\alpha_i), j=\overline{1,n}$ and $b_i^{-}(\alpha_i)$ are left boundaries, and $a_{ij}^{+}(\alpha_i), j=\overline{1,n}$ and $b_i^{+}(\alpha_i)$ are right boundaries of α_i-level sets of fuzzy variables $a_{ij}(\gamma), j=\overline{1,n}$ and $b_i(\gamma)$, respectively.

Now, we assume that $a_{ij}(\gamma), j=\overline{1,n}$ and $b_i(\gamma)$ are mutually T_W-related. Then, by Theorem 2, we have

$$\begin{cases} \sum\limits_{j=1}^{n} a'_{ij}x_j - \max\limits_{j=1,\ldots,n}\{x_j\eta_{ij}\}L^{-1}(\alpha_i) \leq b''_i + \beta_i R^{-1}(\alpha_i), \\ \sum\limits_{j=1}^{n} a''_{ij}x_j + \max\limits_{j=1,\ldots,n}\{x_j\beta_{ij}\}R^{-1}(\alpha_i) \geq b'_i - \eta_i L^{-1}(\alpha_i), \\ x \in \mathbb{E}_+^n. \end{cases} \tag{37}$$

as the equivalent deterministic analogue of the given constraint of model (35).

We prove that $X_W^{\alpha_i} \subseteq X_M^{\alpha_i}$. We take $\tilde{x}=(\tilde{x}_1,\tilde{x}_2,\ldots,\tilde{x}_n) \in X_W^{\alpha_i}$. Since $\tilde{x}$ belongs to $X_W^{\alpha_i}$, both inequalities (37) hold for this value of $\tilde{x}$. We show that both inequalities (36) hold simultaneously for $\tilde{x}$ as well. Thus, we prove that $\tilde{x} \in X_M^{\alpha_i}$. If this is true for any $x \in X_W^{\alpha_i}$, the necessary condition for $X_W^{\alpha_i}$ to be included in $X_M^{\alpha_i}$ is met.

For the fixed $\tilde{x}$, the function $\max\limits_{j=1,\ldots,n}\{\tilde{x}_j\eta_{ij}\}$ takes the particular value $\tilde{x}_s\eta_{is}$. We consider the left-hand side of the first inequality of system of equations (37) for this $\tilde{x}$

$$\sum_{j=1}^{n} a'_{ij}\tilde{x}_j - \tilde{x}_s\eta_{is}L^{-1}(\alpha_i) =$$

$$= a'_{i1}\tilde{x}_1 + a'_{i2}\tilde{x}_2 + \ldots + a'_{is}\tilde{x}_s + \ldots + a'_{in}\tilde{x}_n - \tilde{x}_s\eta_{is}L^{-1}(\alpha_i) =$$

$$= a'_{i1}\tilde{x}_1 + a'_{i2}\tilde{x}_2 + \ldots + \tilde{x}_s(a'_{is} - \eta_{is}L^{-1}(\alpha_i)) + \ldots + a'_{in}\tilde{x}_n.$$

It is not difficult to see what the expression in parentheses means

$$a'_{is} - \eta_{is}L^{-1}(\alpha_i) = t \Rightarrow$$

$$\Rightarrow \frac{a'_{is}-t}{\eta_{is}} = L^{-1}(\alpha_i) \Rightarrow L\left(\frac{a'_{is}-t}{\eta_{is}}\right) = \alpha_i \Longrightarrow t = a_{is}^{-}(\alpha_i).$$

In other words, t is the left boundary of the α_i-level set of the fuzzy variable $a_{is}(\gamma)$. Similarly, we can illustrate what the value in the right-hand side of the first inequality of system (37) means

$$b''_i + \beta_i R^{-1}(\alpha_i) = p \Rightarrow$$

$$\Rightarrow \frac{p-b''_i}{\beta_i} = R^{-1}(\alpha_i) \Rightarrow R\left(\frac{p-b''_i}{\beta_i}\right) = \alpha_i \Longrightarrow p = b_i^{+}(\alpha_i).$$

Taking into account that $a_{ij}^{-}(\alpha_i) \le a'_{ij}$ and $a_{ij}^{+}(\alpha_i) \ge a''_{ij}$ (the left boundary of the α-level set of the fuzzy number does not exceed its left modal value, and the right boundary of the α-level set of the fuzzy number is not less than its right modal value) and $x \in \mathbb{E}_+^N$, we obtain

$$b_i^{+}(\alpha_i) = b''_i + \beta_i R^{-1}(\alpha_i) \ge$$

$$\ge \underbrace{a'_{i1}\tilde{x}_1}_{\ge a_{i1}^{-}(\alpha_i)\tilde{x}_1} + \underbrace{a'_{i2}\tilde{x}_2}_{\ge a_{i2}^{-}(\alpha_i)\tilde{x}_2} + \ldots + \underbrace{\tilde{x}_s(a'_{is} - \eta_{is}L^{-1}(\alpha_i))}_{=a_{is}^{-}(\alpha_i)\tilde{x}_s} + \ldots + \underbrace{a'_{in}\tilde{x}_n}_{\ge a_{in}^{-}(\alpha_i)\tilde{x}_n} \ge$$

$$\ge \sum_{j=1}^{n} a_{ij}^{-}(\alpha_i)x_j.$$

We prove the similar for the second inequality of system (37)

$$b_i^{-}(\alpha_i) = b'_i - \eta_i L^{-1}(\alpha_i) \le$$

$$\le \underbrace{a''_{i1}\tilde{x}_1}_{\le a_{i1}^{+}(\alpha_i)\tilde{x}_1} + \underbrace{a''_{i2}\tilde{x}_2}_{\le a_{i2}^{+}(\alpha_i)\tilde{x}_2} + \ldots + \underbrace{\tilde{x}_s(a''_{is} + \beta_{is}R^{-1}(\alpha_i))}_{=a_{is}^{+}(\alpha_i)\tilde{x}_s} + \ldots + \underbrace{a''_{in}\tilde{x}_n}_{\le a_{in}^{+}(\alpha_i)\tilde{x}_n} \le$$

$$\le \sum_{j=1}^{n} a_{ij}^{+}(\alpha_i)x_j.$$

Thus, any vector $\tilde{x}$ that satisfies the first inequality of system (37) meets the first inequality of system (36) as well, and any vector $\tilde{x}$ for which the second inequality of system (37) holds also meets the second inequality of system (36). Hence, $X_W^{\alpha_i} \subseteq X_M^{\alpha_i}$. It is obvious that

$$X_W^{\alpha_i} \subseteq X_M^{\alpha_i},\ i = \overline{1,m} \Longrightarrow \bigcap_1^m X_W^{\alpha_i} \subseteq \bigcap_1^m X_M^{\alpha_i}.$$

This proves the theorem.

We can prove a similar theorem for the model of criterion as well. We use $F_M^{\alpha_0}(x)$ to denote the range of the criterion, for which (34) holds, given the fact that all $a_{0j}(\gamma), j = \overline{1,n},$, are minrelated, and $F_W^{\alpha_0}(x)$ to denote the range of the criterion that

satisfies model (34), given the fact that all $a_{0j}(\gamma), j = \overline{1,n}$, are mutually T_W-related. We can prove the following theorem.

Theorem 8. $F_W^{\alpha_0}(x) \subseteq F_M^{\alpha_0}(x)$.

Proof. By [25], the equivalent deterministic analogue of model (34) for interminiconnected parameters $a_{0j}(\gamma), j = \overline{1,n}$, has the form

$$\sum_{j=1}^{n} a_{0j}^{-}(\alpha_0)x_j \le k \le \sum_{j=1}^{n} a_{0j}^{+}(\alpha_0)x_j, \tag{38}$$

and by Theorem 3, the equivalent deterministic analogue of model (34) for mutually T_W-related parameters $a_{0j}(\gamma), j = \overline{1,n}$ has the form

$$\sum_{j=1}^{n} a'_{0j}x_j - \max_{j=1,\dots,n}\{x_j\eta_{0j}\}L^{-1}(\alpha_0) \le k \le \sum_{j=1}^{n} a''_{0j}x_j + \max_{j=1,\dots,n}\{x_j\beta_{0j}\}R^{-1}(\alpha_0). \tag{39}$$

Following the line of reasoning similar to Theorem 7, we can show that any vector x that satisfies system (39) satisfies system (38) as well. In other words, the system of inequalities holds

$$\begin{cases} \sum\limits_{j=1}^{n} a_{0j}^{-}(\alpha_0)x_j \le \sum\limits_{j=1}^{n} a'_{0j}x_j - \max\limits_{j=1,\dots,n}\{x_j\eta_{0j}\}L^{-1}(\alpha_0) \le k, \\ \sum\limits_{j=1}^{n} a_{0j}^{+}(\alpha_0)x_j \ge \sum\limits_{j=1}^{n} a''_{0j}x_j + \max\limits_{j=1,\dots,n}\{x_j\beta_{0j}\}R^{-1}(\alpha_0) \ge k. \end{cases}$$

This proves the theorem.

Let $\tilde{F}_M^{\alpha_0}$ and $\tilde{F}_W^{\alpha_0}$ be the optimal values of problem (33)–(35) for minrelated parameters and T_W-related parameters, respectively. We have the following corollary from Theorem 8.

Corollary 1. $\tilde{F}_W^{\alpha_0} \le \tilde{F}_M^{\alpha_0}$.

6.3 Study of Possibilistic Optimization Tasks in the Context of Possibility/Necessity

Feasible region of deterministic equivalent analogue of the optimization task (1)-(3) in the hypotheses of Theorem 4 is defined by corresponding constraints model as a convex set.

Indeed, it is easy to see that for any $i = \overline{1,m}$ function

$$\sum_{j=1}^{n} a''_{ij}x_j + \max_{j=1,\dots,n}\{x_j\beta_{ij}\}R^{-1}(1-\alpha_i)$$

is convex, because it is represented by sum of linear and convex functions. Thus ith constraint defines a convex set. Intersection of convex sets is also a convex set.

This result is important from the point of view of comparative study of equivalent deterministic analogues of the task (1)-(3) in the cases of possibility and necessity measures. As it was shown in the previous section when $\sigma = \pi$ the corresponding constraints model forms a non-convex set.

We use $X_W^{\alpha}(\nu)$ to denote the feasible region defined by the following constrints model

$$\begin{cases} \nu\{f_i(x,\gamma) \le 0\} \ge \alpha_i,\ i = \overline{1,m}, \\ x \in \mathbb{E}_+^n. \end{cases}$$

We use $X_W^{\alpha}(\pi)$ to denote the feasible region defined by

$$\begin{cases} \pi\{f_i(x,\gamma) \le 0\} \ge \alpha_i,\ i = \overline{1,m}, \\ x \in \mathbb{E}_+^n. \end{cases}$$

With the assumptions being made the following theorem holds.

Theorem 9. *Let* $\alpha_i = 0.5,\ i = \overline{1,m}$. *Then*

$$X_W^{\alpha}(\nu) \subseteq X_W^{\alpha}(\pi).$$

Proof. Let $\bar{x} \in X_W^{\alpha}(\nu)$. We show that $\bar{x} \in X_W^{\alpha}(\pi)$. Indeed, because of $\bar{x} \in X_W^{\alpha}(\nu)$, we have that, when switching to equivalent deterministic analogue from the Theorem 4, it turns out that $\forall i = \overline{1,m}$:

$$\sum_{j=1}^{n} a''_{ij}\bar{x}_j + \max_{j=1,\ldots,n}\{\bar{x}_j\beta_{ij}\}R^{-1}(0.5) \le b'_i - \eta_i L^{-1}(0.5).$$

But then $\forall i = \overline{1,m}$ the following inequality holds:

$$\sum_{j=1}^{n} a'_{ij}\bar{x}_j - \max_{j=1,\ldots,n}\{\bar{x}_j\beta_{ij}\}L^{-1}(0.5) \le b''_i + \eta_i R^{-1}(0.5).$$

It is obvious that because $L^{-1}(0.5) = R^{-1}(0.5)$, $\bar{x}$ belongs to intersection of regions defined by later inequalitites. This intersection forms the region $X_W^{\alpha}(\pi)$ as it was shown in the Theorem 7, i.e. $\bar{x} \in X_W^{\alpha}(\pi)$. The theorem is proved.

Remark 2. Threshold level 0.5 accepted in the tasks of stochastic programming [2], as well as in the tasks of possibilistic programming.

7 Conclusion

We performed a comparative study of the models of possibilistic optimization, depending on t-norm that describes interaction (relatedness) of fuzzy parameters. In the long run, it helps "control fuzziness" in optimization and decision-making, which counts for practice. Speaking of further development of this issue, we find it interesting to find methods that will allow choosing appropriate t-norm for modelling fuzzy parameters interaction.

References

1. Dubois, D., Prade, H.: Possibility theory. Tr. from french. Moskva. Radio i svyaz (1990) (in Russian)
2. Ermolyev, Y.M.: Methods of stochastic programming. — M.: Nauka, in Russian (1976) (in Russian)
3. Soldatenko, I.S.: On weighted sum of mutually T_W-related fuzzy variables. Vestnik of Tver State University. Applied Mathematics 5(33), 63–77 (2007) (in Russian)
4. Yazenin, A.V.: To the problem of fuzzy goal achievement level maximization. Izvestia RAN. Theory and Systems for Control 4, 120–123 (1999) (in Russian)
5. Yazenin, A.V.: Fuzzy variables and fuzzy mathematical programming. In: Proceedings of inter-republican scientific conference "Models of alternatives choosing in a fuzzy environment", pp. 57–59. Riga polytechnical institute, Riga (1984) (in russian)
6. Yazenin, A.V.: Models of fuzzy mathematical programming. In: Proceedings of inter-republican scientific conference "Models of alternatives choosing in a fuzzy environment", pp. 51–53. Riga polytechnical institute, Riga (1984)
7. Yazenin, A.V.: Fuzzy mathematical programming. Kalinin State University, Kalinin (1986) (in Russian)
8. De Baets, B., Marková-Stupňanová, A.: Analytical expressions for the addition of fuzzy intervals. Fuzzy Sets and Systems 91, 203–213 (1997)
9. Dubois, D., Prade, H.: Additions of interactive fuzzy numbers. IEEE Trans. Automat. Control 26, 926–936 (1981)
10. Dubois, D., Prade, H.: Fuzzy numbers: an overview. In: Bezdek, J. (ed.) Analysis of Fuzzy Information, vol. 2, pp. 3–39. CRC-Press, Boca Raton (1988)
11. Dubois, D., Prade, H.: Fuzzy Sets and Systems: Theory and Applications. Academic Press, New York (1980)
12. Hong, D.H.: Parameter estimations of mutually T-related fuzzy variables. Fuzzy Sets and Systems 123, 63–71 (2001)
13. Klement, E.P., Mesiar, R., Pap, E.: Triangular norms. Position paper I: basic analytical and algebraic properties. Fuzzy Sets and Systems 143, 5–26 (2004)
14. Klement, E.P., Mesiar, R., Pap, E.: Triangular norms. Position paper II: general constructions and parameterized families. Fuzzy Sets and Systems 145, 411–438 (2004)
15. Klement, E.P., Mesiar, R., Pap, E.: Triangular norms. Position paper III: continuous t-norms. Fuzzy Sets and Systems 145, 439–454 (2004)
16. Liu, B.: Uncertain programming. Wiley and Sons, New York (1999)
17. Mesiar, R.: A note to the T-sum of L − R fuzzy numbers. Fuzzy Sets and Systems 79, 259–261 (1996)
18. Mesiar, R.: Computation of $L-R$-fuzzy numbers. In: Proc. 5th Internat. Workshop on Current Issues in Fuzzy Technologies, Trento, Italy, pp. 165–176 (1995)
19. Mesiar, R.: Shape preserving additions of fuzzy intervals. Fuzzy Sets and Systems 86, 73–78 (1997)
20. Mesiar, R.: Triangular-norm-based addition of fuzzy intervals. Fuzzy Sets and Systems 91, 231–237 (1997)
21. Nahmias, S.: Fuzzy Variables. Fuzzy Sets and Systems 1, 97–110 (1978)
22. Nguyen, H.T., Walker, E.A.: A first cours in fuzzy logic. CRC Press, Boca Raton (1997)

23. Rao, M., Rashed, A.: Some comments on fuzzy variables. Fuzzy Sets and Systems 6, 285–292 (1981)
24. Yazenin, A.V.: Fuzzy and stochastic programming. Fuzzy Sets and Systems 22, 171–180 (1987)
25. Yazenin, A.V.: On the problem of possibilistic optimization. Fuzzy Sets and Systems 81, 133–140 (1996)
26. Yazenin, A.V., Wagenknecht, M.: Possibilistic optimization. Brandenburgische Technische Universitat, Cottbus, Germany (1996)
27. Zadeh, L.A.: Fuzzy sets as a basis for a theory of possibility. Fuzzy Sets and Systems 1, 3–28 (1978)

A Parametrized Model for Optimization with Mixed Fuzzy and Possibilistic Uncertainty

Elizabeth Untiedt

Abstract. Fuzzy and possibilistic uncertainty are very closely related, and sometimes coexist in optimization under uncertainty problems. Fuzzy uncertainty in mathematical programming problems typically represents flexibility on the part of the decision maker. On the other hand, possibilistic uncertainty generally expresses a lack of information about the values the parameters will assume.

Several models for mixed fuzzy and possibilistic programming problems have previously been published. The semantic interpretation of these models, however, is of questionable value. The mixed models in the literature find solutions in which the fuzzy uncertainty (or flexibility) and the possibilistic uncertainty (or lack of confidence in the outcome) are held to the same levels.

This chapter proposes a new mixed model which allows a trade-off between fuzzy and possibilistic uncertainty. This trade-off corresponds to a semantic interpretations consistent with human decision-making. The new model shares characteristics with multi-objective programming and Markowitz models. Model structure, semantic justification, and solution approaches are covered.

1 Introduction

In the application of optimization theory, parameters are often not known with certainty. Fuzzy and possibilistic uncertainty are very closely related, and sometimes coexist in optimization under uncertainty problems. Fuzzy uncertainty in mathematical programming problems typically represents flexibility on the part of the decision maker. On the other hand, possibilistic uncertainty generally expresses a lack of information about the value a parameter will assume.

Several models for mixed fuzzy and possibilistic programming problems have been previously published. The semantic interpretation of these models,

W.A. Lodwick & J. Kacprzyk (Eds.): Fuzzy Optimization, STUDFUZZ 254, pp. 193–208.
springerlink.com

however, is of questionable value. This chapter proposes a semantic interpretation of the mixed fuzzy and possibilistic linear programming problem that is not fully addressed by any of the existing models. The newe *parameterized model* shares characteristics with multi-objective problems and programming and Markowitz models.

In Section two of this chapter, we will provide background definitions, define the problem, and review the current state of the literature on fuzzy and possibilistic programming problems. Chepter 3 presents tha model, along with its derivation and justification. Chapter 4 examines problem structure, solution methods, and an industrial strength application.

2 Background

Fuzzy uncertainty describes vagueness, or a softening of the concept of belonging to a set. A fuzzy set $\tilde{X}$ is defined by its membership function, $\mu_{\tilde{X}}$. An element s has its degree of membership in $\tilde{X}$ described by $\mu_{\tilde{X}}(s)$, with 1 indicating full membership, 0 indicating full non-membership, and numbers between 0 and 1 indicating partial membership. In decision making applications, a fuzzy set may be used to indicate flexibility on the part of the decision maker. A goal or constraint may be softened via a fuzzy inequality.

Possibilistic uncertainty, on the other hand, describes ambiguity. A possibilistic variable is not random– its value is pre-determined, but is not known with certainty. The likelihood that the value of a possibilistic variable lies in a particular interval is described by a possibility distribution. Klir relates possibility distributions to fuzzy sets as follows [4].

> Given a universe Y, let y be a variable which takes on values in Y. Now let F be a fuzzy set on Y, then let $F(x)$ describe the extent to which x is a member of F. Then if we say, "y is F", $F(x)$ for each x in Y is the possibility that x is F.

Though they are related, possibilistic uncertainty and fuzzy sets are used to represent very different things. In optimization problems, a possibilistic variable is used to indicate a parameter whose value is fixed, but is not known with certainty, (for example: a measurement with a margin of error). A possibility distribution might also be used to reflect a decision maker's estimation of the distribution of a random variable when the exact distribution is too costly or impossible to determine.

2.1 Semantics of the Mixed Fuzzy and Possibilistic Linear Programming Problem

Occasionally, fuzzy and possibilistic uncertainty occur in the same optimization problem. Consider the linear program (LP):

$$\begin{aligned} \max\ & c^{\mathrm{T}}x \\ \text{subject to}\ & Ax \leq b \\ & x \geq 0. \end{aligned} \tag{1}$$

Possibilistic uncertainty (denoted in this chapter with a "hat") can occur in the parameters represented by $\hat{c}$, $\hat{b}$, and/or $\hat{A}$. (This happens when the value of the parameters is fixed, but the decision maker has incomplete information about its value.)

On the other hand, fuzzy uncertainty (denoted in this chapter with a "tilde") can occur in any of the parameters, $\tilde{A}$, $\tilde{b}$, and/or $\tilde{c}$. This happens when the values of the parameters are not sharp. Consider c_i, which represents the cost or value of x_i, in the problem. When there is a fuzzy interval of values which, to varying degrees, represent the cost, $\tilde{c}_i$ is fuzzy. Fuzzy parameters appear in optimization problems on occasion, but most fuzzy uncertainty occurs in the inequality. A fuzzy less than constraint ($\tilde{\leq}$) can be interpreted as "approximately less than." When this uncertainty represents a willingness on the part of the decision-maker to bend the constraints, optimization problem with fuzzy inequalities are sometimes called "flexible programs," and the constraints are called "soft constraints."

A common mixed fuzzy and possibilistic linear program, then, might assume the following form:

$$\begin{aligned} \max\ & \hat{c}^{\mathrm{T}}x \\ \text{subject to}\ & \hat{A}x \ \tilde{\leq}\ \hat{b} \\ & x \geq 0. \end{aligned} \tag{2}$$

An element's membership in a fuzzy set (or the degree to which a fuzzy inequality is satisfied) is quantified by the membership function, and represented by $\alpha \in [0, 1]$. The likelihood that an interval contains a possibilistic variable is quantified by the possibility distribution, but is also represented by an $\alpha \in [0, 1]$. The fuzzy α and the possibilistic α, however, mean very different things. The fuzzy α represents the level at which the decision-maker's requirements are satisfied. The possibilistic α, on the other hand, represents the likelihood that the parameters will take on values which will result in that level of satisfaction. With that in mind, let us examine some previously published approaches for mixed fuzzy and possibilistic linear programming problems.

2.2 Existing Models for Mixed Fuzzy and Possibilistic Programming

Delgado, Verdegay, and Villa [5] propose the following formulation for dealing with ambiguity in the constraint coefficients and right-hand sides, as well as vagueness in the inequality relationship:

$$\begin{aligned} &\text{maximize} && c^{\mathrm{T}}x \\ &\text{subject to} && \hat{A}x \tilde{\leq} \hat{b} \\ & && x \geq 0. \end{aligned} \tag{3}$$

In addition to (3), membership functions $\pi_{a_{ij}}$ are defined for the possible values of each possibilistic element of $\hat{A}$, membership functions π_{b_i} are defined for the possible values of each possibilistic element of $\hat{b}$, and membership function μ_i gives the degree to which the fuzzy constraint i is satisfied. Stated another way, μ_i is the membership function of the fuzzy inequality. Recall that the uncertainty in the $\tilde{a}_{ij}$s and the $\tilde{b}_i$s is due to ambiguity concerning the actual value of the parameter, while the uncertainty in the $\tilde{\leq}_i$s is due to the decision maker's flexibility regarding the necessity of satisfying the constraints in full.

Delgado, *et al.* solve the problem parametrically on α. For each $\alpha \in [0, 1]$ (or practically speaking, for a finite subset of $\alpha \in [0, 1]$), a set of constraints and objective function are produced. The resut is a fuzzy solution to the fuzzy problem. Since a fuzzy solution cannot be implemented, the decision maker must select an α and implement the corresponding solution. *This implementation has a likelihood α of satisfying the constraints of the problem at a level α.* The likelihood that the constraints will be satisfied and the level at which they are satisfied are two completely separate concepts, but this model holds both to the same α level.

Another mixed formulation is what Inuiguchi [3] refers to as the "fuzzy robust programming" problem [1, 8]. This is a mathematical program with possibilistic constraint coefficients $\hat{a}_{ij}$ that satisfy fuzzy constraints, $\tilde{b}_i$ as follows:

$$\max c^{\mathrm{T}}x \tag{4}$$

$$\text{subject to } \hat{a}_i' x' \subseteq \tilde{b}_i \tag{5}$$

$$x' = (1, x^t)t \geq 0.$$

Zadeh [9] defines the set-inclusion relation $\tilde{M} \subseteq \tilde{N}$ as $\mu_{\tilde{M}}(r) \leq \mu_{\tilde{N}}(r)$ for all r. Robust programming interprets the set-inclusive constraint to mean that the region in which $\tilde{a}_i' x'$ can possibly occur is restricted to $\tilde{b}_i$, a region which is tolerable to the decision maker. Therefore, the left side of (5) is possibilistic, and the right side is fuzzy.

Negoita [8] defines the fuzzy right hand side as follows:

$$\tilde{b}_i = \{r \in \mathcal{R} | r \geq b_i\}. \tag{6}$$

As a result, we can interpret $\hat{a}_i' x' \subseteq \tilde{b}_i$ as an extension of an inequality constraint. The set-inclusive constraint (5) is reduced to

$$a_i^+(\alpha)x \leq b_i^+(\alpha) \tag{7}$$

$$a_i^-(\alpha)x \geq b_i^-(\alpha)$$
$$\text{for all } \alpha \in (0,1].$$

If the membership functions are linear, it suffices to satisfy the constraints for $\alpha = 1$ and for $\alpha = \epsilon$, where ϵ is close to zero, since all $\alpha \in (\epsilon, 1)$ will be satisfied by interpolation. If the membership functions are not linear, however, we have an infinitely constrained problem. If we abide by Negoita's definition of $\tilde{b}$ ([6]), $b_i^+ = \infty$ for all values of α, so we can drop the first constraint in ([7]). Nonetheless, we still have an infinitely constrained program, with a constraint for each value of $\alpha \in (0,1]$.

Consider the semantics of the second inequality in [6]. It requires that the α level of a possibilistic $\hat{a}_i$ multiplied by x be greater than the α level of a fuzzy $\tilde{b}_i$. In other words, there is a likelihood greater than or equal to α that a_i has a value which leads to constraint satisfaction at a level α. Like Delgado's model, the fuzzy robust model holds two very different types of uncertainty to the same α level. That is to say, the optimal implementaion will have a likelihood α of satisfying the constraints at a level α. The decision maker might want to allow a lower level of constraint satisfaction in order to have a greater guarantee of his/her result, or vice versa.

This leads us to ask the following questions. What is it that makes this model "robust"? And how can robust optimization theory inform a more practical approach to solving the mixed optimization under uncertainty problem?

2.3 Robust Optimization

The goal of robust optimization, which has its roots in stochastic optimization, is to produce a solution whose quality will withstand a wide variety of parameter realizations. Robust optimization seeks to mitigate the effects of uncertainty rather than merely anticipating it. Hence, robustness reflects a tendency to hedge against uncertainty, sacrificing some performance in order to avoid excessive volatility [7]. Robust formulations are designed to yield solutions that are less sensitive to model data than classical mathematical programming formulations. Robust programs fall into two broad categories– *solution robust* programs seek to minimize variance in solution optimality, while *model robust* programs aim to decrease variance in feasibility.

The robust fuzzy optimization model in [7] is called "robust" (in the model robust sense) because it seeks a solution which guarantees compliance with constraints at every possibility level (every α level). Unfortunately, at lower possibility levels, it is held to lower standards of constraint compliance, bringing into question its "robust" designation. To shed light on this possible misnomer, let us examine a classic robust model– the Markowitz model, which is robust in the solution robust sense.

2.4 Markowitz Model

In 1952, Markowitz [6] proposed a novel approach to financial portfolio optimization. He makes the case that a traditional linear programming approach to portfolio optimization will never prefer a diversified portfolio to an undiversified portfolio. He observes that simply diversifying among top return solutions will not result in a reliable portfolio, since the returns are too interrelated for the law of large numbers to apply. He proposes that both maximizing the expected value of the return and minimizing the historical variance (risk) are valid objectives. An efficient combination, then, is one which has the minimum risk for a return greater than or equal to a given level; or one which has the maximum return for a risk less than or equal to a given level. The decision maker can move among these efficient combinations, or along the efficient frontier, according to hisor her degree of risk aversion.

3 Main Results

3.1 Model Concept

In the spirit of the Markowitz model, we wish to allow a trade-off between the potential reward of the outcome and the reliability of the outcome, with the weights of the two competing objectives determined by the decision maker's risk aversion. The desire is to obtain an objective function like the following:

$$\text{maximize}: \ reward \ + \ (\lambda \times reliability), \tag{8}$$

where λ is a parameter indicating risk aversion.

The reward variable is the α-level associated with the fuzzy constraints and goal(s). It tells the decision maker how satisfactory the solution is. The reliability variable is the α-level associated with the possibilistic parameters. It tells the decision maker how likely it is that the solution will actually be satisfactory. To avoid confusion, let us refer to the fuzzy constraint membership level as α and the possibilistic parameter membership level as β.

In addition, let $\lambda \in [0, 1]$ be an indicator of the decision maker's valuation of reward and risk-avoidance, with 0 indicating that the decision maker cares exclusively about the reward, and 1 indicating that only risk avoidance is important. Using this notation, the desired objective is

$$\max \ (1 - \lambda)\alpha \ + \ \lambda\beta. \tag{9}$$

There is a frontier, which we call *the efficient frontier* in accordance with the literature, along which a trade-off occurs. The choice of λ determines which solution along the frontier is chosen.

3.2 Model Implementation

Suppose we begin with the mixed problem:

$$\begin{aligned} \max\ & \hat{c}^{\mathrm{T}} x \\ \text{subject to}\ & \hat{A}x \ \tilde{\leq}\ b \\ & x \ \geq 0. \end{aligned} \tag{10}$$

Let us for the moment ignore the possibilistic parameters and deal with the fuzzy constraints according to Bellman and Zadeh. We first introduce a goal, g, for the objective function value and state the objective function and a fuzzy goal, $\hat{c}^{\mathrm{T}} x \ \tilde{\geq}\ g$. Now together the goal and the constraints form the decision space, and we wish to maximize the α-level at which the least satisfied constraint or goal is met. Allow (as a slight abuse of notation) $\tilde{\leq}_\alpha$ to denote the α-level at which a constraint or goal is met from a pessimistic point of view. The problem is now,

$$\begin{aligned} \max\ & \alpha \\ \text{subject to}\ & -\hat{c}^{\mathrm{T}} x \ \tilde{\leq}_\alpha\ -g \\ & \hat{A}x \ \tilde{\leq}_\alpha\ b \\ & x, \alpha \ \geq\ 0. \end{aligned} \tag{11}$$

(12)

Now let $\hat{c}_\beta$ denote the right end-point of the $(1-\beta)$-cut of $-\hat{c}$, and $\hat{a}_{ij\beta}$ denote the right end-point of the $(1-\beta)$-cut of $\hat{a}_{ij}$ as illustrated in Figure 1.

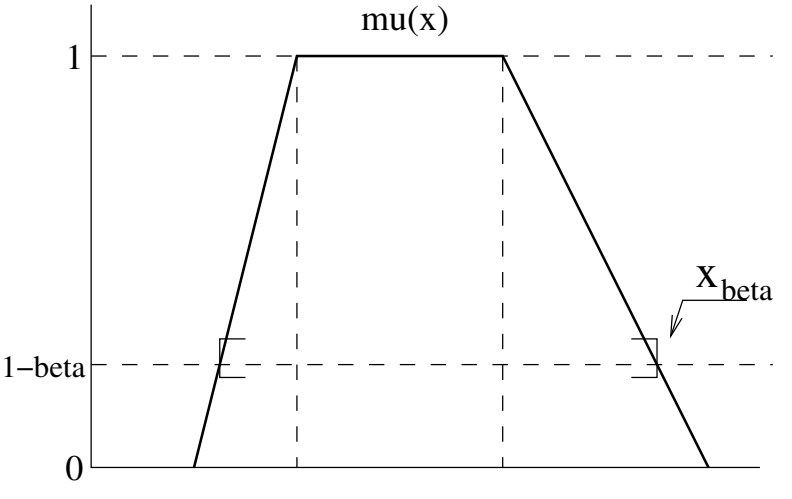


Fig. 1 The right end-point on the $1-\beta$-cut of the $\hat{c}$

We can now complete our formulation as follows:

$$\begin{array}{lrll} & \max\ (1-\lambda)\alpha & + & \lambda\beta \\ \text{subject to} & -\hat{c}_\beta^{\mathrm{T}} x & & \tilde{\leq}_\alpha -g \\ & \hat{A}_\beta x & & \tilde{\leq}_\alpha\ b \\ & x & & \geq 0 \\ & \alpha, \beta & & \in [0,1]. \end{array} \tag{13}$$

The constraints are functions of x, α, and β, and although the objective function is linear (it does not have to be– some aggregation other than addition could have been chosen for the α and β terms), we shall soon see that the variables are quite entangled in the constraints, resulting in a non-linear formulation.

Let us suppose, for the sake of simplicity, that the possibility distribution for each a_{ij} (and b_i) is trapezoidal, with support (w_{ij}, z_{ij}) (or (w_i, z_i)), and core (u_{ij}, v_{ij}) (or (x_i, y_i)), as illustrated in Figure 2.

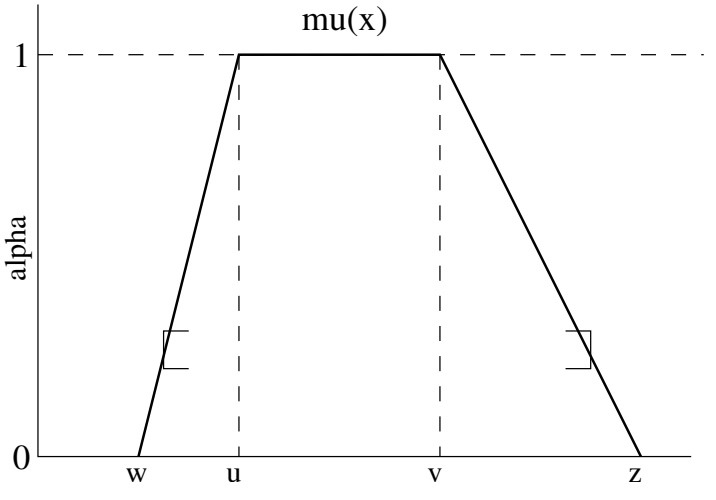


Fig. 2 Trapezoidal fuzzy number

Also, suppose that we have a single goal, and that the membership functions for the fuzzy constraint and goal are trapezoidal, with d_0 denoting the maximum acceptable deviation from the goal, and d_i denoting the maximum acceptable deviation from constraint i, as illustrated in Figure 3.

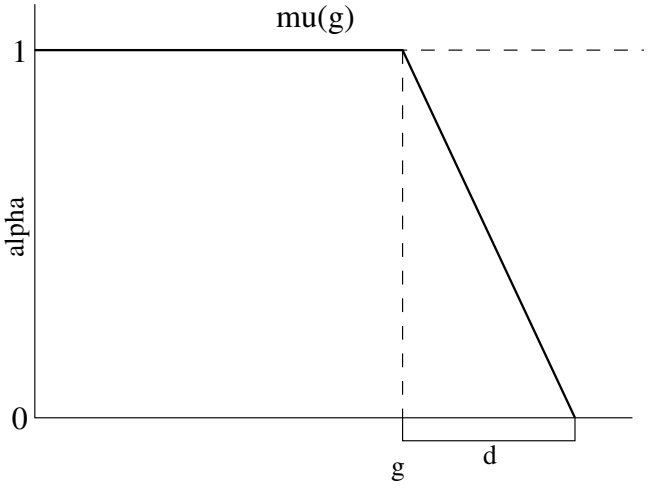


Fig. 3 Trapezoidal fuzzy goal

Then, incorporating fuzziness in the tradition of Zimmerman [10], we get

$$\begin{aligned} \max \quad & (1-\lambda)\alpha \ + \ \lambda\beta \\ \text{subject to} \quad \alpha \ & \leq -\frac{g}{d_0} + \sum_j \frac{c_{j,\beta}}{d_0} x_j \\ \alpha \ & \leq \frac{b_i}{d_i} - \sum_j \frac{a_{ij,\beta}}{d_i} x_j, \quad \forall i \\ x \ & \geq 0 \\ \alpha, \beta & \in [0,1]. \end{aligned} \tag{14}$$

And incorporating a pessimistic view of possibility (if the problem has no solution, we return again with an optimistic point of view), we get

$$\begin{aligned} \max \quad & (1-\lambda)\alpha \ + \ \lambda\beta \\ \text{subject to} \quad \alpha \ & \leq -\frac{g}{d_0} + \sum_j \frac{u_j}{d_0} x_j + \sum_j \frac{u_j - w_j}{d_0} x_j \beta \\ \alpha \ & \leq \frac{b_i}{d_i} - \sum_j \frac{v_{ij}}{d_i} x_j - \sum_j \frac{z_{ij} - v_{ij}}{d_i} x_j \beta \\ x \ & \geq 0 \\ \alpha, \beta & \in [0,1]. \end{aligned} \tag{15}$$

3.3 A Simple Example

Let us examine a simple, two-variable toy problem as an example. For the sake of comparison, consider the numerical example treated by Delgado, et al. [5]:

$$\begin{aligned} \text{maximize } z &= 5x_1 + 6x_2 \\ \text{subject to} \quad & \hat{3}x_1 + \hat{4}x_2 \tilde{\leq} \hat{18}, \\ & \hat{2}x_1 + \hat{1}x_2 \tilde{\leq} \hat{7} \\ & x_1, x_2 \geq 0, \end{aligned} \tag{16}$$

where

$$\begin{aligned} &\hat{3} = (3, 2, 4),\ \hat{4} = (4, 2.5, 5.5),\ \hat{18} = (18, 16, 19) \\ &\hat{2} = (2, 1, 3),\ \ \hat{1} = (1, 0.5, 2),\ \ \hat{7} = (7, 6, 9), \end{aligned} \tag{17}$$

and the maximum violation of the first constraint (our d_1) is 3, while the maimum allowable violation of the second constraint (our d_2) is 1.

Our parametrized model is not yet developed to handle possibilistic right-hand sides, so let us modify Delgado, *et al*'s problem by making the right-hand sides crisp. Also, the parametrized mixed formulation requires that the objective function be reformulated as a goal. The solution to the associated crisp problem yields an objective function value of 28, so let the goal be 23, with maximum violation (d_0) of 5. Finally, introduce possibilistic uncertainty in the objective function coefficients.

The toy problem to solve is then:

$$\begin{aligned} \text{maximize } z &= \hat{5}x_1 + \hat{6}x_2 \\ \text{subject to} \quad & \hat{3}x_1 + \hat{4}x_2 \tilde{\leq} 18, \\ & \hat{2}x_1 + \hat{1}x_2 \tilde{\leq} 7 \\ & x_1, x_2 \geq 0, \end{aligned} \tag{18}$$

where

$$\begin{aligned} &\hat{5} = (5, 4, 6),\ \hat{6} = (6, 5, 7)\hat{3} = (3, 2, 4),\ \hat{4} = (4, 2.5, 5.5), \\ &\hat{2} = (2, 1, 3), \qquad \hat{1} = (1, 0.5, 2), \\ &\quad d_1 = 3, \qquad\quad d_2 = 1 \\ &\quad g = 28, \qquad\quad d_0 = 5. \end{aligned}$$

Reformulating (19) according to (15) yields:

$$\begin{aligned} \max \quad & (1-\lambda)\alpha \ + \ \lambda\beta \\ \text{subject to} \quad \alpha \ & \leq -\frac{23}{5} + x_1 + \frac{6}{5}x_2 + \frac{1}{5}x_1\beta + \frac{1}{5}x_2\beta \\ \alpha \ & \leq \frac{18}{3} - x_1 - \frac{4}{3}x_2 - \frac{1}{3}x_1\beta - \frac{1}{6}x_2\beta \\ \alpha \ & \leq 7 - 2x_1 - 1x_2 - 1x_1\beta - 1x_2\beta \\ x \ & \geq 0 \\ \alpha, \beta & \in [0,1]. \end{aligned} \tag{19}$$

This toy problem was formulated in GAMS for $\lambda = 0, \lambda = .5$ and $\lambda = 1$ and solved with the non-linear programming solver. The results are summarized in Table 1.

Table 1 GAMS results for toy problem. Risk aversion is represented by λ, the reformulated objective by $z = (1-\lambda)\alpha + \lambda\beta$, and the original (crisp) objective by o. Reliability, represented by β ranges between 0 and 1, and reward, represented by α also ranges between 0 and 1.

λ	z	o	α	β	x_1	x_2
1	1	19.0	0	1	.2	3.1
0.5	0.6	22.2	0.49	0.7	0	3.8
0	0.5	25.6	0.52	0	1.9	2.7

It is clear, from the results, that as risk aversion decreases, the objective function level that can possibly be attained (represented by o) increases. However, the certainty of attaining that level, represented by β, decreases. This is the desired and expected result for the model. In fact, in order for the model to make sense, we need to know that α and β are always inversely related. This is in contrast to other mixed models, in which α and β are equal.

4 Problem Structure and Its Relation to Solution Methods

Many optimization problems have special structures which are exploited by efficient solution algorithms. In the search for practical solution methods to (15), we first evaluate the structure of the problem.

The last terms in each of the constraints in (15) contain βx, so the constraints are bi-linear. If the possibility distributions were non-linear (i.e. not trapezoidal or triangular), the system would be non-linear rather than bi-linear. Most optimization models for fuzzy or possibilistic uncertainty (that

have linear possibility distributions or membership functions, as we assume here) are linear programs. The fact that the fuzzy robust model results in a bi-linear program is a distinct disadvantage.

On the bright side, there are some simplicities to the model that may result in a specialized solution method. The objective function is linear. Also, the non-linearity in the constraints results from the product of a function of β taken independently with each component of x. There are no mixed terms, which would lead to quadratic constraints. Unfortunately, the bi-linear constraints (in the simplest case) form a non-convex feasible region, which makes for a very hard optimization problem.

One possibility is to try to convert the fuzzy robust model into a bi-linear program by adding the constraints to the objective function with penalties. The disadvantages of this approach are two-fold. First, the introduction of an auxiliary penalty does not really make sense in the scope of the problem, since our objective in the first place in to minimize a kind of penalty (uncertainty). Second, the problem with penalized constraint violations in the objective function would require sequential solutions, which may result in greater complexity than other non-linear programming methods. For these reasons the sequential solution of the fuzzy robust problem with penalties was not pursued, but may be an avenue for further research.

Convex Programming

We've observed that the fuzzy robust problem is particularly difficult because the βx term results in a non-convex feasible region. The shape of the feasible region is directly related to the trapezoidal shape of the possibility distribution for the possibilistic parameters in the constraint matrix.

However, if the left and right-hand sides of the possibility distributions were not linear, but were defined by sufficiently convex functions, the feasible region would be convex. Specifically, consider the special case in which the left- and right-hand sides of the possibility distributions are each bounded above by $\frac{c}{x}$ for some constant c. Then the constraints will be concave functions bounded above by the linear function cx.

For example, let the right hand side of the possibility function for a parameter a be defined as $a_v + (\frac{2}{\beta+1} - 1)(a_z - a_v)$ (see figure 4). Then a_β (the right end-point of the $(1-\beta)$-cut defined in section 3.2) will be $v + *(\frac{-2}{\beta} - 1)(z - v)$. This leads a problem with constraints of the form

$$Vx + (\frac{-2}{\beta + 1} - 1)(Z - V)x \leq b + (\alpha - 1)d,$$

which are concave. In this situation, convex programming algorithms may be used to solve the problem.

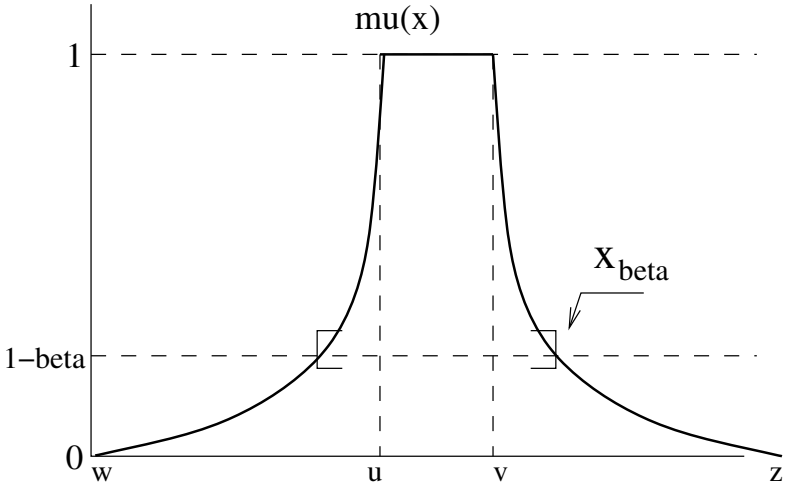


Fig. 4 The right end-point on the $1-\beta$-cut of the $\hat{c}$ when the right-hand-side of the possibility distribution is defined as $a_v + (\frac{2}{\beta+1} - 1)(a_z - a_v)$

Since membership functions are not typically determined empirically, but are chosen to represent the opinion of the decision maker, and often, the convenience of the modeler, this is a significant observation. Modelers often arbitrarily define linear membership functions for their models, but for the fuzzy robust model they could arbitrarily define sufficiently convex membership functions.

4.1 Testing the Model: The Radiation Therapy Problem

The fuzzy robust model with both non-convex and convex feasible regions was tested on the radiation therapy planning problem, using the MATLAB Optimization Toolbox function fmincon which minimizes non-linear constrained problems.

The use of particle beams to treat tumors is called the radiation therapy planning (RTP) problem [2]. Beams of particles are oriented at a variety of angles and with varying intensities to deposit radiation dose (measured as energy/unit mass) to the tumor. The goal is to deposit a tumorcidal dose to the tumor while minimizing damage to surrounding non-tumor tissue.

A *treatment plan* is the identification of a set of beam angles and weights that provides a lethal dose to the tumor cells while sparing healthy tissue, with a resulting dose distribution acceptable to the radiation oncologist. A *dose transfer matrix* A, specific to the patient's geometry, represents how a unit of radiation in beamlet j is deposited in body pixel i. The components of A are determined by the fraction of pixel i which intersects with beamlet j,

attenuated by the distance of the pixel from the place where the beam enters the body. The dose transfer matrix A can be divided into the following: a matrix T which contains dose transfer information to tumor pixels only, matrices C_1 through C_K which contain does transfer information to pixel in critical organs 1 though K, and body matrix B which contains dose transfer information for all non-tumor and non-critical-organ pixels in the body. The variable vector x represents the beamlet intensities, and the right hand side vector b represents the dosage requirements.

The constraints for the crisp (non-fuzzy) formulation of the RTP is

$$\begin{aligned} \text{subject to} \quad B &\leq b_{body} \\ C_1 &\leq b_{C_1} \\ &\ldots \\ C_K &\leq b_{C_K} \\ T &\leq b_{tumor} \\ -T &\leq -b_{tumor}. \end{aligned} \tag{20}$$

To test the fuzzy robust model, we interpret the radiation therapy problem in which the radiation oncologist is flexible regarding the dose limits (fuzziness in the inequality) and the components of the attenuation matrix are based on incomplete information (possibilistic parameters), as below.

$$\max \quad (1-\lambda)\alpha \ + \ \lambda\beta \tag{21}$$

$$\begin{aligned} \text{subject to} \quad (1-\beta)V_{body}x + \beta Z_{body}x &\leq b_{body} + (\alpha-1)d_{body} \\ (1-\beta)V_{C_1}x + \beta Z_{C_1}x &\leq b_{C_1} + (\alpha-1)d_{C_1} \\ &\ldots \\ (1-\beta)V_{C_k}x + \beta Z_{C_k}x &\leq b_{C_k} + (\alpha-1)d_{C_k} \\ (1-\beta)V_{tumor}x + \beta Z_{tumor}x &\leq b_{tumor} + (\alpha-1)d_{tumor} \\ (\beta-1)V_{tumor}x - \beta Z_{tumor}x &\leq -b_{tumor} + (1-\alpha)d_{tumor} \\ x &\geq 0 \\ \alpha, \beta &\in [0,1]. \end{aligned} \tag{22}$$

The fuzzy robust model for the radiation therapy planning problem was solved in MATLAB. The starting point was found by solving a non-fuzzy version of the radiation therapy planning problem using linear programming. The code was tested on an image of one tumor and two critical organs with 64×64 pixel resolution the radiation beam discretized into 10 beamlets. In addition, the model was tested with concave constraints as in (20).

4.2 Testing the Model: Results

MATLAB found a feasible, acceptable solution to the Radiation Therapy problem for all values of λ between 0 and 1. The problem had 72 variables and 997 constraints, which made it a medium-to-large-scale problem. MATLAB prefers to use trust region methods for large scale problem, but because of the non-convex feasible region, it had to use a line-search method. MATLAB required over 20,000 function evaluations to converge to a solution, even though it was close after 2,000. The time taken to solve the fuzzy robust model was an order of magnitude larger than the time taken to solve other mixed fuzzy and possibilistic programming models, all of which had linear programming formulations. With concave constraints, MATLAB still used line-search methods to solve the problem, but required only 6,000 function evaluations to converge to a solution. The solutions found by the concave and non-concave formulations were not identical– but the resultant doses were equally satisfactory.

5 Conclusion

This chapter introduces a model for problems with both fuzzy and possibilistic variables. The model is semantically meaningful, and puts an additional fine-tuning parameter in the hands of the decision maker.

In addition, it introduces the idea of selecting a membership function that will facilitate the solution of the problem. The fact that a slight perturbation in the shape of the distribution of the possibilistic parameters improved model performance appears to be a novel observation.

Avenues for further research include:

- How can the possibilistic right hand side be incorporated into the mixed robust model? Is there a way to simultaneously represent constraint flexibility with α and right-hand-side imprecision with β?
- Both fuzzy and possibilistic intervals are upper-semi-continuous, so varying α-levels imply moving up or down either the left slope (or profile), or the right slope (or profile), but not over the entire interval. Because the decision maker is seeking to minimize risk, the current formulation selects whichever profile represents the pessimistic point of view. Is there a way to appropriately parametrize movement over the entire fuzzy or possibilistic interval so that an optimistic point of view can also be represented?

References

1. Dubois, D., Prade, H.: Systems of linear fuzzy constraints. Fuzzy Sets and Systems 3, 37–48 (1980)
2. Holder, A., Salter, B.: A tutorial on radiation oncology and optimization. In: Greenberg, H. (ed.) Tutorials on Emerging Methodologies and Application in Operations Research, ch. 4 (2004)

3. Inuiguchi, M.: Fuzzy linear programming: What, why and how? Tatra Mountains Mathematical Publications, 123–167 (1997)
4. Klir, G., Yuan, B.: Fuzzy Sets and Fuzzy Logic: Theory and Applications. Prentice-Hall, Englewood Cliffs (1995)
5. Vila, M.A., Delgado, M., Verdegay, J.L.: A general model for fuzzy linear programming. Fuzzy Sets and Systems, 21–29 (1989)
6. Markowitz, H.: Portfolio selection. The Journal of Finance 7(1), 77–91 (1952)
7. Morrison, T.: Personal communication (July 2006)
8. Negoita, C.V.: The current interest in fuzzy optimization. Fuzzy Sets and Systems 6, 261–269 (1981)
9. Zadeh, L.A.: Fuzzy sets. Information and Control 8, 338–353 (1965)
10. Zimmermann, H.J.: Description and optimization of fuzzy systems. International Journal of General Systems 2, 209–215 (1985)

On Solving Optimization Problems with Ordered Average Criteria and Constraints

Włodzimierz Ogryczak and Tomasz Śliwiński

Abstract. The problem of aggregating multiple numerical attributes to form overall measure is of considerable importance in many disciplines. The ordered weighted averaging (OWA) aggregation, introduced by Yager, uses the weights assigned to the ordered values rather than to the specific attributes. This allows one to model various aggregation preferences, preserving simultaneously the impartiality (neutrality) with respect to the individual attributes. However, importance weighted averaging is a central task in multiattribute decision problems of many kinds. It can be achieved with the Weighted OWA (WOWA) aggregation though the importance weights make the WOWA concept much more complicated than the original OWA. In this paper we analyze solution procedures for optimization problems with the ordered average objective functions or constraints. We show that the WOWA aggregation with monotonic preferential weights can be reformulated in a way allowing to introduce linear programming optimization models, similar to the optimization models we developed earlier for the OWA aggregation. Computational efficiency of the proposed models is demonstrated.

Keywords: OWA, WOWA, Optimization, Linear Programming.

1 Introduction

Consider a decision problem defined characterized by m attribute functions $f_i(\mathbf{x})$. That means there is given a feasible set $\mathcal{F} \subset R^q$ of decision vectors $\mathbf{x}$ (vectors of decision variables). The feasible set is usually defined by some constraints on the

Włodzimierz Ogryczak · Tomasz Śliwiński
Warsaw University of Technology, Institute of Control & Computation Engineering,
00-665 Warsaw, Poland
e-mail: wogrycza@ia.pw.edu.pl, tsliwins@ia.pw.edu.pl

W.A. Lodwick & J. Kacprzyk (Eds.): Fuzzy Optimization, STUDFUZZ 254, pp. 209–230.
springerlink.com

decision variables. Further $\mathbf{f}(\mathbf{x}) = (f_1(\mathbf{x}), f_2(\mathbf{x}), \ldots, f_m(\mathbf{x}))$ is a vector function that maps the feasible set $\mathscr{F}$ into the outcome attribute space R^m. In order to make the multiple attribute model operational for the decision support process, one needs to assume some aggregation function for multiple attributes: $a : R^m \to R$. The aggregated attribute values can be either bounded or optimized (maximized or minimized).

The most commonly used aggregation is based on the weighted mean where positive importance weights p_i $(i = 1, \ldots, m)$ are allocated to several attributes

$$A_{\mathbf{p}}(\mathbf{y}) = \sum_{i=1}^{m} y_i p_i \tag{1}$$

The weights are typically normalized to the total 1 ($\sum_{i=1}^{m} p_i = 1$). However, the weighted mean allowing to define the importance of attributes does not allow to model the decision maker's preferences regarding distribution of outcomes. The latter is crucial when aggregating (normalized) uniform achievement criteria like those used in the fuzzy optimization methodologies [29] as well as in the goal programming and the reference point approaches to the multiple criteria decision support [13]. In the stochastic problems uniform objectives may represent various possible values of the same (uncertain) outcome under several scenarios [14].

The preference weights can be effectively introduced with the so-called Ordered Weighted Averaging (OWA) aggregation developed by Yager [25]. In the OWA aggregation the weights are assigned to the ordered values (i.e. to the smallest value, the second smallest and so on) rather than to the specific attributes. Since its introduction, the OWA aggregation has been successfully applied to many fields of decision making [29, 30, 2]. When applying the OWA aggregation to optimization problems with attributes modeled by variables the weighting of the ordered outcome values causes that the OWA operator is nonlinear even for linear programming (LP) formulation of the original constraints and criteria. Yager [26] has shown that the nature of the nonlinearity introduced by the ordering operations allows one to convert the OWA optimization into a mixed integer programming problem. We have shown [18] that the OWA optimization with monotonic weights can be formed as a standard linear program of higher dimension.

The OWA operator allows to model various aggregation functions from the maximum through the arithmetic mean to the minimum. Thus, it enables modeling of various preferences from the optimistic to the pessimistic one. On the other hand, the OWA does not allow to allocate any importance weights to specific attributes. Actually, the weighted mean (1) cannot be expressed in terms of the OWA aggregations.

Importance weighted averaging is a central task in multicriteria decision problems of many kinds, such as selection, classification, object recognition, and information retrieval. Therefore, several attempts have been made to incorporate importance weighting into the OWA operator [28, 6]. Finally, Torra [22] has introduced the Weighted OWA (WOWA) aggregation as a particular case of Choquet integral

using a distorted probability as the measure. The WOWA averaging is defined by two weighting vectors: the preferential weights $\mathbf{w}$ and the importance weights $\mathbf{p}$. It covers both the weighted means (defined with $\mathbf{p}$) and the OWA averages (defined with $\mathbf{w}$) as special cases. Actually, the WOWA average is reduced to the weighted mean in the case of equal all the preference weights and it becomes the standard OWA average in the case of equal all the importance weights. Since its introduction, the WOWA operator has been successfully applied to many fields of decision making [24, 15, 19, 20] including metadata aggregation problems [1, 10].

In this paper we analyze solution procedures for optimization problems with the ordered average objective functions or constraints. Exactly we consider optimization problems

$$\max \{a(\mathbf{y}) \ : \ \mathbf{y} = \mathbf{f}(\mathbf{x}), \quad \mathbf{x} \in \mathscr{F}\} \tag{2}$$

or

$$\max\{g(\mathbf{y}) \ : \ a(\mathbf{y}) \leq \rho, \quad \mathbf{y} = \mathbf{f}(\mathbf{x}), \quad \mathbf{x} \in \mathscr{F}\} \tag{3}$$

As an aggregation we consider both OWA and WOWA, both with possible some generalization due to possible arbitrary density of grid of the preference weights. We show that the concepts of the LP formulations for the OWA optimization with monotonic preferential weights [18] can easily be extended to cover optimization problems with the WOWA bounds and objectives with arbitrary importance weights and arbitrary density of the preferential weights. A special attention will be paid to problems with linear attribute functions $f_i(\mathbf{x}) = \mathbf{c}_i\mathbf{x}$ and polyhedral feasible sets:

$$\mathbf{y} = \mathbf{f}(\mathbf{x}) = \mathbf{C}\mathbf{x} \quad \text{and} \quad \mathscr{F} = \{\mathbf{x} \in R^q \ : \ \mathbf{A}\mathbf{x} = \mathbf{b}, \quad \mathbf{x} \geqq \mathbf{0}\} \tag{4}$$

where $\mathbf{C}$ is an $m \times q$ matrix (consisting of rows $\mathbf{c}_i$), $\mathbf{A}$ is a given $r \times q$ matrix and $\mathbf{b} = (b_1, \ldots, b_r)^T$ is a given RHS vector.

The paper is organized as follows. In the next section we introduce formally the WOWA operator and its generalization on arbitrary density grids of the preferential weights. We derive some alternative computational formula based on the Lorenz curves and analyze the orness/andness properties of the WOWA operator with monotonic preferential weights. In Section 3 we introduce the LP formulations for bounds or minimization of the WOWA aggregations with decreasing preferential weights. Similarly in Section 4, the LP formulations for bounds or maximization of the WOWA aggregations with increasing weights are given. Finally, in Section 5 we demonstrate computational efficiency of the introduced models.

2 The Ordered Weighted Averages

2.1 OWA and WOWA Aggregations

Let $\mathbf{w} = (w_1, \ldots, w_m)$ be a weighting vector of dimension m such that $w_i \geq 0$ for $i = 1, \ldots, m$ and $\sum_{i=1}^{m} w_i = 1$. The corresponding OWA aggregation of attributes $\mathbf{y} = (y_1, \ldots, y_m)$ can be mathematically formalized as follows [25]. First, we

introduce the ordering map $\Theta : R^m \to R^m$ such that $\Theta(\mathbf{y}) = (\theta_1(\mathbf{y}), \theta_2(\mathbf{y}), \ldots, \theta_m(\mathbf{y}))$, where $\theta_1(\mathbf{y}) \geq \theta_2(\mathbf{y}) \geq \cdots \geq \theta_m(\mathbf{y})$ and there exists a permutation τ of set I such that $\theta_i(\mathbf{y}) = y_{\tau(i)}$ for $i = 1, \ldots, m$. Further, we apply the weighted sum aggregation to ordered achievement vectors $\Theta(\mathbf{y})$, i.e. the OWA aggregation has the following form:

$$A_{\mathbf{w}}(\mathbf{y}) = \sum_{i=1}^{m} w_i \theta_i(\mathbf{y}) \tag{5}$$

The OWA aggregation (5) allows to model various aggregation functions from the maximum ($w_1 = 1$, $w_i = 0$ for $i = 2, \ldots, m$) through the arithmetic mean ($w_i = 1/m$ for $i = 1, \ldots, m$) to the minimum ($w_m = 1$, $w_i = 0$ for $i = 1, \ldots, m-1$). However, the weighted mean (1) cannot be expressed as an OWA aggregation. Actually, the OWA aggregations are symmetric (impartial, neutral) with respect to the individual attributes and it does not allow to represent any importance weights allocated to specific attributes.

Further, let $\mathbf{w} = (w_1, \ldots, w_m)$ and $\mathbf{p} = (p_1, \ldots, p_m)$ be weighting vectors of dimension m such that $w_i \geq 0$ and $p_i \geq 0$ for $i = 1, \ldots, m$ as well as $\sum_{i=1}^{m} w_i = 1$ and $\sum_{i=1}^{m} p_i = 1$. The corresponding Weighted OWA aggregation of outcomes $\mathbf{y} = (y_1, \ldots, y_m)$ is defined as follows [22]:

$$A_{\mathbf{w},\mathbf{p}}(\mathbf{y}) = \sum_{i=1}^{m} \omega_i \theta_i(\mathbf{y}) \tag{6}$$

where the weights ω_i are defined as

$$\omega_i = w^*\Big(\sum_{k \leq i} p_{\tau(k)}\Big) - w^*\Big(\sum_{k < i} p_{\tau(k)}\Big) \tag{7}$$

with w^* a monotone increasing function that interpolates points $(\frac{i}{m}, \sum_{k \leq i} w_k)$ together with the point (0.0) and τ representing the ordering permutation for $\mathbf{y}$ (i.e. $y_{\tau(i)} = \theta_i(\mathbf{y})$). Function w^* is required to be a straight line whenever the points can be interpolated in this way. Due to this requirement, the WOWA aggregation covers the standard weighted mean (1) with weights p_i as a special case of equal preference weights ($w_i = 1/m$ for $i = 1, \ldots, m$). Actually, the WOWA operator is a particular case of Choquet integral using a distorted probability as the measure [4].

Note that function w^* can be expressed as $w^*(\alpha) = \int_0^\alpha g(\xi)\, d\xi$ where g is a generation function. Let us introduce breakpoints $\beta_i = \sum_{k \leq i} p_{\tau(k)}$ and $\beta_0 = 0$. This allows one to express weights ω_i as

$$\omega_i = \int_0^{\beta_i} g(\xi)\, d\xi - \int_0^{\beta_{i-1}} g(\xi)\, d\xi = \int_{\beta_{i-1}}^{\beta_i} g(\xi)\, d\xi$$

and the entire WOWA aggregation as

$$A_{\mathbf{w},\mathbf{p}}(\mathbf{y}) = \sum_{i=1}^{m} \theta_i(\mathbf{y}) \int_{\beta_{i-1}}^{\beta_i} g(\xi)\, d\xi = \int_0^1 g(\xi) F_{\mathbf{y}}^{(-1)}(\xi)\, d\xi \tag{8}$$

where $F_{\mathbf{y}}^{(-1)}$ is the stepwise function $F_{\mathbf{y}}^{(-1)}(\xi) = \theta_i(\mathbf{y})$ for $\beta_{i-1} < \xi \leq \beta_i$. It can also be mathematically formalized as follows. First, we introduce the left-continuous right tail cumulative distribution function (cdf):

$$F_{\mathbf{y}}(d) = \sum_{i \in I} p_i \delta_i(d) \quad \text{where} \quad \delta_i(d) = \begin{cases} 1 & \text{if } y_i \geq d \\ 0 & \text{otherwise} \end{cases} \tag{9}$$

which for any real (outcome) value d provides the measure of outcomes greater or equal to d. Next, we introduce the quantile function $F_{\mathbf{y}}^{(-1)}$ as the right-continuous inverse of the cumulative distribution function $F_{\mathbf{y}}$:

$$F_{\mathbf{y}}^{(-1)}(\xi) = \sup\{\eta : F_{\mathbf{y}}(\eta) \geq \xi\} \quad \text{for } 0 < \xi \leq 1.$$

Formula (8) provides the most general expression of the WOWA aggregation allowing for expansion to continuous case. The original definition of WOWA allows one to build various interpolation functions w^* [23] thus to use different generation functions g in formula (8). We focus our analysis on the simplest case of linear interpolation leading to the piecewise linear function w^*. Note, however, that the piecewise linear functions may be built with various number of breakpoints, not necessarily m. Thus, any nonlinear function can be well approximated by a piecewise linear function with appropriate number of breakpoints. Therefore, we will consider weights vectors $\mathbf{w}$ of dimension n not necessarily equal to m. Any such piecewise linear interpolation function w^* can be expressed with the stepwise generation function

$$g(\xi) = nw_k \quad \text{for } (k-1)/n < \xi \leq k/n, \quad k = 1, \ldots, n \tag{10}$$

This leads us to the following specification of formula (8):

$$A_{\mathbf{w},\mathbf{p}}(\mathbf{y}) = \int_0^1 g(\xi) F_{\mathbf{y}}^{(-1)}(\xi)\, d\xi = \sum_{k=1}^{n} w_k n \int_{(k-1)/n}^{k/n} F_{\mathbf{y}}^{(-1)}(\xi)\, d\xi \tag{11}$$

We will treat formula (11) as a formal definition of the WOWA aggregation of m-dimensional outcomes $\mathbf{y}$ defined by m-dimensional importance weights $\mathbf{p}$ and n-dimensional preferential weights $\mathbf{w}$. Note that quantities $n \int_{(k-1)/n}^{k/n} F_{\mathbf{y}}^{(-1)}(\xi)\, d\xi$ express the conditional means within the corresponding quantiles $(k-1)/n$ and k/n. In the case of $n = m$ and equal importance weights $p_i = 1/n$, formula (11) represents the standard definition of the OWA aggregation (5), since $F_{\mathbf{y}}^{(-1)}(\xi) = \theta_k(\mathbf{y})$ for $(k-1)/n \leq \xi < k/n$. Although formula (11) allows one to express general WOWA aggregations by using the preferential weights to redefine $F_{\mathbf{y}}^{(-1)}(\xi) = \theta_k(\mathbf{y})$ accordingly. Moreover, various values of n possibly different from the number of attributes m allows one to generalize both the WOWA aggregation as well as the OWA aggregation (with equal importance weights $p_i = 1/n$).

When in (8) using the integrals from the left end rather than those on intervals one gets

$$A_{\mathbf{w},\mathbf{p}}(\mathbf{y}) = \sum_{k=1}^{n} nw_k(L(\mathbf{y},\mathbf{p},\frac{k}{n}) - L(\mathbf{y},\mathbf{p},\frac{k-1}{n})) \tag{12}$$

where $L(\mathbf{y},\mathbf{p},\beta)$ is defined by left-tail integrating $F_{\mathbf{y}}^{(-1)}$, i.e.

$$L(\mathbf{y},\mathbf{p},0) = 0 \quad \text{and} \quad L(\mathbf{y},\mathbf{p},\beta) = \int_0^{\beta} F_{\mathbf{y}}^{(-1)}(\alpha)d\alpha \quad \text{for } 0 < \beta \leq 1 \tag{13}$$

In particular, $L(\mathbf{y},\mathbf{p},1) = \int_0^1 F_{\mathbf{y}}^{(-1)}(\alpha)d\alpha = A_{\mathbf{p}}(\mathbf{y})$. Graphs of functions $L(\mathbf{y},\mathbf{p},\beta)$ (with respect to β) take the form of concave curves, the so-called (upper) absolute Lorenz curves. In the case of $n = m$ and equal importance weights $p_i = 1/n$ thus representing the standard OWA aggregation, one gets $L(\mathbf{y},\mathbf{p},\frac{k}{n}) = \frac{1}{n}\sum_{i=1}^{k}\theta_i(\mathbf{y})$ and formula (12) reduces to (5). Although, for $n \neq m$ one gets more complicated formula for $L(\mathbf{y},\mathbf{p},\frac{k}{n})$ even in the case of equal importance weights $p_i = 1/n$ thus representing the generalized OWA aggregation.

Alternatively, one may refer in formula (11) to the integrals from the right end instead of intervals thus getting

$$A_{\mathbf{w},\mathbf{p}}(\mathbf{y}) = \sum_{k=1}^{n} nw_k(\overline{L}(\mathbf{y},\mathbf{p},1-\frac{k-1}{n}) - \overline{L}(\mathbf{y},\mathbf{p},1-\frac{k}{n})) \tag{14}$$

where $\overline{L}(\mathbf{y},\mathbf{p},\beta)$ is defined by right tail integrating $F_{\mathbf{y}}^{(-1)}$, i.e.

$$\overline{L}(\mathbf{y},\mathbf{p},0) = 0 \quad \text{and} \quad \overline{L}(\mathbf{y},\mathbf{p},\beta) = \int_0^{1-\beta} F_{\mathbf{y}}^{(-1)}(1-\alpha)d\alpha \quad \text{for } 0 < \beta \leq 1 \tag{15}$$

One may easily notice that for any $0 \leq \beta \leq 1$

$$L(\mathbf{y},\mathbf{p},\beta) + \overline{L}(\mathbf{y},\mathbf{p},1-\beta) = \int_0^1 F_{\mathbf{y}}^{(-1)}(\alpha)d\alpha = A_{\mathbf{p}}(\mathbf{y})$$

Hence, $\overline{L}(\mathbf{y},\mathbf{p},1) = A_{\mathbf{p}}(\mathbf{y})$. Graphs of functions $\overline{L}(\mathbf{y},\mathbf{p},\beta)$ (with respect to β) take the form of convex curves, the (lower) absolute Lorenz curves. In the case of the standard OWA aggregation represented by $n = m$ and equal importance weights $p_i = 1/n$, one gets $\overline{L}(\mathbf{y},\mathbf{p},1-\frac{k}{n}) = \frac{1}{n}\sum_{i=k}^{n}\theta_i(\mathbf{y})$ thus reducing formula (12) to (5).

2.2 *The Orness and Andness Properties*

The OWA aggregation may model various preferences from the optimistic (max) to the pessimistic (min). Yager [25] introduced a well appealing concept of the orness measure to characterize the OWA operators. The degree of orness associated with the OWA operator $A_{\mathbf{w}}(\mathbf{y})$ is defined as

$$\text{orness}(\mathbf{w}) = \sum_{i=1}^{m} \frac{m-i}{m-1} w_i \tag{16}$$

For the max aggregation representing the fuzzy 'or' operator with weights $\mathbf{w} = (1,0,\ldots,0)$ one gets $\text{orness}(\mathbf{w}) = 1$ while for the min aggregation representing the fuzzy 'and' operator with weights $\mathbf{w} = (0,\ldots,0,1)$ one has $\text{orness}(\mathbf{w}) = 0$. For the average (arithmetic mean) one gets $\text{orness}((1/m,1/m,\ldots,1/m)) = 1/2$. Actually, one may consider a complementary measure of andness defined as $\text{andness}(\mathbf{w}) = 1 - \text{orness}(\mathbf{w})$. OWA aggregations with orness greater or equal 1/2 are considered or-like whereas the aggregations with orness smaller or equal 1/2 are treated as and-like. The former correspond to rather optimistic preferences while the latter represents rather pessimistic preferences.

The OWA aggregations with monotonic weights are either or-like or and-like. Exactly, decreasing weights $w_1 \geq w_2 \geq \ldots \geq w_m$ define an or-like OWA operator, while increasing weights $w_1 \leq w_2 \leq \ldots \leq w_m$ define an and-like OWA operator. Actually, the orness and the andness properties of the OWA operators with monotonic weights are total in the sense that they remain valid for any subaggregations defined by subsequences of their weights. Namely, for any $2 \leq k \leq m$ and any k-dimensional normalized weights subvector $\mathbf{w}^k = \frac{1}{\bar{w}^k}(w_{i_1}, w_{i_2}, \ldots, w_{i_k})$ with $1 \leq i_1 < i_2 < \ldots < i_k \leq m$ and $\bar{w}^k = \sum_{j=1}^{k} w_{i_j}$, one gets $\text{orness}(\mathbf{w}^k) \geq 1/2$ for the OWA operators with decreasing or $\text{orness}(\mathbf{w}^k) \leq 1/2$ for the OWA operators with increasing weights, respectively. Moreover, appropriate weights monotonicity is necessary to achieve the above total orness or andness properties. Therefore, we will refer to the OWA aggregation with decreasing weights as the totally or-like OWA operator, and to the OWA aggregation with increasing weights as the totally and-like OWA operator.

Yager [27] proposed to define the OWA weighting vectors via the regular increasing monotone (RIM) quantifiers, which provide a dimension independent description of the aggregation. A fuzzy subset Q of the real line is called a RIM quantifier if Q is (weakly) increasing with $Q(0) = 0$ and $Q(1) = 1$. The OWA weights can be defined with a RIM quantifier Q as $w_i = Q(i/m) - Q((i-1)/m)$ and the orness measure can be extended to a RIM quantifier (according to $m \to \infty$) as follows [27]

$$\text{orness}(Q) = \int_0^1 Q(\alpha)\, d\alpha \tag{17}$$

Thus, the orness of a RIM quantifier is equal to the area under it. The measure takes the values between 0 (achieved for $Q(1) = 1$ and $Q(\alpha) = 0$ for all other α) and 1 (achieved for $Q(0) = 1$ and $Q(\alpha) = 0$ for all other α). In particular, $\text{orness}(Q) = 1/2$ for $Q(\alpha) = \alpha$ which is generated by equal weights $w_k = 1/n$. Formula (17) allows one to define the orness of the WOWA aggregation (6) which can be viewed with the RIM quantifier $Q(\alpha) = w^*(\alpha)$ [7]. Let us consider piecewise linear function $Q = w^*$ defined by weights vectors $\mathbf{w}$ of dimension n according to the stepwise generation function (10). One may easily notice that decreasing weights $w_1 \geq w_2 \geq \ldots \geq w_n$ generate a strictly increasing concave curve $Q(\alpha) \geq \alpha$ thus guaranteeing the or-likeness of the WOWA operator. Similarly, increasing weights

$w_1 \le w_2 \le \ldots \le w_n$ generate a strictly increasing convex curve $Q(\alpha) \le \alpha$ thus guaranteeing the and-likeness of the WOWA operator. Actually, the monotonic weights generate the totally or-like and and-like operators, respectively, in the sense that that they remain valid for any subaggregations defined with respect to subintervals of the interval $[0,1]$. Namely, for any interval $[a,b]$, where $0 \le a < b \le 1$, and the corresponding part of Q renormalized to represent a RIM quantifier

$$Q_a^b(\alpha) = \frac{Q(a+\alpha(b-a)) - Q(a)}{Q(b) - Q(a)}$$

one gets orness$(Q_a^b) \ge 1/2$ for the OWA operators with decreasing or orness$(Q_a^b) \le 1/2$ for the OWA operators with increasing weights w_i, respectively. Moreover, in the case of piecewise linear function $Q = w^*$ defined by weights vectors $\mathbf{w}$ of dimension n according to the stepwise generation function (10), we consider, appropriate weights monotonicity is necessary to achieve the total orness or andness properties. Therefore, we will refer to the WOWA aggregation with decreasing preferential weights as the totally or-like WOWA operator, and to the WOWA aggregation with increasing preferential weights as the totally and-like WOWA operator.

3 Totally Or-Like Ordered Weighted Aggregations

Consider a totally or-like WOWA aggregation defined by decreasing weights $w_1 \ge w_2 \ge \ldots \ge w_n$. Following formula (12) the WOWA aggregation may be expressed as

$$A_{\mathbf{w},\mathbf{p}}(\mathbf{y}) = \sum_{k=1}^{n} n w_k (L(\mathbf{y},\mathbf{p},\frac{k}{n}) - L(\mathbf{y},\mathbf{p},\frac{k-1}{n})) = \sum_{k=1}^{n} w'_k L(\mathbf{y},\mathbf{p},\frac{k}{n}) \qquad (18)$$

where $w'_n = n w_n$, $w'_k = n(w_k - w_{k+1})$. Due to formula (13), values of function $L(\mathbf{y},\mathbf{p},\alpha)$ for any $0 \le \alpha \le 1$ can be found by optimization:

$$L(\mathbf{y},\mathbf{p},\alpha) = \max_{u_i} \{ \sum_{i=1}^{m} y_i u_i : \sum_{i=1}^{m} u_i = \alpha, \quad 0 \le u_i \le p_i \quad \forall\, i \} \qquad (19)$$

The above problem is an LP for a given outcome vector $\mathbf{y}$ while it becomes nonlinear for $\mathbf{y}$ being a vector of variables. This difficulty can be overcome by taking advantages of the LP dual to (19). Introducing dual variable t corresponding to the equation $\sum_{i=1}^{m} u_i = \alpha$ and variables d_i corresponding to upper bounds on u_i one gets the following LP dual of problem (19):

$$L(\mathbf{y},\mathbf{p},\alpha) = \min_{t,d_i} \{ \alpha t + \sum_{i=1}^{m} p_i d_i : t + d_i \ge y_i,\ d_i \ge 0 \quad \forall\, i \} \qquad (20)$$

Due the LP duality theory, the following assertion is valid.

Lemma 1. *For any value ρ, vector $\mathbf{y}$ fulfills inequality $L(\mathbf{y},\mathbf{p},\xi) \le \rho$ if and only if there exist t and d_i $(i = 1,\dots,m)$ such that*

$$\xi t + \sum_{i=1}^{m} p_i d_i \le \rho \quad \text{and} \quad t + d_i \ge y_i,\ d_i \ge 0 \quad \forall\, i$$

Note that following (18) the WOWA with increasing weights w_k takes the form

$$A_{\mathbf{w},\mathbf{p}}(\mathbf{y}) = \sum_{k=1}^{n} w'_k L(\mathbf{y},\mathbf{p},\frac{k}{n})$$

with positive weights w'_k. Therefore, the following assertions can be proven.

Theorem 1. *Any totally or-like WOWA aggregation $A_{\mathbf{w},\mathbf{p}}$ defined by decreasing weights $w_1 \ge w_2 \ge \dots \ge w_n$ is a piecewise linear convex function of $\mathbf{y}$.*

Proof. Note that for any given $\mathbf{p}$ and ξ, due to formula (20), $L(\mathbf{y},\mathbf{p},\xi)$ is a piecewise linear convex function of $\mathbf{y}$. Hence, due to decreasing preferential weights, following formula (37) the entire WOWA aggregation is a piecewise linear convex function of $\mathbf{y}$ as a linear combination of functions $L(\mathbf{y},\mathbf{p},\xi)$ for $\xi = k/n$, $k = 1,2,\dots,n$ with nonnegative weights w'_k.

Theorem 2. *For any totally or-like WOWA aggregation $A_{\mathbf{w},\mathbf{p}}$ defined by decreasing weights $w_1 \ge w_2 \ge \dots \ge w_n$ and any constant ρ inequality $A_{\mathbf{w},\mathbf{p}}(\mathbf{y}) \le \rho$ is valid if and only if there exist t_k and d_{ik} $(i = 1,\dots,m; k = 1,2,\dots,n)$ such that*

$$\begin{aligned} &\sum_{k=1}^{n} w'_k [\frac{k}{n} t_k + \sum_{i=1}^{m} p_i d_{ik}] \le \rho \\ &t_k + d_{ik} \ge y_i,\ d_{ik} \ge 0 \qquad \text{for } i = 1,\dots,m;\ k = 1,\dots,n \end{aligned} \tag{21}$$

Proof. Assume that there exist t_k^0 and d_{ik}^0 $(i = 1,\dots,m; k = 1,2,\dots,n)$ satisfying the requirements (21). Then, according to Lemma 1,

$$L(\mathbf{y},\mathbf{p},\frac{k}{n}) \le \frac{k}{n} t_k^0 + \sum_{i=1}^{m} p_i d_{ik}^0 \quad \forall\, k$$

Hence, due to nonnegative weights w'_k,

$$A_{\mathbf{w},\mathbf{p}}(\mathbf{y}) = \sum_{k=1}^{n} w'_k L(\mathbf{y},\mathbf{p},\frac{k}{n}) \le \sum_{k=1}^{n} w'_k [\frac{k}{n} t_k^0 + \sum_{i=1}^{m} p_i d_{ik}^0] \le \rho$$

which proves the required inequality.

Assume now that the inequality $A_{\mathbf{w},\mathbf{p}}(\mathbf{y}) \le \rho$ holds. Define t_k^0 and d_{ik}^0 $(i = 1,\dots,m; k = 1,2,\dots,n)$ as optimal solutions to problems (20) for $\xi = k/n$ $(k = 1,\dots,n)$, respectively. They obviously fulfill conditions (21).

Consider an optimization problem with an upper bound on a totally or-like WOWA aggregation

$$\max\{g(\mathbf{y}) \;:\; A_{\mathbf{w},\mathbf{p}}(\mathbf{y}) \le \rho, \quad \mathbf{y} = \mathbf{f}(\mathbf{x}), \quad \mathbf{x} \in \mathscr{F}\} \tag{22}$$

Following Theorem 2 it can be reformulated as

$$\begin{aligned} \max_{t_k, d_{ik}, y_i, x_j} \quad & g(\mathbf{y}) \\ \text{s.t.} \quad & \sum_{k=1}^{n} w'_k [\frac{k}{n} t_k + \sum_{i=1}^{m} p_i d_{ik}] \le \rho \\ & t_k + d_{ik} \ge y_i, \; d_{ik} \ge 0 \quad \text{for } \; i = 1, \ldots, m; \; k = 1, \ldots, n \\ & \mathbf{y} \le \mathbf{f}(\mathbf{x}), \quad \mathbf{x} \in \mathscr{F} \end{aligned}$$

In the case of model (4) with linear function $g(\mathbf{y}) = \sum_{i=1}^{m} g_i y_i$ this leads us to the following LP formulation of the optimization problem (22):

$$\max \sum_{i=1}^{m} g_i y_i \tag{23}$$

$$\text{s.t.} \quad \mathbf{A}\mathbf{x} = \mathbf{b} \tag{24}$$

$$\mathbf{y} - \mathbf{C}\mathbf{x} = \mathbf{0} \tag{25}$$

$$\sum_{k=1}^{n} \frac{k}{n} w'_k t_k + \sum_{k=1}^{n} \sum_{i=1}^{m} w'_k p_i d_{ik} \le \rho \tag{26}$$

$$d_{ik} \ge y_i - t_k \quad \text{for } i = 1, \ldots, m; \; k = 1, \ldots, n \tag{27}$$

$$d_{ik} \ge 0 \quad \text{for } i = 1, \ldots, m; \; k = 1, \ldots, n; \; x_j \ge 0 \; \forall \; j \tag{28}$$

Model (23)–(28) is an LP problem with $mn + m + n + q$ variables and $mn + m + r + 1$ constraints. In the case of multiple WOWA constraints one gets additional $mn + m$ variables and $mn + 1$ inequalities per each constraint. Thus, for problems with not too large number of attributes (m) and preferential weights (n) it can be solved directly.

Consider now minimization of a totally or-like WOWA aggregation

$$\min\{A_{\mathbf{w},\mathbf{p}}(\mathbf{y}) \;:\; \mathbf{y} = \mathbf{f}(\mathbf{x}), \quad \mathbf{x} \in \mathscr{F}\} \tag{29}$$

Taking advantages of Theorem 2, minimization of the WOWA criterion may be expressed as the following problem with auxiliary linear inequalities:

$$\begin{aligned} \min_{\zeta, t_k, d_{ik}, y_i, x_j} \quad & \zeta \\ \text{s.t.} \quad & \sum_{k=1}^{n} w'_k [\frac{k}{n} t_k + \sum_{i=1}^{m} p_i d_{ik}] \le \zeta \\ & t_k + d_{ik} \ge y_i, \; d_{ik} \ge 0 \quad \text{for } \; i = 1, \ldots, m; \; k = 1, \ldots, n \\ & \mathbf{y} \le \mathbf{f}(\mathbf{x}), \quad \mathbf{x} \in \mathscr{F} \end{aligned}$$

While eliminating the ζ variable this leads us to the following LP formulation of the WOWA problem:

$$\begin{aligned} \min_{t_k,d_{ik},y_i,x_j} & \sum_{k=1}^{n} w'_k[\frac{k}{n}t_k + \sum_{i=1}^{m} p_i d_{ik}] \\ \text{s.t. } & t_k + d_{ik} \geq y_i,\ d_{ik} \geq 0 \quad \text{for } i = 1,\ldots,m;\ k = 1,\ldots,n \\ & \mathbf{y} \leq \mathbf{f}(\mathbf{x}), \quad \mathbf{x} \in \mathscr{F} \end{aligned}$$

When taking into account the linear attributes and constraints (4) we get the following LP formulation of the WOWA optimization problem (29):

$$\min \sum_{k=1}^{n} \frac{k}{n} w'_k t_k + \sum_{k=1}^{n}\sum_{i=1}^{m} w'_k p_i d_{ik} \tag{30}$$

$$\text{s.t. } \mathbf{Ax} = \mathbf{b} \tag{31}$$

$$\mathbf{y} - \mathbf{Cx} = \mathbf{0} \tag{32}$$

$$d_{ik} \geq y_i - t_k \quad \text{for } i = 1,\ldots,m;\ k = 1,\ldots,n \tag{33}$$

$$d_{ik} \geq 0 \quad \text{for } i = 1,\ldots,m;\ k = 1,\ldots,n;\ x_j \geq 0\ \forall\ j \tag{34}$$

This LP problem contains $mn + m + n + q$ variables and $mn + m + r$ constraints. Thus, for not too large values of m and n it can be solved directly. Actually, the LP model is quite similar to that introduced in [18] for the OWA optimization (c.f., model (30)–(34)).

The number of constraints in problem (30)–(34) is similar to the number of variables. However, the crucial number of variables (mn variables d_{ik}) is associated with singleton columns. Therefore, it may be better to deal with the dual of (30)–(34) where the corresponding rows become simple upper bounds, thus reducing dramatically the LP problem size. While introducing the dual variables: $\mathbf{u} = (u_1,\ldots,u_r)$, $\mathbf{v} = (v_1,\ldots,v_m)$ and $\mathbf{z} = (z_{ik})_{i=1,\ldots,m;\ k=1,\ldots,n}$ corresponding to the constraints (31), (32) and (33), respectively, we get the following dual:

$$\begin{aligned} & \max\ \mathbf{ub} \\ \text{s.t. } & \mathbf{uA} - \mathbf{vC} \overset{\leq}{=} \mathbf{0} \\ & v_i - \sum_{k=1}^{n} z_{ik} = 0 \quad \text{for } i = 1,\ldots,m \\ & \sum_{i=1}^{m} z_{ik} = \frac{k}{n} w'_k \quad \text{for } k = 1,\ldots,n \\ & 0 \leq z_{ik} \leq p_i w'_k \quad \text{for } i = 1,\ldots,m;\ k = 1,\ldots,n \end{aligned} \tag{35}$$

The dual problem (35) is consisted of only $m + n + q$ structural constraints on $mn + r + m$ variables. Since the average complexity of the simplex method depends on the number of constraints, the dual model (35) can be directly solved for quite large values of m and n. Moreover, the columns corresponding to mn variables z_{ik} form the network (node-link incidence) matrix thus allowing one to employ special techniques of the network embedded simplex algorithm [3].

Similar to the case of minimization of the totally or-like WOWA, it may be also introduced the dual of (23)–(28) representing the WOWA constraints. Indeed, while introducing the dual variables: $\mathbf{u} = (u_1,\ldots,u_r)$, $\mathbf{v} = (v_1,\ldots,v_m)$, ξ and

$\mathbf{z} = (z_{ik})_{i=1,\ldots,m;\ k=1,\ldots,n}$ corresponding to the constraints (24), (25), (26) and (27), respectively, we get the following dual:

$$\begin{aligned} &\min\ \mathbf{ub} + \rho\xi \\ \text{s.t. } &\mathbf{uA} - \mathbf{vC} \stackrel{\geq}{=} \mathbf{0} \\ &v_i + \sum_{k=1}^{n} z_{ik} = g_i \qquad \text{for } i = 1,\ldots,m \\ &\sum_{i=1}^{m} z_{ik} - \frac{k}{n} w'_k \xi = 0 \qquad \text{for } k = 1,\ldots,n \\ &z_{ik} \leq w'_k p_i \xi \qquad \text{for } i = 1,\ldots,m;\ k = 1,\ldots,n \\ &\xi \geq 0, \quad z_{ik} \geq 0 \qquad \text{for } i = 1,\ldots,m;\ k = 1,\ldots,n \end{aligned} \tag{36}$$

However, the mn rows corresponding to variables d_{ik} represent variable upper bounds ([21, 11]) instead of simple upper bounds. Thus the model simplification is not so dramatic.

4 Totally And-Like Ordered Weighted Aggregations

Consider a totally and-like WOWA aggregation defined by increasing weights $w_1 \leq w_2 \leq \ldots \leq w_n$. By consideration of $-\mathbf{y}$ instead of $\mathbf{y}$ such an aggregation may be viewed as a negative to the totally or-like WOWA aggregation defined by decreasing weights

$$A_{\mathbf{w},\mathbf{p}}(\mathbf{y}) = -A_{\bar{\mathbf{w}},\mathbf{p}}(\mathbf{y}) \quad \text{where} \quad \bar{w}_k = w_{n-k+1} \quad \text{for } k = 1,\ldots,n$$

Alternatively, taking advantages of formula (14) the WOWA aggregation may be expressed as

$$A_{\mathbf{w},\mathbf{p}}(\mathbf{y}) = \sum_{k=1}^{n} n w_k \left(\overline{L}(\mathbf{y},\mathbf{p},1-\frac{k-1}{n}) - \overline{L}(\mathbf{y},\mathbf{p},1-\frac{k}{n})\right) = \sum_{k=1}^{n} w''_k \overline{L}(\mathbf{y},\mathbf{p},\frac{k}{n}) \tag{37}$$

with weights $w''_k = -w'_{n-k} = n(w_{n-k+1} - w_{n-k})$ for $k = 1,\ldots,n-1$ and $w''_n = n w_1$ while values of function $\overline{L}(\mathbf{y},\mathbf{p},\xi)$ for any $0 \leq \xi \leq 1$ are given by optimization:

$$\overline{L}(\mathbf{y},\mathbf{p},\xi) = \min_{u_i} \{ \sum_{i=1}^{m} y_i u_i : \sum_{i=1}^{m} u_j = \xi, \quad 0 \leq u_i \leq p_i \quad \forall\, i \} \tag{38}$$

Introducing dual variable t corresponding to the equation $\sum_{i=1}^{m} u_i = \xi$ and variables d_i corresponding to upper bounds on u_i one gets the following LP dual expression of $\overline{L}(\mathbf{y},\mathbf{p},\xi)$

$$\overline{L}(\mathbf{y},\mathbf{p},\xi) = \max_{t,d_i} \{\xi t - \sum_{i=1}^{m} p_i d_i : t - d_i \leq y_i,\ d_i \geq 0 \quad \forall\, i\} \tag{39}$$

Due the duality theory, for any given vector $\mathbf{y}$ the cumulated ordered coefficient $\overline{L}(\mathbf{y},\mathbf{p},\xi)$ can be found as the optimal value of the above LP problem. Actually, relation (39) can be expressed as the following assertion.

Lemma 2. *For any value ρ, vector $\mathbf{y}$ fulfills inequality $\overline{L}(\mathbf{y},\mathbf{p},\xi) \geq \rho$ if and only if there exist t and d_i $(i = 1,\ldots,m)$ such that*

$$\xi t - \sum_{i=1}^{m} p_i d_i \geq \rho \quad \text{and} \quad t - d_i \leq y_i,\ d_i \geq 0 \quad \forall\, i$$

Note that following (37) the WOWA with increasing weights w_k takes the form

$$A_{\mathbf{w},\mathbf{p}}(\mathbf{y}) = \sum_{k=1}^{n} w''_k \overline{L}(\mathbf{y},\mathbf{p},\frac{k}{n})$$

with positive weights w''_k. This enables the following statements.

Theorem 3. *Any totally and-like WOWA aggregation $A_{\mathbf{w},\mathbf{p}}(\mathbf{y})$ defined by increasing preferential weights $w_1 \leq w_2 \leq \ldots \leq w_n$ is a piecewise linear concave function of $\mathbf{y}$.*

Proof. Note that for any given $\mathbf{p}$ and ξ, due to formula (39), $\overline{L}(\mathbf{y},\mathbf{p},\xi)$ is a piecewise linear concave function of $\mathbf{y}$. Hence, due to increasing preferential weights, following formula (37) the entire WOWA aggregation is a piecewise linear concave function of $\mathbf{y}$ as a linear combination of functions $\overline{L}(\mathbf{y},\mathbf{p},\xi)$ for $\xi = k/n$, $k = 1,2,\ldots,n$ with nonnegative weights w'_k.

Theorem 4. *For any totally and-like WOWA aggregation $A_{\mathbf{w},\mathbf{p}}$ defined by increasing weights $w_1 \leq w_2 \leq \ldots \leq w_n$ and any constant ρ inequality $A_{\mathbf{w},\mathbf{p}}(\mathbf{y}) \geq \rho$ is valid if and only if there exist t_k and d_{ik} $(i = 1,\ldots,m; k = 1,2,\ldots,n)$ such that*

$$\begin{aligned} &\sum_{k=1}^{n} w''_k [\frac{k}{n} t_k - \sum_{i=1}^{m} p_i d_{ik}] \geq \rho \\ &t_k - d_{ik} \leq y_i,\ d_{ik} \geq 0 \qquad \text{for } i = 1,\ldots,m;\ k = 1,\ldots,n \end{aligned} \tag{40}$$

Proof. Assume that there exist t^0_k and d^0_{ik} $(i = 1,\ldots,m; k = 1,2,\ldots,n)$ satisfying the requirements (40). Then, according to Lemma 2,

$$\overline{L}(\mathbf{y},\mathbf{p},\frac{k}{n}) \geq \frac{k}{n} t^0_k - \sum_{i=1}^{m} p_i d^0_{ik} \quad \forall\, k$$

Hence, due to nonnegative weights w''_k,

$$A_{\mathbf{w},\mathbf{p}}(\mathbf{y}) = \sum_{k=1}^{n} w''_k \overline{L}(\mathbf{y},\mathbf{p},\frac{k}{n}) \geq \sum_{k=1}^{n} w''_k [\frac{k}{n} t^0_k - \sum_{i=1}^{m} p_i d^0_{ik}] \geq \rho$$

which proves the required inequality.

Assume now that the inequality $A_{\mathbf{w},\mathbf{p}}(\mathbf{y}) \geq \rho$ holds. Define t_k^0 and d_{ik}^0 ($i = 1,\ldots,m; k = 1,2,\ldots,n$) as optimal solutions to problems (39) for $\xi = k/n$ ($k = 1,\ldots,n$), respectively. They obviously fulfill conditions (40).

Consider an optimization problem with a lower bound on a totally and-like WOWA aggregation

$$\max\{g(\mathbf{y}) \,:\, A_{\mathbf{w},\mathbf{p}}(\mathbf{y}) \geq \rho, \quad \mathbf{y} = \mathbf{f}(\mathbf{x}), \quad \mathbf{x} \in \mathscr{F}\} \tag{41}$$

Following Theorem 4 it can be reformulated as

$$\begin{aligned} \max_{t_k, d_{ik}, y_i, x_j} \quad & g(\mathbf{y}) \\ \text{s.t.} \quad & \sum_{k=1}^{n} w_k''\Big[\frac{k}{n} t_k - \sum_{i=1}^{m} p_i d_{ik}\Big] \geq \rho \\ & t_k - d_{ik} \leq y_i,\ d_{ik} \geq 0 \quad \text{for } i = 1,\ldots,m;\ k = 1,\ldots,n \\ & \mathbf{y} \leq \mathbf{f}(\mathbf{x}), \quad \mathbf{x} \in \mathscr{F} \end{aligned}$$

For model (4) with linear function $g(\mathbf{y}) = \sum_{i=1}^{m} g_i y_i$ this leads us to the following LP formulation of the optimization problem (41):

$$\max \sum_{i=1}^{m} g_i y_i \tag{42}$$

$$\text{s.t.} \quad \mathbf{A}\mathbf{x} = \mathbf{b} \tag{43}$$

$$\mathbf{y} - \mathbf{C}\mathbf{x} = \mathbf{0} \tag{44}$$

$$\sum_{k=1}^{n} \frac{k}{n} w_k'' t_k - \sum_{k=1}^{n} \sum_{i=1}^{m} w_k'' p_i d_{ik} \geq \rho \tag{45}$$

$$d_{ik} \geq t_k - y_i \quad \text{for } i = 1,\ldots,m;\ k = 1,\ldots,n \tag{46}$$

$$d_{ik} \geq 0 \quad \text{for } i = 1,\ldots,m;\ k = 1,\ldots,n;\ x_j \geq 0\ \forall\ j \tag{47}$$

Model (42)–(47) is an LP problem with $mn + m + n + q$ variables and $mn + m + r + 1$ constraints. In the case of multiple WOWA constraints one gets additional $mn + m$ variables and $mn + 1$ inequalities per each constraint. Thus, for problems with not too large number of attributes (m) and preferential weights (n) it can be solved directly.

Consider now maximization of a totally and-like WOWA aggregation

$$\max\{A_{\mathbf{w},\mathbf{p}}(\mathbf{y}) \,:\, \mathbf{y} = \mathbf{f}(\mathbf{x}), \quad \mathbf{x} \in \mathscr{F}\} \tag{48}$$

Maximization of the WOWA aggregation (48) can be expressed as follows

$$\begin{aligned}\max_{\zeta,t_k,d_{ik},y_i,x_j} \quad & \zeta \\ \text{s.t. } \zeta &\le \sum_{k=1}^{n} w''_k[\frac{k}{n}t_k - \sum_{i=1}^{m} p_i d_{ik}] \\ & t_k - d_{ik} \le y_i,\ d_{ik} \ge 0 \quad \text{for } i=1,\ldots,m;\ k=1,\ldots,n \\ & \mathbf{y} \le \mathbf{f}(\mathbf{x}), \quad \mathbf{x} \in \mathscr{F}\end{aligned}$$

While eliminating the ζ variable this leads us to the following LP formulation of the WOWA problem:

$$\begin{aligned}\max_{t_k,d_{ik},y_i,x_j} \quad & \sum_{k=1}^{n} w''_k[\frac{k}{n}t_k - \sum_{i=1}^{m} p_i d_{ik}] \\ \text{s.t. } & t_k - d_{ik} \le y_i,\ d_{ik} \ge 0 \quad \text{for } i=1,\ldots,m;\ k=1,\ldots,n \\ & \mathbf{y} \le \mathbf{f}(\mathbf{x}), \quad \mathbf{x} \in \mathscr{F}\end{aligned}$$

In the case of model (4) this leads us to the following LP formulation of the WOWA maximization problem (48):

$$\max \sum_{k=1}^{n} \frac{k}{n} w''_k t_k - \sum_{k=1}^{n}\sum_{i=1}^{m} w''_k p_i d_{ik} \tag{49}$$

$$\text{s.t. } \mathbf{Ax} = \mathbf{b} \tag{50}$$

$$\mathbf{y} - \mathbf{Cx} = \mathbf{0} \tag{51}$$

$$d_{ik} \ge t_k - y_i \quad \text{for } i=1,\ldots,m;\ k=1,\ldots,n \tag{52}$$

$$d_{ik} \ge 0 \quad \text{for } i=1,\ldots,m;\ k=1,\ldots,n;\ x_j \ge 0\ \forall\ j \tag{53}$$

The problem has the identical structure as that of (30)–(34) differing only with some negative signs in the objective function (49) and the deviation variable definition (52). While in (30)–(34) variables d_{ik} represent the upperside deviations from the corresponding targets t_k, here they represent the downside deviations for those targets. Note that WOWA model (49)–(53) differs from the analogous deviational model for the OWA optimization [18] only due to coefficients within the objective function (49) and the possibility of different values of m and n. In other words, the OWA deviational model [18] can easily be expanded to accommodate the importance weighting of WOWA.

Model (49)–(53) is an LP problem with $mn+m+n+q$ variables and $mn+m+r$ constraints. Thus, for problems with not too large number of criteria (m) and preferential weights (n) it can be solved directly. However, similar to the case of minimization of the or-like WOWA, it may be better to deal with the dual of (49)–(53) where mn rows corresponding to variables d_{ik} represent only simple upper bounds. Indeed, while introducing the dual variables: $\mathbf{u} = (u_1,\ldots,u_r)$, $\mathbf{v} = (v_1,\ldots,v_m)$ and $\mathbf{z} = (z_{ik})_{i=1,\ldots,m;\ k=1,\ldots,n}$ corresponding to the constraints (50), (51) and (52), respectively, we get the following dual:

$$\begin{aligned} &\min \mathbf{ub} \\ \text{s.t. } &\mathbf{uA} - \mathbf{vC} \stackrel{\geq}{=} \mathbf{0} \\ &v_i - \sum_{k=1}^{n} z_{ik} = 0 \quad \text{for } i = 1,\dots,m \\ &\sum_{i=1}^{m} z_{ik} = \frac{k}{n} w''_k \quad \text{for } k = 1,\dots,n \\ &0 \leq z_{ik} \leq w''_k p_i \quad \text{for } i = 1,\dots,m;\ k = 1,\dots,n \end{aligned} \tag{54}$$

The dual problem (54), similar to (35), contains $mn + r + m$ variables and $m + n + q$ structural constraints. Therefore, it can be directly solved for quite large values of m and n.

Similar to the case of maximization of the totally and-like WOWA, it may be also introduced the dual of (42)–(47) representing the WOWA constraints. Indeed, while introducing the dual variables: $\mathbf{u} = (u_1,\dots,u_r)$, $\mathbf{v} = (v_1,\dots,v_m)$, ξ and $\mathbf{z} = (z_{ik})_{i=1,\dots,m;\ k=1,\dots,n}$ corresponding to the constraints (43), (44), (45) and (46), respectively, we get the following dual:

$$\begin{aligned} &\min \mathbf{ub} - \rho\xi \\ \text{s.t. } &\mathbf{uA} - \mathbf{vC} \stackrel{\geq}{=} \mathbf{0} \\ &v_i - \sum_{k=1}^{n} z_{ik} = g_i \quad \text{for } i = 1,\dots,m \\ &\sum_{i=1}^{m} z_{ik} - \frac{k}{n} w''_k \xi = 0 \quad \text{for } k = 1,\dots,n \\ &z_{ik} \leq w''_k p_i \xi \quad \text{for } i = 1,\dots,m;\ k = 1,\dots,n \\ &\xi \geq 0, \quad z_{ik} \geq 0 \quad \text{for } i = 1,\dots,m;\ k = 1,\dots,n \end{aligned} \tag{55}$$

However, the mn rows corresponding to variables d_{ik} represent variable upper bounds ([21, 11]) instead of simple upper bounds. Thus the model simplification is not so dramatic.

5 Computational Tests

In order to examine computational performances of the LP models for the WOWA optimization we have solved randomly generated problems with varying number q of decision variables and number m of attributes. The core LP feasible set has been defined by a single knapsack-type constraint

$$A\{\mathbf{y} = \mathbf{f}(\mathbf{x}) \ : \ \sum_{j=1}^{q} x_j = 1, \quad x_j \geq 0 \quad \text{for } j = 1,\dots,q\} \tag{56}$$

where $f_i(\mathbf{x}) = \mathbf{c}_i\mathbf{x} = \sum_{j=1}^{q} c_{ij}x_j$. Such problems may be interpreted as resource allocation decisions [17] as well as relocation ones [5] or portfolio selection

Table 1 WOWA criterion optimization times [s]: primal model (49)–(53)

Number of attributes (m)	Number of variables (q)							
	10	20	50	100	150	200	300	400
10	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
20	0.0	0.0	0.0	0.0	0.0	0.2	0.0	0.0
50	0.8	1.0	1.4	1.6	1.4	1.4	1.4	1.6
100	22.6	27.2	35.6	37.8	48.8	71.6	71.2	111.6
150	196.8	259.2	359.8	355.8	387.6	484.4	446.0	^{3}558.4

problem [12] when several attributes represent the unique scenario realizations under various scenarios. Assuming the attributes represent some desired quantities we have considered totally and-like WOWA aggregation defined by increasing weights $w_1 \leq w_2 \leq \ldots \leq w_n$. We have analyzed both the WOWA maximization problem

$$\max \{A_{\mathbf{w},\mathbf{p}}(\mathbf{y}) \ : \ \mathbf{y} \in A\} \tag{57}$$

as well as the WOWA lower bounded problem of the weighted mean maximization

$$\max \{A_{\mathbf{p}}(\mathbf{y}) \ : \ A_{\mathbf{w},\mathbf{p}}(\mathbf{y}) \geq \rho, \ \mathbf{y} \in A\} \tag{58}$$

The former correspond to the multiple conditional value-at-risk performance measure maximization [9] while the latter represents more traditional approach to the portfolio optimization where the expected return is maximized with some risk measure bound [8].

For our computational tests we have randomly generated problems (57) and (58). Coefficients c_{ij} were generated as follows. First, for each j the upper bound r_j was generated as a random number uniformly distributed in the interval $[0.05, 0.15]$. Next, individual coefficients c_{ij} were generated as uniformly distributed in the interval $[-0.75r_j, r_j]$. In order to generate strictly increasing and positive preference weights w_k, we generated randomly the corresponding increments $\delta_k = w_k - w_{k-1}$. The latter were generated as uniformly distributed random values in the range of 1.0 to 2.0, except from a few (5 on average) possibly larger increments ranged from 1.0 to $n/3$. Importance weights p_i were generated according to the exponential smoothing scheme, $p_i = \alpha(1-\alpha)^{i-1}$ for $i = 1, 2, \ldots, m$ and the parameter α is chosen for each test problem size separately to keep the smallest weight p_m around 0.001. The ρ value in (58) was always set to 90% of the objective of the corresponding problem (57).

For each number of decision variables q and number of attributes m we solved 5 randomly generated problems (57) and (58). All computations were performed on a PC with the Pentium 4 2.4GHz processor employing the CPLEX 9.1 package with standard settings and with the time limit of 600 seconds.

In Tables 1 and 2 we show the solution times for the primal (49)–(53) and the dual (54) forms of the computational model, being the averages of 5 randomly generated problems. Upper index in front of the time value indicates the number of tests

Table 2 WOWA criterion optimization times [s]: dual model (54)

Number of attributes (m)	Number of variables (q)							
	10	20	50	100	150	200	300	400
10	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
20	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
50	0.0	0.2	0.2	0.4	0.2	0.4	0.4	0.4
100	0.4	0.6	1.0	7.0	8.2	9.8	10.6	15.4
150	1.4	2.2	3.2	25.2	50.6	53.4	62.0	71.0
200	3.8	5.2	8.8	65.4	156.4	217.8	291.6	253.2
300	10.2	18.0	30.6	132.8	[2]486.6	–	–	–
400	29.6	38.8	88.2	–	–	–	–	–

Table 3 WOWA bound optimization times [s]: primal model (42)–(47)

Number of attributes (m)	Number of variables (q)							
	10	20	50	100	150	200	300	400
10	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
20	0.0	0.0	0.0	0.0	0.2	0.0	0.2	0.0
50	1.0	1.6	1.8	2.4	2.4	2.4	2.8	3.0
100	30.0	52.2	111.4	97.6	88.0	81.4	76.4	82.2
150	189.6	309.8	[1]539.6	[1]513.8	[1]499.6	[2]528.0	[1]532.6	[4]234.6

among 5 that exceeded the time limit. The empty cell (minus sign) shows that this has occurred for all 5 instances. As one can see, the dual form of the model performs much better in each tested problem size. It behaves very well with increasing number of variables if the number of attributes does not exceed 150, and satisfactory if the number of attributes equals 200. Similarly, the model performs very well with increasing number of attributes if only the number of variables does not exceed 50.

Tables 3 and 4 contain solution times for the primal (42)–(47) and the dual (55) form of the model of the weighted mean maximization with the WOWA lower boundary.

As one can see the primal approach requires similar computation effort for the WOWA problem as well as for the weighted mean with the WOWA lower bound. The dual approach, however, is much better suited for the standard WOWA problem then for the weighted mean with the WOWA lower bound. The reason for that is the change of singleton columns to the doubleton ones resulting from the introduction of the WOWA constraint. But still, for the weighted mean with the WOWA lower bound the dual model is a better choice.

In order to examine how much importance weighting of the WOWA complicates our optimization models we have rerun all the tests assuming equal importance weights thus restricting the models to the standard OWA optimization according to [18]. Tables 5 to 8 show the solution times for the primal (49)–(53) and the dual (54) OWA models as well as for the primal (42)–(47) and the dual (55) OWA bounded

Table 4 WOWA bound optimization times [s]: dual model (55)

Number of attributes (m)	Number of variables (q)							
	10	20	50	100	150	200	300	400
10	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
20	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
50	1.2	1.0	0.4	0.6	0.8	0.8	0.8	1.0
100	20.4	42.2	41.2	26.4	20.8	22.0	22.4	23.8
150	138.8	186.4	319.6	149.6	171.8	175.4	178.6	[1]221.2
200	422.2	[3]529.4	–	[4]578.2	[3]554.6	[3]589.4	–	–

Table 5 OWA criterion optimization times [s]: primal model (49)–(53)

Number of attributes (m)	Number of variables (q)							
	10	20	50	100	150	200	300	400
10	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
20	0.0	0.0	0.0	0.0	0.0	0.0	0.2	0.0
50	0.6	0.8	1.0	1.2	1.2	1.2	1.2	1.4
100	18.6	26.4	37.2	40.6	40.2	50.6	49.4	72.4
150	170.8	246.4	355.8	305.0	330.2	365.2	380.0	387.8
200	[2]537.4	–	–	–	–	–	–	–

Table 6 OWA criterion optimization times [s]: dual model

Number of attributes (m)	Number of variables (q)							
	10	20	50	100	150	200	300	400
10	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
20	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
50	0.2	0.0	0.4	0.2	0.4	0.4	0.6	0.8
100	0.4	0.8	2.4	7.2	7.8	12.2	13.4	18.4
150	1.4	2.2	5.8	26.8	44.6	48.8	73.6	102.2
200	3.4	4.8	9.8	62.6	107.2	179.2	246.0	197.6
300	8.6	15.8	29.2	223.0	[2]503.0	[4]592.4	–	–
400	21.6	35.6	67.4	[2]315.0	–	–	–	–

weighted mean optimization models, respectively, with equal importance weights while all the other parameters remain the same.

One may notice that in the case of the primal model the WOWA optimization times (Table 1 and Table 3) are 10–30% longer than the corresponding OWA optimization times (Table 5 and Table 7). On the other hand, in the case of the dual model the WOWA optimization times (Table 2 and Table 4) turn out to be shorter than the corresponding OWA times (Table 6 and Table 8), and frequently even shorter.

The optimization times were analyzed for various size parameters m and q. The basic tests were performed for the standard WOWA model with $n = m$. However, we

Table 7 OWA bound optimization times [s]: primal model (42)–(47)

Number of attributes (m)	Number of variables (q)							
	10	20	50	100	150	200	300	400
10	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
20	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
50	0.8	1.2	1.6	2.0	1.8	1.8	2.0	2.2
100	29.4	51.0	85.6	70.2	73.4	58.4	64.6	59.2
150	162.4	303.0	[1]508.0	440.4	393.0	365.2	380.0	414.0
200	[2]321.0	–	–	–	–	–	–	–

Table 8 OWA bound optimization times [s]: dual model (55)

Number of attributes (m)	Number of variables (q)							
	10	20	50	100	150	200	300	400
10	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
20	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
50	1.2	1.2	0.6	1.0	1.0	1.0	1.4	1.4
100	23.2	40.0	50.2	27.6	25.0	26.0	48.4	64.2
150	131.4	244.6	411.0	236.6	[1]213.6	132.0	145.2	170.4
200	437.2	–	–	–	–	[3]535.6	–	–

Table 9 WOWA optimization times [s]: varying number of preferential weights ($m = 100$, $q = 50$)

Model	Number of preferential weights (n)									
	3	5	10	20	50	100	150	200	300	400
WOWA criterion	0.0	0.0	0.0	0.2	1.0	1.0	1.4	2.2	3.2	4.8
WOWA bound	0.2	0.0	0.2	0.8	7.8	42.2	259.8	185.6	[2]500.2	–

also analyzed the case of larger n for more detailed preferences modeling, as well as the case of smaller n thus representing a rough preferences model.

Table 9 presents solution times for the dual model with different numbers of the preferential weights for problems with 100 attributes and 50 variables. One may notice that the computational efficiency can be improved by reducing the number of preferential weights which can be reasonable in non-automated decision making support systems. On the other hand, in case of the WOWA optimization (but not WOWA bounded weighted mean optimization) increasing the number of preferential weights and thus the number of breakpoints in the interpolation function does not induce the massive increase in the computational complexity.

6 Concluding Remarks

The problem of aggregating multiple attributes to form overall functions is of considerable importance in many disciplines. The WOWA aggregation [22] represents a universal tool allowing to take into account both the preferential weights allocated to ordered attribute values and the importance weights allocated to several attributes. The ordered aggregation operators are generally hard to implement when applied to variables. We have shown that the WOWA aggregations with the monotonic weights can be modeled by introducing auxiliary linear constraints. Exactly, the OWA LP-solvable models introduced in [18] can be expanded to accommodate the importance weighting of the WOWA aggregation used within the inequalities or objective functions.

Our computational experiments have shown that the formulations enable to solve effectively medium size problems. While taking advantages of the dual model the WOWA problems with up to 100 attributes have been solved directly by general purpose LP code within less than half a minute. Although the problems with the WOWA constraints have required typically more time than similar problems with the WOWA objective function. Further research on efficient algorithm for LP problems with the WOWA bounds seems to be promising.

References

1. Damiani, E., De Capitani di Vimercati, S., Samarati, P., Viviani, M.: A WOWA-based aggregation technique on trust values connected to metadata. Electronic Notes Theor. Comp. Sci. 157, 131–142 (2006)
2. Fodor, J., Roubens, M.: Fuzzy Preference Modelling and Multicriteria Decision Support. Kluwer A.P., Dordrecht (1994)
3. Glover, F., Klingman, D.: The simplex SON method for LP/embedded network problems. Math. Progr. Study 15, 148–176 (1981)
4. Grabisch, M., Orlovski, S.A., Yager, R.R.: Fuzzy aggregation of numerical preferences. In: Fuzzy sets in decision analysis, operations research and statistics, pp. 31–68. Kluwer AP, Dordrecht (1999)
5. Kostreva, M.M., Ogryczak, W., Tonkyn, D.W.: Relocation Problems Arising in Conservation Biology. Computers Math. Appls. 37, 135–150 (1999)
6. Larsen, H.L.: Importance weighted OWA aggregation of multicriteria queries. In: Proc. North American Fuzzy Info. Proc. Soc. Conf. (NAFIPS 1999), pp. 740–744 (1999)
7. Liu, X.: Some properties of the weighted OWA operator. IEEE Trans. Systems, Man and Cyber. B 368, 118–127 (2006)
8. Mansini, R., Ogryczak, W., Speranza, M.G.: LP Solvable Models for Portfolio Optimization: A Classification and Computational Comparison. IMA J. Manage. Math. 14, 187–220 (2003)
9. Mansini, R., Ogryczak, W., Speranza, M.G.: Conditional Value at Risk and Related Linear Programming Models for Portfolio Optimization. Annals of Oper. Res. 152, 227–256 (2007)
10. Nettleton, D., Muniz, J.: Processing and representation of meta-data for sleep apnea diagnosis with an artificial intelligence approach. Medical Informatics 63, 77–89 (2001)

11. Ogryczak, W.: An Implementation of Variable Upper Bounds via SUB Methodology. J. Information & Optimization Sci. 13, 29–47 (1992)
12. Ogryczak, W.: Multiple criteria linear programming model for portfolio selection. Annals Oper. Res. 97, 143–162 (2000)
13. Ogryczak, W.: On Goal Programming Formulations of the Reference Point Method. J. Opnl. Res. Soc. 52, 691–698 (2001)
14. Ogryczak, W.: Multiple criteria optimization and decisions under risk. Control & Cyber. 31, 975–1003 (2002)
15. Ogryczak, W.: WOWA enhancement of the preference modeling in the Reference Point Method. In: Torra, V., Narukawa, Y. (eds.) MDAI 2008. LNCS (LNAI), vol. 5285, pp. 38–49. Springer, Heidelberg (2008)
16. Ogryczak, W., Ruszczyński, A.: Dual stochastic dominance and related mean-risk models. SIAM J. Optimization 13, 60–78 (2002)
17. Ogryczak, W., Śliwiński, T.: On equitable approaches to resource allocation problems: the conditional minimax solution. J. Telecom. Info. Tech. 3(02), 40–48 (2002)
18. Ogryczak, W., Śliwiński, T.: On solving linear programs with the ordered weighted averaging objective. Eur. J. Opnl. Res. 148, 80–91 (2003)
19. Ogryczak, W., Śliwński, T.: On Optimization of the importance weighted OWA aggregation of multiple criteria. In: Gervasi, O., Gavrilova, M.L. (eds.) ICCSA 2007, Part I. LNCS, vol. 4705, pp. 804–817. Springer, Heidelberg (2007)
20. Ogryczak, W., Śliwński, T.: On decision support under risk by the WOWA optimization. In: Mellouli, K. (ed.) ECSQARU 2007. LNCS (LNAI), vol. 4724, pp. 779–790. Springer, Heidelberg (2007)
21. Schrage, L.: Implicit representation of variable upper bounds in linear programming. Math. Programming Study 4, 118–132 (1975)
22. Torra, V.: The weighted OWA operator. Int. J. Intell. Syst. 12, 153–166 (1997)
23. Torra, V.: The WOWA operator and the interpolation function W*: Chen and Otto's interpolation method revisited. Fuzzy Sets Syst. 113, 389–396 (2000)
24. Valls, A., Torra, V.: Using classification as an aggregation tool for MCDM. Fuzzy Sets Syst. 115, 159–168 (2000)
25. Yager, R.R.: On ordered weighted averaging aggregation operators in multicriteria decision making. IEEE Trans. Systems, Man and Cyber. 18, 183–190 (1988)
26. Yager, R.R.: Constrained OWA aggregation. Fuzzy Sets Syst. 81, 89–101 (1996)
27. Yager, R.R.: Quantifier guided aggregation using OWA operators. Int. J. Intell. Syst. 11, 49–73 (1996)
28. Yager, R.R.: Including Importances in OWA Aggegations Using Fuzzy Systems Modeling. IEEE Trans. Fuzz. Syst. 6, 286–294 (1998)
29. Yager, R.R., Filev, D.P.: Essentials of Fuzzy Modeling and Control. Wiley, New York (1994)
30. Yager, R.R., Kacprzyk, J.: The Ordered Weighted Averaging Operators: Theory and Applications. Kluwer AP, Dordrecht (1997)

Fuzzy Dynamic Programming Problem for Extremal Fuzzy Dynamic System

Gia Sirbiladze

Abstract. This work deals with the problems of the Expremal Fuzzy Continuous Dynamic System (EFCDS) optimization problems and briefly discuss the results developed by G. Sirbiladze [31]–[38]. The basic properties of extended extremal fuzzy measure are considered and several variants of their representation are given. In considering extremal fuzzy measures, several transformation theorems are represented for extended lower and upper Sugeno integrals. Values of extended extremal conditional fuzzy measures are defined as a levels of an expert knowledge reflections of EFCDS states in the fuzzy time intervals. The notions of extremal fuzzy time moments and intervals are introduced and their monotone algebraic structures that form the most important part of the fuzzy instrument of modeling extremal fuzzy dynamic systems are discussed. New approaches in modeling of EFCDS are developed. Applying the results of [31] and [32], fuzzy processes with possibilistic uncertainty, the source of which is extremal fuzzy time intervals, are constructed. The dynamics of EFCDS's is described. Questions of the ergodicity of EFCDS's are considered. Fuzzy-integral representations of controllable extremal fuzzy processes are given. Sufficient and necessary conditions are presented for the existence of an extremal fuzzy optimal control processes, for which we use R. Bellman's optimality principle and take into consideration the gain-loss fuzzy process. A separate consideration is given to the case where an extremal fuzzy control process acting on the EFCDS does not depend on an EFCDS state. Applying Bellman's optimality principle and assuming that the gain-loss process exists for the EFCDS, a variant of the fuzzy integral representation of an optimal control is given for the EFCDS. This variant employs the instrument of extended extremal fuzzy composition measures constructed in [32]. An example of constructing of the EFCDS optimal control is presented.

Gia Sirbiladze
Department of Computer Sciences, I. Javakhishvili Tbilisi State University, 2, University St., Tbilisi 0143, Georgia
e-mail: gsirbiladze@tsu.ge

W.A. Lodwick & J. Kacprzyk (Eds.): Fuzzy Optimization, STUDFUZZ 254, pp. 231–270.
springerlink.com

1 Introduction

In recent years, both the dynamics of fuzzy system and the modeling issue received an increased attention. Dynamics is an obvious problem in control; moreover, its interest goes far beyond control applications. Applications of the dynamics of fuzzy systems and of the modeling of dynamic systems by fuzzy systems range from physics to biology to economics to pattern recognition and to time series prediction.

Evidence exists that fuzzy models can explain cooperative processes, such as in biology, chemistry, material sciences, or in economy. Relationships between dynamics of fuzzy systems and the performance of decision support systems were found, and chaotic processes in various classes of fuzzy systems were shown as a powerful tool in analyzing complex, weakly structurable systems, as anomal and extremal processes.

To make the decision-making effective in the framework of computer systems supporting this process, we must solve analytic problems of optimization, state evaluation, model identification, complex dynamic system control, optimal control, filtering and so on.

It is well recognized that optimization and other decision support technologies have been playing an important role in improving almost every aspect of human society. Intensive study over the past several years has resulted in significant progress in both the theory and applications of optimization and decision science.

Optimization and decision-making problems are traditionally handled by either the deterministic or probabilistic approaches. The former provides an approximate solution, completely ignoring uncertainty, while the latter assumes that any uncertainty can be represented as a probability distribution. Of course, both approaches only partially capture reality uncertainty (such as stock price, commodity, cost, natural resource availability and so on) that indeed exist but not in the form of known probability distributions.

In the Preface of the Journal of Fuzzy Optimization and Decision Making (vol. I, 2002, pp. 11–12) Professor L. A. Zadeh had said: "My 1970 paper with R.E. Bellman, "Decision-Making in a Fuzzy Environment" was intended to suggest a framework based on the theory of fuzzy sets for dealing with imprecision and partial truth in optimization and decision analysis. In the intervening years, a voluminous literature on applications of fuzzy logic to decision analysis has come into existence."

In alternative classical approaches to modeling and when working with complex systems the main accent is placed on the assumption of fuzziness. As the complexity of systems increases, our ability to define exactly their behaviour drops to a certain level, below which such characteristics of information as exactness and uncertainty become mutually excluding. In such situations an exact quantitative analysis of real complex systems is apt to be not quite plausible. Hence, a conclusion comes to mind that problems of this kind should be solved by means of analytic-fundamental methods of fuzzy mathematics, while the system approach to constructing models of complex systems with fuzzy uncertainty guarantees the creation of computer-aided systems forming the instrumental basis of intelligent technology solutions of

expert-analytic problems. It is obvious that the source of fuzzy-statistical samples is the population of fuzzy characteristics of expert knowledge. Fuzziness arises from observations of time moments as well as from other expert measurements.

Problems of making an optimal solution for systems with fuzzy uncertainty are difficult because it frequently happens that the controllable object possesses conflicting properties which might include:

1) imperfection of a control process due to information uncertainty;
2) unreliable elements of a control system;
3) nonuniqueness and the applicability of many criteria encountered in a control process;
4) restriction of possibilities (resources) of a control system;
5) loss of the ability of a control system to solve arising control problems.

Fuzzy programming problems have been discussed widely in literature ([1]–[3], [5], [7], [10], [11], [23], [25], [26], [35], [37], [39], [45]–[47] and so on) and applied in such various disciplines as operations research, economic management, business administration, engineering and so on. B. Liu [25] presents a brief review on fuzzy programming models, and classifies them into three broad classes: expected value models, chance-constrained programming and chance-dependent programming.

Our further study belongs to the first class, where we use the instrument of fuzzy measures and integrals ([8], [14]–[16], [31]–[33], [38], [40]–[42], [44] and so on) or, speaking more exactly, extremal fuzzy measures and Sugeno's type integrals along with extremal fuzzy expected values.

Therefore in the paper the new approach to the study of weakly structurable fuzzy dynamic systems optimization is presented (Extremal Fuzzy Continuous Dynamic System). This approach is based on the six papers published in the Int. Journal of General Systems (by G. Sirbiladze, "Modeling of Extremal Fuzzy Dynamic Systems". Parts I-VI: 34,2, 2005, 107-138; 139-167; 169-198; 35, 4, 2006, 435-459; 35, 5, 2006, 529-554; 36,1 2007, 19-58). Different from other approaches where the source of fuzzy uncertainty in dynamic systems is expert, this approach considers time as long as an expert to be the source of fuzzy uncertainty. This notably widens the area of studied problems. All these is connected to the incomplete, imprecise, anomal and extremal processes in nature and society, where connections between the system's objects are of subjective (expert) nature, which is caused by lack of objective information about the evolution of studied system, for example in 1) economics and business of developing countries, conflictology, sociology, medical diagnosis etc; 2) management of evacuation processes in catastrophe areas, estimation of disease spreading in epidemical regions; 3) research of complex systems of applied physics, etc. One of our purposes is to create scenarios describing possible evolution of EFCDS using methods of optimization developed in this paper by the framework of expert-possibilistic theory. This includes construction of algorithms of logical-possibilistic simulations of anomal and extremal process analysis.

Our attention is focused on the rapidly developing theory of fuzzy measures and integrals ([8], [14]–[16], [31]–[33], [38], [40]–[42], [44] and so on). The application of fuzzy measures and integrals as an instrument of constructing the

intelligent decision-making systems is not a novel idea ([8], [13]–[16], [18], [20], [22], [25], [29], [30], [33]–[37], [39]–[42], [44], [46] and so on). We employ the part of the theory of fuzzy measures which concerns extremal fuzzy measures ([31]–[33], [38]) and which, in our opinion, is rather seldom used. We have constructed a new instrument of a fuzzy measure, the extension of which is based on Sugeno lower and upper integrals.

In the framework of this theory a new apparatus of extended fuzzy measures was constructed on the basis of Sugeno's upper and lower integrals ([31]–[33], [38]). Using this apparatus new fuzzy extremal models of weakly structurable dynamic system control were created, where fuzziness is represented in time. Here the structure of time is represented by monotone extremal classes of measurable sets. On such structures uncertainty is described by extremal fuzzy measures and problems of fuzzy analysis of extremal fuzzy processes: 1. Fuzzy Optimization, 2. Fuzzy Identification, 3. Fuzzy Filtration and so on. We will deal with the fuzzy control problems of fuzzy dynamic systems (EFCDS) ([33]–[36], [39]), where fuzzy uncertainty arises with time and time structures are monotone classes of measurable sets.

As known (Subsection 2.2 and [31]), in fuzzy dynamic processes where fuzziness participates as a time factor, an important role is assigned to the structures of extremal fuzzy time intervals $\{\widetilde{\mathscr{F}}I_*(T), \succeq, \underset{*}{\otimes}\}\rangle)$, $(\langle\{\widetilde{\mathscr{F}}I^*(T), \preceq, \overset{*}{\otimes}\}$. As the fuzzy time flows, the process of expert knowledge measurement on the system states with respect to time is affected by the incompleteness of the obtained information. The polar characteristics of this information manifest themselves as imprecision and uncertainty. A degree of information imprecision is defined by current fuzzy time moments $(\widetilde{t} \in \widetilde{\mathscr{B}}_1^*)$ and future fuzzy time moments $(\widetilde{t} \in \widetilde{\mathscr{B}}_{1*})$, while an uncertainty degree is defined by current fuzzy time intervals $(\widetilde{r} \in \widetilde{\mathscr{B}}_2^*)$ and future fuzzy time intervals $(\widetilde{r} \in \widetilde{\mathscr{B}}_{2*})$. We have constructed the corresponding fuzzy monotone structures

$$\{\widetilde{\mathscr{F}}I_*(T), \succeq, \underset{*}{\oplus}\} \quad \text{and} \quad \{\widetilde{\mathscr{F}}I^*(T), \preceq, \overset{*}{\oplus}\}, \tag{1}$$

in which sequential extremal fuzzy time intervals are calculated recurrently.

Here only note that when expert describes the dynamics of complex objects and "measure" system states, where the source of uncertainty is fuzzy measurements with respect to time, it is necessary to carry out "extremal" "dual" measurements, namely, measurements in extended current and compressed future fuzzy time intervals [31].

In the present paper, we represent the controllable extremal fuzzy processes defined in [35]–[37], [39]. The subject/matter of our investigation is the existence of an optimal control for EFCDS's. Section 2 contains some necessary preliminary concepts presented in [31]–[33], [38]. Sections 3 and 4 describe models of extremal fuzzy continuous dynamic system. Section 5 deals with problems of EFCDS optimization when the control parameter does not depend on a state in which an EFCDS is. Questions of the existence of an optimal control are studied, and variants of their fuzzy integral representation are proposed. Section 6 contains an example in which the EFCDS fuzzy optimal control process is constructed.

2 Preliminary Concepts

All definitions and results see in [31]–[33], [38].

2.1 On the Space of Extended Extremal Fuzzy Measures

Definition 1. Let X be some nonempty set.

a) We call some class $\mathscr{B}^* \subset 2^X$ of subsets of X an upper σ^*-monotone class if (i) $\varnothing, X \in \mathscr{B}^*$; (ii) $\forall A, B \in \mathscr{B}^* \Rightarrow A \cup B \in \mathscr{B}^*$; (iii) $\forall \{A_n\} \in \mathscr{B}^*$, $n = 1,2,\ldots$, $A_n \uparrow A \Rightarrow A \in \mathscr{B}^*$.

b) We call some class $\mathscr{B}_* \subset 2^X$ of subsets of X a lower σ_*-monotone class if (i) $\varnothing, X \in \mathscr{B}_*$; (ii) $\forall A, B \in \mathscr{B}_* \Rightarrow A \cap B \in \mathscr{B}_*$; (iii) $\forall \{A_n\} \in \mathscr{B}_*$, $n = 1,2,\ldots$, $A_n \downarrow A \Rightarrow A \in \mathscr{B}_*$.

Definition 2. We call the classes $\mathscr{B}^*$ and $\mathscr{B}_*$ extremal if and only if

$$\forall A \in \mathscr{B}^* \Leftrightarrow \overline{A} \in \mathscr{B}_*.$$

Remark 1. Let $\mathscr{B} \subseteq 2^X$ be some σ-algebra. Then B is both a σ^*-monotone class and a σ_*-monotone class.

Definition 3. 1) $(X, \mathscr{B}^*)$ is called an upper measurable space;

2) $(X, \mathscr{B}_*)$ is called a lower measurable space;

3) If $\mathscr{B}^*$ and $\mathscr{B}_*$ are extremal σ^*- and σ_*-monotone classes, then $(X, \mathscr{B}_*, \mathscr{B}^*)$ is called an extremal measurable space.

Example 1

$$\mathscr{B}_1^* \stackrel{\Delta}{=} \left\{A \subset \mathbb{R}_0^+ \mid A = (\alpha; +\infty),\ \alpha \in \mathbb{R}_0^+\right\} \cup \{\varnothing\} \cup \{\mathbb{R}_0^+\} \text{ is a } \sigma^*\text{-monotone class,}$$

$$\mathscr{B}_{1*} \stackrel{\Delta}{=} \left\{A \subset \mathbb{R}_0^+ \mid A = [0; \alpha],\ \alpha \in \mathbb{R}_0^+\right\} \cup \{\varnothing\} \cup \{\mathbb{R}_0^+\} \text{ is a } \sigma_*\text{-monotone class.}$$

$\mathscr{B}_1^*$ and $\mathscr{B}_{1*}$ are respectively called a Borel σ^*-monotone class and a Borel σ_*-monotone class of first kind. Clearly, $\mathscr{B}_1^*$ and $\mathscr{B}_{1*}$ are extremal.

Example 2

$$\mathscr{B}_2^* \stackrel{\Delta}{=} \left\{A \subset \mathbb{R}_0^+ \mid A = [0; \alpha),\ \alpha \in \mathbb{R}_0^+\right\} \cup \{\varnothing\} \cup \{\mathbb{R}_0^+\} \text{ is a } \sigma^*\text{-monotone class,}$$

$$\mathscr{B}_{2*} \stackrel{\Delta}{=} \left\{A \subset \mathbb{R}_0^+ \mid A = [\alpha; +\infty),\ \alpha \in \mathbb{R}_0^+\right\} \cup \{\varnothing\} \cup \{\mathbb{R}_0^+\} \text{ is a } \sigma_*\text{-monotone class.}$$

$\mathscr{B}_2^*$ and $\mathscr{B}_{2*}$ are respectively called a Borel σ^*- and a Borel σ_*-monotone class of second kind. It is obvious that $\mathscr{B}_2^*$ and $\mathscr{B}_{2*}$ are extremal.

Definition 4. Let $(X, \mathscr{B}^*)$ be some upper measurable space. A function $g^* : \mathscr{B}^* \to [0;1]$ is called an upper fuzzy measure if: (i) $g^*(\varnothing) = 0$, $g^*(X) = 1$; (ii) $\forall A, B \in \mathscr{B}^*$,

$A \subset B \Rightarrow g^*(A) \leq g^*(B)$; (iii) $\forall\{A_n\} \in \mathscr{B}^*,\ n = 1,2,\ldots,\ A_n \uparrow A \Rightarrow g^*(A) = \lim\limits_{n\to\infty} g^*(A_n)$.

Definition 5. Let $(X,\mathscr{B}_*)$ be some lower measurable space. A function $g_* : \mathscr{B}_* \to [0;1]$ is called a lower fuzzy measure if: (i) $g_*(\varnothing) = 0$, $g_*(X) = 1$; (ii) $\forall A,B \in \mathscr{B}_*$, $A \subset B \Rightarrow g_*(A) \leq g_*(B)$; (iii) $\forall\{A_n\} \in \mathscr{B}_*$, $n = 1,2,\ldots,\ A_n \downarrow A \Rightarrow g_*(A) = \lim\limits_{n\to\infty} g_*(A_n)$.

Definition 6. Let $(X,\mathscr{B}_*,\mathscr{B}^*)$ be some extremal measurable space, g_* be a lower and g^* an upper fuzzy measure.

Then:

a) $g_* : \mathscr{B}_* \to [0;1]$ and $g^* : \mathscr{B}^* \to [0;1]$ is called extremal if and only if

$$\forall A \in \mathscr{B}_* \ :\ g_*(A) = 1 - g^*(\overline{A}).$$

b) $(X,\mathscr{B}_*,\mathscr{B}^*,g_*,g^*)$ is called a space of extremal fuzzy measures.

Definition 7. Let $(X_1,\mathscr{B}'_*,\mathscr{B}'^*)$ and $(X_2,\mathscr{B}''_*,\mathscr{B}''^*)$ be some extremal measurable spaces; $h : X_1 \to X_2$ is called measurable if

$$\forall A \in \mathscr{B}''^*,\ B \in \mathscr{B}''_* \ :\ h^{-1}(A) \in \mathscr{B}'^*,\ h^{-1}(B) \in \mathscr{B}'_*.$$

Definition 8. Let $(X,\mathscr{B}_*,\mathscr{B}^*)$ be some extremal measurable space. Then:

a) The function $h : X \to \mathbb{R}_0^*$ is called upper measurable if and only if h is measurable with respect to the spaces $(X,\mathscr{B}_*,\mathscr{B}^*)$ and $(\mathbb{R}_0^+,\mathscr{B}_{1*},\mathscr{B}_1^*)$. Then

$$\forall\alpha \geq 0 \quad h^{-1}((\alpha;+\infty)) \in \mathscr{B}^*,\quad h^{-1}([0;\alpha]) \in \mathscr{B}_*.$$

b) The function $h : X \to \mathbb{R}_0^+$ is called lower measurable if and only if h is measurable with respect to the spaces $(X,\mathscr{B}_*,\mathscr{B}^*)$ and $(\mathbb{R}_0^+,\mathscr{B}_{2*},\mathscr{B}_2^*)$. Then

$$\forall\alpha \geq 0 \quad h^{-1}([0;\alpha)) \in \mathscr{B}^*,\quad h^{-1}([\alpha;+\infty)) \in \mathscr{B}_*.$$

Definition 9. Let $(X,\mathscr{B}_*,\mathscr{B}^*)$ be some extremal measurable space.

a) The class of fuzzy subsets $\widetilde{A} \subset X$ with lower measurable compatibility functions

$$\begin{aligned}\widetilde{\mathscr{B}}_* &= \left\{\widetilde{A} \subset X \mid \mu_{\widetilde{A}} \text{ is lower measurable}\right\}\\ &= \left\{\widetilde{A} \in X \mid \forall\, 0 \leq \alpha \leq 1,\ \mu_{\widetilde{A}}^{-1}([0;\alpha)) \in \mathscr{B}^*,\ \mu_{\widetilde{A}}^{-1}([\alpha;+\infty)) \in \mathscr{B}^*\right\}\end{aligned}$$

is called an extension of the σ_*-monotone class $\mathscr{B}_*$.

b) The class of fuzzy subsets $\widetilde{A} \subset X$ with upper measurable compatibility functions

$$\widetilde{\mathscr{B}}^* = \left\{\widetilde{A} \subset X \mid \mu_{\widetilde{A}} \text{ is upper measurable}\right\}$$
$$= \left\{\widetilde{A} \in X \mid \forall\, 0 \le \alpha \le 1,\ \mu_{\widetilde{A}}^{-1}([0;\alpha]) \in \mathscr{B}_*,\ \mu_{\widetilde{A}}^{-1}((\alpha;+\infty)) \in \mathscr{B}^*\right\}$$

is called an extension of the σ^*-monotone class $\mathscr{B}^*$.

Definition 10. An extremal measurable space $(X, \widetilde{\mathscr{B}}_*, \widetilde{\mathscr{B}}^*)$ is called an extension of an extremal measurable space $(X, \mathscr{B}_*, \mathscr{B}^*)$.

Using the Sugeno integral, we next introduce the notion of extension of fuzzy extremal measures.

Definition 11. Let $(X, \mathscr{B}_*, \mathscr{B}^*, g_*, g^*)$ be some space of extremal fuzzy measures, and $(X, \widetilde{\mathscr{B}}_*, \widetilde{\mathscr{B}}^*)$ be an extension of the extremal measurable space $(X, \mathscr{B}_*, \mathscr{B}^*)$. Then:

a) the function

$$\widetilde{g}_*(\widetilde{A}) \equiv \int\limits_{X}^{*} \mu_{\widetilde{A}}(x) \circ g_*(\cdot) \stackrel{\Delta}{=} \bigvee_{0<\alpha\le 1} \left[\alpha \wedge g_*([\widetilde{A}]_{\bar{\alpha}})\right], \quad \forall \widetilde{A} \in \widetilde{\mathscr{B}}_*; \tag{2}$$

is called an extension of the fuzzy measure g_* on $\widetilde{\mathscr{B}}_*$;

b) the function

$$\widetilde{g}^*(\widetilde{A}) \equiv \int\limits_{X}^{*} \mu_{\widetilde{A}}(x) \circ g^*(\cdot) \stackrel{\Delta}{=} \bigwedge_{0<\alpha\le 1} \left[\alpha \vee g^*([\widetilde{A}]_{\alpha})\right], \quad \forall \widetilde{A} \in \widetilde{\mathscr{B}}^*, \tag{3}$$

is called an extension of the fuzzy measure g^* on $\widetilde{\mathscr{B}}^*$.

Here $[\widetilde{A}]_\alpha = \{x \in X \mid \mu_{\widetilde{A}}(x) > \alpha\}$, $[\widetilde{A}]_{\bar{\alpha}} = \{x \in X \mid \mu_{\widetilde{A}}(x) \ge \alpha\}$, $0 < \alpha \le 1$.

Definition 12. A space of extremal fuzzy measures $(X, \widetilde{\mathscr{B}}_*, \widetilde{\mathscr{B}}^*, \widetilde{g}_*, \widetilde{g}^*)$ is called an extension of the space $(X, \mathscr{B}_*, \mathscr{B}^*, g_*, g^*)$.

Let $(X, \mathscr{B}_*, \mathscr{B}^*, g_*, g^*)$ be some space of extremal fuzzy measures and $(X, \widetilde{\mathscr{B}}_*, \widetilde{\mathscr{B}}^*, \widetilde{g}_*, \widetilde{g}^*)$ be its extension.

Definition 13. a) Let $\widetilde{A}, \widetilde{B} \in \widetilde{\mathscr{B}}_*$ be any fuzzy sets. Then the lower fuzzy Sugeno integral of the compatibility function $\mu_{\widetilde{B}}$ on the fuzzy set $\widetilde{A}$ is defined with respect to a lower fuzzy measure $\widetilde{g}_*$ by the formula

$$\int\limits_{\widetilde{A}}^{*} \mu_{\widetilde{B}}(x) \circ \widetilde{g}_*(\cdot) \stackrel{\Delta}{\equiv} \bigvee_{0<\alpha\le 1} \left[\alpha \wedge \widetilde{g}_*(\widetilde{A} \cap [\widetilde{B}]_{\bar{\alpha}})\right]. \tag{4}$$

b) Let $\widetilde{A}, \widetilde{B} \in \widetilde{\mathscr{B}}^*$ be any fuzzy sets. Then the upper fuzzy Sugeno integral of the compatibility function $\mu_{\widetilde{B}}$ on the fuzzy set $\widetilde{A}$ is defined with respect to a upper fuzzy measure $\widetilde{g}^*$ by the formula

$$\int\limits_{\widetilde{A}}^{*} \mu_{\widetilde{B}}(x) \circ \widetilde{g}^{*}(\cdot) \stackrel{\Delta}{=} \bigwedge_{0<\alpha\leq 1} \left[\alpha \vee \widetilde{g}^{*}(\widetilde{A} \cup [\widetilde{B}]_{\alpha})\right]. \tag{5}$$

Definition 14. Let $(X, \mathscr{B}_{*}, \mathscr{B}^{*}, g_{*}, g^{*})$ be some space of extremal fuzzy measures.

a) Let $h \in \widetilde{\mathscr{B}}_{*}$ be some fuzzy set. The measure

$$\forall A \in \widetilde{\mathscr{B}}_{*} \; : \; \widetilde{g}_{h*}(\widetilde{A}) \stackrel{\Delta}{=} \int\limits_{\widetilde{A}}^{*} \mu_{h}(x) \circ \widetilde{g}_{*}(\cdot) = \int\limits_{h}^{*} \mu_{\widetilde{A}}(x) \circ \widetilde{g}_{*}(\cdot) = \int\limits_{X}^{*} \mu_{h\cap\widetilde{A}}(x) \circ \widetilde{g}_{*}(\cdot) \tag{6}$$

is called the lower extension of g_{*} on $\widetilde{\mathscr{B}}_{*}$ with respect to h.

b) Let $h \in \widetilde{\mathscr{B}}^{*}$ be some fuzzy set. The measure

$$\forall A \in \widetilde{\mathscr{B}}^{*} \; : \; \widetilde{g}^{*}_{h}(\widetilde{A}) \stackrel{\Delta}{=} \int\limits_{\widetilde{A}}^{*} \mu_{h}(x) \circ \widetilde{g}^{*}(\cdot) = \int\limits_{h}^{*} \mu_{\widetilde{A}}(x) \circ \widetilde{g}^{*}(\cdot) = \int\limits_{X}^{*} \mu_{h\cup\widetilde{A}}(x) \circ \widetilde{g}^{*}(\cdot) \tag{7}$$

is called the upper extension of g^{*} on $\widetilde{\mathscr{B}}^{*}$ with respect to h.

2.2 On the Composition Products of Spaces of Extremal Fuzzy Measures

Let $(X_1, \mathscr{B}'_{*}, \mathscr{B}'^{*}, g'_{*}, g'^{*})$ and $(X_2, \mathscr{B}''_{*}, \mathscr{B}''^{*}, g''_{*}, g''^{*})$ be any two spaces of extremal fuzzy measures.

Definition 15. Let some subset $H \subset X_1 \times X_2$ be a binary relation. We introduce the following mappings $\forall x_0 \in X_1$ and $\forall y_0 \in X_2$:

$$\begin{aligned} E_H(x_0, \cdot) &\stackrel{\Delta}{=} \{y \in X_2 \mid (x_0, y) \in H\}, \\ E_H(\cdot, y_0) &\stackrel{\Delta}{=} \{x \in X_1 \mid (x, y_0) \in H\}. \end{aligned} \tag{8}$$

a) A binary relation $H \subset X_1 \times X_2$ is called lower measurable if $\forall A \in \mathscr{B}''_{*}$ and $\forall B \in \mathscr{B}'_{*}$ there exist sequences $\{x_n\}_{n\in N} \subset B$, $\{y_n\}_{n\in N} \subset A$ such that $E_H(x_n, \cdot) \supset E_H(x_{n+1})$, $E_H(\cdot, y_n) \supset E_H(\cdot, y_{n+1})$, $n = 1, 2, \ldots$. We have

$$\Gamma_{H*}(A) \stackrel{\Delta}{=} \{x \in X_1 \mid \forall y \in A : (x, y) \in H\} \equiv \bigcap_{y\in A} E_H(\cdot, y) = \bigcap_{n=1}^{\infty} E_H(\cdot, y_n) \in \mathscr{B}'_{*} \tag{9}$$

and

$$\Gamma'_{H*}(B) \stackrel{\Delta}{=} \{y \in X_2 \mid \forall x \in B : (x, y) \in H\} \equiv \bigcap_{x\in B} E_H(x, \cdot) = \bigcap_{n=1}^{\infty} E_H(x_n, \cdot) \in \mathscr{B}''_{*}. \tag{10}$$

b) Denote by $\mathscr{B}'_* \otimes \mathscr{B}''_*$ the set of all binary lower measurable relations from $X_1 \times X_2$ and call it the composition product of measurable spaces $\mathscr{B}'_*$ and $\mathscr{B}''_*$.

a′) A binary relation $H \subset X_1 \times X_2$ is called upper measurable if $\forall A \in \mathscr{B}''^*$ and $\forall B \in \mathscr{B}'^*$ there exist sequences $\{x_n\}_{n\in N} \subset B$, $\{y_n\}_{n\in N} \subset A$ such that $E_H(x_n,\cdot) \subset E_H(x_{n+1})$, $E_H(\cdot,y_n) \subset E_H(\cdot,y_{n+1})$, $n = 1,2,\ldots$. We have

$$\Gamma^*_H(A) \stackrel{\Delta}{=} \{x \in X_1 \mid \exists y \in A : (x,y) \in H\} \equiv \bigcup_{y\in A} E_H(\cdot,y) = \bigcup_{n=1}^{\infty} E_H(\cdot,y_n) \in \mathscr{B}'^* \quad (11)$$

and

$$\begin{aligned}\Gamma'^*_H(B) &\stackrel{\Delta}{=} \{y \in X_2 \mid \exists x \in B : (x,y) \in H\} \\ &\equiv \bigcup_{x\in B} E_H(x,\cdot) = \bigcup_{n=1}^{\infty} E_H(x_n,\cdot) \in \mathscr{B}''^*. \end{aligned} \quad (12)$$

b′) Denote by $\mathscr{B}'^* \otimes \mathscr{B}''^*$ the set of all binary upper measurable relations from $X_1 \times X_2$ and call it the composition product of measurable spaces $\mathscr{B}'^*$ and $\mathscr{B}''^*$.

It is not difficult to verify that $\mathscr{B}'_* \otimes \mathscr{B}''_*$ is a lower σ_*-monotone class and $\mathscr{B}'^* \otimes \mathscr{B}''^*$ is a upper σ^*-monotone class.

Theorem 1. *Let $(X_1,\mathscr{B}'_*,g'_*)$ and $(X_2,\mathscr{B}''_*,g''_*)$ be two spaces of lower fuzzy measures. Then on the composition lower measurable space $(X_1 \times X_2,\mathscr{B}'_* \otimes \mathscr{B}''_*)$ the measure $g_* : \forall H \in \mathscr{B}'_* \otimes \mathscr{B}'_*$ defined by*

$$\begin{aligned} g_*(H) \equiv g'_* \otimes g''_*(H) &\stackrel{\Delta}{=} \bigvee_{E\in\mathscr{B}'_*} \left\{g'_*(E) \wedge g''_*(\Gamma'_{H*}(E))\right\} \\ &\equiv \bigvee_{F\in\mathscr{B}''_*} \left\{g'_*(\Gamma_{H*}(F)) \wedge g''_*(F)\right\} \end{aligned} \quad (13)$$

is a lower fuzzy measure.

Theorem 2. *Let $(X_1,\mathscr{B}'^*,g'^*)$ and $(X_2,\mathscr{B}''^*,g''^*)$ be two spaces of upper fuzzy measures. Then, on the composition upper measurable space $(X_1 \times X_2,\mathscr{B}'^* \otimes \mathscr{B}''^*)$, the measure $g^* : \forall H \in \mathscr{B}'^* \otimes \mathscr{B}''^*$ defined by*

$$\begin{aligned} g^*(H) \equiv g'^* \otimes g''^*(H) &\stackrel{\Delta}{=} \bigwedge_{E\in\mathscr{B}'^*} \left\{g'^*(E) \vee g''^*(\Gamma'^*_H(E))\right\} \\ &= \bigwedge_{F\in\mathscr{B}''^*} \left\{g'^*(\Gamma^*_H(F)) \vee g''^*(F)\right\} \end{aligned} \quad (14)$$

is an upper fuzzy measure.

Theorem 3. a) *Let $H \in \mathscr{B}'_* \otimes \mathscr{B}''_*$ be some binary lower measurable relation ($H \subset X_1 \times X_2$). Then the value of the measure $g'_* \otimes g''_*$ on H is represented through g'_* and g''_* as the following composition:*

$$g'_* \otimes g''_*(H) = \int^*_{X_2} g'_*(E_H(\cdot,y)) \circ g''_*(\cdot) = \int^*_{X_1} g''_*(E_H(x,\cdot)) \circ g'_*(\cdot); \quad (15)$$

b) *Let $H \in \mathscr{B}'^* \otimes \mathscr{B}''^*$ be some binary upper measurable relation. Then the value of the measure $g'^* \otimes g''^*$ on H is represented through g'^* and g''^* as the following composition:*

$$g'^* \otimes g''^*(H) = \overset{*}{\int\!\!\!\int}_{X_2} g'^*(E_H(\cdot, y)) \circ g''^*(\cdot) = \overset{*}{\int\!\!\!\int}_{X_1} g''^*(E_H(x, \cdot)) \circ g'^*(\cdot). \tag{16}$$

Now let us proceed to defining fuzzy binary relations on $X_1 \times X_2$.

Definition 16. a) A fuzzy set $\widetilde{H} \subset X_1 \times X_2$ is called a lower fuzzy binary relation if the compatibility function $\mu_{\widetilde{H}} : X_1 \times X_2 \to [0;1]$ is lower measurable;

b) A fuzzy set $\widetilde{H} \subset X_1 \times X_2$ is called an upper fuzzy binary relation if the compatibility function $\mu_{\widetilde{H}}$ is upper measurable.

We have constructed the compositional space of extremal extended fuzzy measures $(X_1 \times X_2, \widetilde{\mathscr{B}}'_* \otimes \widetilde{\mathscr{B}}''_*, \widetilde{\mathscr{B}}^{*\prime} \otimes \widetilde{\mathscr{B}}^{*\prime\prime}, \widetilde{g}'_* \otimes \widetilde{g}''_*, \widetilde{g}^{*\prime} \otimes \widetilde{g}^{*\prime\prime})$.

2.3 Extremal Fuzzy Time Moments and Intervals, and Their Structures

The questions investigated in the preceding paragraphs enable us to consider some extremal interval structures, in particular, extremal fuzzy time moments and intervals.

We would like to say just a few words about the origination of these structures and their importance in studying dynamic processes.

A person who makes a decision always gives an "incomplete" prognosis about a time moment for extremal, crisis, anomalous and other situations that may occur in the future. The person (expert) who makes a decision connects all such situations with future fuzzy time moments and intervals. Clearly, his/her prognosis is of fuzzy nature and the corresponding decisions should be obtained by possibilistic-statistical analysis or, speaking more exactly, by analysis of fuzzy time intervals, for which we need to construct a new mathematical fuzzy instrument.

When we make decisions on the basis of our past knowledge, we recall certain facts, reference data and the like. When doing so, we perform certain "expert measurements" ("expert reflections") of our knowledge. These measurements are connected with past time moments, which as a rule are fuzzy. Hence the results of such "measurements" may frequently be also fuzzy and these results of recollections are in the end reflected in experimental data (samples). It is understood that the source of such samples is the population of fuzzy characteristics of our knowledge. This can be explained mainly by the following two reasons: first, in terms of dynamics, moments of recollections of facts and moments of "expert measurements" are fuzzy moments; second, on frequent occasions the results of "measurements" are fuzzy. Let us illustrate this viewpoint by examples. Suppose that prior to diagnosing the disease the examining physician (expert) asked the patient to present data on his temperature distribution in time. If the patient measured his temperature but

for various reasons did not record the time of measurements, then his replies would sound like this: "In the morning my temperature varied approximately from 38°C to 38.5°C, at noon it dropped to something like 37° and in the evening it was not higher than 39°". Clearly, the results of such "measurements" are fuzzy both in time and in numerical values. It might happen that the patient made measurements of his temperature during the whole day (measurement results are objective data with uncertainty of probabilistic-statistical nature), but he did not record the time moments at which his temperature was measured. Therefore, when asking the patient to present this information in dynamics, we deal with fuzzy time moments. In such situations objective data are characterized by possibilistic uncertainty.

It is clear that decisions (prognoses) made about a future state of the object (prognoses) on the basis of such data by means of the classical statistical methods are less plausible for one reason: the source from which data of this kind originate is the person. The nature of data uncertainty is dual. It is only statistical-and-possibilistic methods that can give us more or less plausible estimates and prognoses.

With this aim in view, we begin our study of fuzzy time moments and intervals and their structures. For convenience, the observation time is identified with the set of nonnegative real numbers: $T \stackrel{\Delta}{=} \mathbb{R}_0^+$. Any time moment $t \in T = \mathbb{R}_0^+$ is assumed to be a nonnegative number.

Our notion of a fuzzy time moment is based on the definition presented in [8].

Definition 17. A fuzzy nonnegative real number $\widetilde{t}$ with the compatibility function

$$\mu_{\widetilde{t}} : \mathbb{R}_0^+ \to [0;1] \tag{17}$$

with the following properties:

(i) $\mu_{\widetilde{t}}(0) = 0$;

(ii) $\bigvee\limits_{\tau \geq 0} \mu_{\widetilde{t}}(\tau) = 1$ (normed);

(iii) $\forall \tau_0 \in \mathbb{R}_0^+$, $\mu_{\widetilde{t}}(\tau_0) = \bigvee\limits_{\tau < \tau_0} \mu_{\widetilde{t}}(\tau)$ (left continuity);

(iv) $\mu_{\widetilde{t}}(\tau)$ is a nonincreasing function on $\mathbb{R}_0^+ \equiv T$,

is called a fuzzy time moment.

The set of all fuzzy time moments is denoted by $\widetilde{\mathscr{F}M}_0(\mathbb{R}_0^+) \equiv \widetilde{\mathscr{F}M}_0^*(T)$.

Now, let us consider the extremal measurable Borel space of first kind $(\mathbb{R}_0^+, \mathscr{B}_{1*}, \mathscr{B}_1^*)$ and its extension $(\mathbb{R}_0^+, \widetilde{\mathscr{B}}_{1*}, \widetilde{\mathscr{B}}_1^*)$. If $\widetilde{a} \in \widetilde{\mathscr{B}}_1^*$ is a fuzzy number, then $\forall \alpha \geq 0$, $\mu_{\widetilde{a}}^{-1}((\alpha;+\infty)) \equiv (\tau;+\infty) \in \mathscr{B}_1^*$ and $\mu_{\widetilde{a}}^{-1}([0;\alpha]) \equiv [0,\tau] \in \mathscr{B}_{1*}$, i.e., $\mu_{\widetilde{a}}$ is an upper measurable function (or $\mu_{\widetilde{a}} : \mathscr{B}_1^* \to \mathscr{B}_1^*$, $\mathscr{B}_{1*} \to \mathscr{B}_{1*}$ is measurable). It is not difficult to verify that the compatibility function of the fuzzy moment $\widetilde{t}$ is upper measurable, i.e., the fuzzy time moment $\widetilde{t}$ is an upper fuzzy number on $T = \mathbb{R}_0^+$. We obtain $\widetilde{\mathscr{F}M}_0(\mathbb{R}_0^+) \subset \widetilde{\mathscr{B}}_1^*$.

Let us consider the negation of the fuzzy moment $\widetilde{t}$. It clearly follows that $\overline{\widetilde{t}} \in \widetilde{\mathscr{B}}_{1*}$ or $\forall \alpha \geq 0$, $\mu_{\overline{\widetilde{t}}}^{-1}([\alpha;+\infty)) \equiv [0,\tau] \in \mathscr{B}_{1*}$ and $\mu_{\overline{\widetilde{t}}}([0;\alpha)) \equiv (\tau;+\infty) \in \mathscr{B}_1^*$, where $\mu_{\overline{\widetilde{t}}}$ is lower measurable.

In terms of information, the negation of the fuzzy time moment $\bar{\bar{\tilde{t}}}$ can be interpreted as follows: it describes a measurement time medium, where the fuzzy time moment $\tilde{t}$ is excluded.

The relation between the time moment t and the time interval $[0;\tau)$ (and, accordingly, $[\tau;+\infty)$) is one-to-one:

$$t \in [0;\tau) \Longleftrightarrow t \notin [\tau;+\infty).$$

Therefore we may suppose that there exists a relation between the fuzzy time moment $\tilde{t}$ and the intervals $[0;\tau)$ and $[\tau;+\infty)$. As indicated in [31], for the fuzzy time moment $\tilde{t}$ its compatibility level $\mu_{\tilde{t}}(\tau)$, $\tau \geq 0$, is understood as a level of belonging of the fuzzy time moment $\tilde{t}$ to the time interval $[0;\tau)$ (a compatibility level). Our interpretation is as follows: $\mu_{\tilde{t}}(\tau)$ is a level of "measurement" imprecision, a level of finding the fuzzy time moment $\tilde{t}$ in the time interval $[0;\tau)$. A high compatibility level $\mu_{\tilde{t}}(\tau)$ gives more plausibility that the fuzzy time moment $\tilde{t}$ "is measured" up to the real moment τ in the time interval $[0;\tau)$. We call this interval the current time interval. Formally, it can be written as $\forall \tau \geq 0$

$$\mu_{\tilde{t}}(\tau) := \langle \text{an imprecise measure of } (\tilde{t} \in [0;\tau) := \text{the current time interval}) \rangle. \quad (18)$$

Now let us consider the class of complements to fuzzy time moments $\bar{\bar{\tilde{t}}}$. Since $\tilde{t} \in \widetilde{\mathscr{F}M}_0^*(T) \subset \mathscr{B}_1^*$, we denote this class by $\widetilde{\mathscr{F}M}_{0*}(T) \subset \widetilde{\mathscr{B}}_{1*}$. We call $\widetilde{\mathscr{F}M}_0^*(T)$ the class of upper fuzzy time moments, and $\widetilde{\mathscr{F}M}_{0*}(T)$ the class of lower fuzzy time moments.

Extending the above arguments to lower fuzzy time moments, we say that for a fuzzy time moment $\bar{\bar{\tilde{t}}}$ its compatibility level $\mu_{\tilde{t}}(\tau)$ is understood as a level of belonging of the fuzzy time moment $\tilde{t}$ to the interval $[\tau;+\infty)$, i.e., $\mu_{\bar{\bar{E}}}(\tau)$ is an imprecision level of measurement or, in other words, a level of finding a fuzzy time moment $\tilde{t}$ in the time interval $[\tau;+\infty)$. A high compatibility level $\mu_{\tilde{t}}(\tau)$ makes it more plausible that the fuzzy time moment $\tilde{t}$ will be "measured" after the real moment τ in the time interval $[\tau;+\infty)$, which we call the future time interval. If $\tilde{t} \in \widetilde{\mathscr{F}M}_{0*}(T)$, then $\forall \alpha \geq 0$, $\mu_{\tilde{t}}^{-1}([\alpha;+\infty)) = [0;\tau] \in \mathscr{B}_{1*}$, $\mu_{\tilde{t}}^{-1}([0;\alpha]) = [\tau;+\infty)$, i.e., $\mu_{\tilde{t}}$ is a $\mathscr{B}_{1*} \to \mathscr{B}_{2*}$, $\mathscr{B}_1^* \to \mathscr{B}_2^*$-measurable function.

We call the moment $\tilde{t} \in \widetilde{\mathscr{F}M}_{0*}(T) \subset \widetilde{\mathscr{B}}_{1*}$ a lower fuzzy time moment, while $\tilde{t} \in \widetilde{\mathscr{F}M}_{0*}(T)$ and $\bar{\bar{\tilde{t}}} \in \widetilde{\mathscr{F}M}_0^*(T)$ extremal fuzzy time moments.

If $\tilde{t} \in \widetilde{\mathscr{F}M}_{0*}(T)$, then, formally, this can be written as follows:

$$\begin{aligned} \mu_{\tilde{t}}(\tau) := \langle\, &\text{an imprecise measure of } (\tilde{t} \in [\tau;+\infty) \\ &:= \text{is the future time interval}) \rangle, \quad \tau \geq 0. \end{aligned} \quad (19)$$

In the process of expert measurement with respect to time the values of the compatibility functions $\mu_{\tilde{t}}(\tau)$ and $\mu_{\bar{\bar{t}}}(\tau)$, $\tau \geq 0$, are degrees of imprecision of finding the fuzzy time moment $\tilde{t}$ in the future time interval $([\tau;+\infty))$ and the current time interval $([0;\tau))$, respectively.

When we discuss fuzzy time moments in the process of time flow, we should specially mention the pair of extremal fuzzy time moments $(\overline{\tilde{t}},\underline{\tilde{t}})$. By the measurement of a fuzzy moment with respect to the real time τ we understand its measurement in the current time interval $[0;\tau)$ and in the future time interval $[\tau;+\infty)$ by (18) and (19). The extremal classes of fuzzy time moments $\widetilde{\mathscr{F}M}_{0*}(T)$ and $\widetilde{\mathscr{F}M}_0^*(T)$ are the classes of complementary fuzzy time moments

$$\tilde{t} \in \widetilde{\mathscr{F}M}_{0*}(T) \Leftrightarrow \overline{\underline{\tilde{t}}} \in \widetilde{\mathscr{F}M}_0^*(T).$$

Let us consider the structures of current and future fuzzy time intervals. By Definition 3 (Example 2) we know that

$$\mathscr{B}_2^* \stackrel{\Delta}{\equiv} \{[0;\tau),\ \tau \geq 0\} \quad \text{and} \quad \mathscr{B}_{2*} \stackrel{\Delta}{\equiv} \{\tau;+\infty),\ \tau \geq 0\}$$

are Borel σ_*- and σ^*-algebras of second kind. Clearly, the spaces of current and future time intervals are measurable or, speaking more exactly, coincide with extremal Borel spaces of second kind $(\mathbb{R}_0^+,\mathscr{B}_{2*},\mathscr{B}_2^*)$.

Further, we introduce the notion of extremal fuzzy time interval in terms of extension $(\mathbb{R}_0^+,\widetilde{\mathscr{B}}_{2*},\widetilde{\mathscr{B}}_2^*)$.

Definition 18. a) Any fuzzy positive number $\tilde{r} \equiv \widetilde{[0,\tau)} \in \widetilde{\mathscr{B}}_2^*$ is called an extended fuzzy current time interval.

b) Any fuzzy positive number $\tilde{r} \equiv \widetilde{[\tau;+\infty)} \in \widetilde{\mathscr{B}}_{2*}$ is called an extended fuzzy future time interval.

Obviously, if $\tilde{r} \in \widetilde{\mathscr{B}}_2^*$, then $\forall\alpha \geq 0$ we have $\mu_{\tilde{r}}^{-1}([\alpha;+\infty)) \equiv [0,t) \in \mathscr{B}_2^*$, $\mu_{\tilde{r}}^{-1}([0;\alpha]) \equiv [t;+\infty) \in \mathscr{B}_{2*}$, i.e., $\mu_{\tilde{r}}$ is the $\mathscr{B}_{2*} \to \mathscr{B}_{1*}$, $\mathscr{B}_2^* \to \mathscr{B}_1^*$-measurable function and if $\tilde{r} \in \widetilde{\mathscr{B}}_{2*}$, then $\forall\alpha \geq 0$ we have $\mu_{\tilde{r}}^{-1}([\alpha;+\infty)) \equiv [t;+\infty)$, $\mu_{\tilde{r}}^{-1}([0;\alpha)) \equiv [0;t) \in \mathscr{B}_{2*}$, i.e., $\mu_{\tilde{r}}$ is the $\mathscr{B}_{2*} \to \mathscr{B}_{2*}$, $\mathscr{B}_2^* \to \mathscr{B}_2^*$-measurable function. The fuzzy intervals $\tilde{r} \in \widetilde{\mathscr{B}}_2^*$ and $\overline{\underline{\tilde{r}}} \in \mathscr{B}_{2*}$ are called extremal.

Let us discuss the relation between fuzzy extremal time moments and intervals. Let $\tilde{t} \in \widetilde{\mathscr{F}M}_0^*(T)$ and $\tilde{r} \in \widetilde{\mathscr{B}}_2^*$ be respectively the fuzzy current time moment and the future fuzzy time interval. As has been mentioned above, $\mu_{\tilde{t}}(\tau)$ is a degree of imprecision of finding the fuzzy moment $\tilde{t}$ in the current time interval $[0;\tau)$ in the process of time flow. We think that the value $\mu_{\tilde{r}}(\tau)$ defines the level of compatibility that the current fuzzy time interval $\tilde{r}$ is not covered by the current time interval $[0;\tau)$. Moreover, $\mu_{\tilde{r}}(\tau)$ is a degree of uncertainty that $\tilde{r} \not\subset [0;\tau)$. $\forall\tau \geq 0$:

$$\begin{aligned}\mu_{\tilde{r}}(\tau) := \langle\, &\text{an uncertainty measure of } (\tilde{r} \not\subset [0;\tau) \\ &:= \text{the current time interval})\rangle. \end{aligned} \tag{20}$$

Let $\tilde{t} \in \widetilde{\mathscr{F}M}_{0*}(T)$ and $\tilde{r} \in \widetilde{\mathscr{B}}_{2*}$ be the fuzzy future time moment and the fuzzy time interval, respectively. As has been mentioned above, $\mu_{\tilde{t}}(\tau)$ is a degree of imprecision of finding the fuzzy moment $\tilde{t}$ in the future time interval $[\tau;+\infty)$ in the process of

time flow. We think that the value $\mu_{\widetilde{r}}(\tau)$ defines the level of compatibility that the fuzzy future time interval $\widetilde{r}$ is not covered by the future time interval $[\tau;+\infty)$. More exactly, $\mu_{\widetilde{r}}(\tau)$ is a degree of uncertainty that $\widetilde{r} \not\subset [\tau;+\infty)$ and $\forall \tau \geq 0$

$$\mu_{\widetilde{r}}(\tau) := \langle \text{ an uncertainty measure of } (\widetilde{r} \not\subset [\tau;+\infty) := \text{the future time interval}) \rangle.$$

Note that in the time flow process, the values of the compatibility function of extended extremal fuzzy time intervals $\widetilde{r} \in \widetilde{\mathscr{B}}_2^*$ and $\overline{\widetilde{r}} \in \widetilde{\mathscr{B}}_{2*}$ are degrees of uncertainty that these intervals do not belong to the respective current and future time intervals $[0;\tau)$ and $[\tau;+\infty)$. When speaking of the calculus of fuzzy time intervals, we will mean the pair of extremal fuzzy time intervals $(\widetilde{r},\overline{\widetilde{r}})$, where $\widetilde{r}$ is the current fuzzy time interval $(\widetilde{r} \in \widetilde{\mathscr{B}}_2^*)$, and $\overline{\widetilde{r}}$ is the future fuzzy time interval $(\overline{\widetilde{r}} \in \widetilde{\mathscr{B}}_{2*})$.

In the sequel, we will make use of the following concrete subclass of extended extremal fuzzy time intervals.

Definition 19. The class of fuzzy nonnegative numbers $\widetilde{\mathscr{F}I}^*(T)$ with the properties $(\widetilde{r} \in \widetilde{\mathscr{F}I}^*(T))$:

(i) $\mu_{\widetilde{r}}(0) = 1$;

(ii) $\forall \tau_0 \geq 0$, $\mu_{\widetilde{r}}(\tau_0) = \bigvee\limits_{\tau > \tau_0} \mu_{\widetilde{r}}(\tau)$ (right continuity);

(iii) $\mu_{\widetilde{r}}$ is nonincreasing on $T = \mathbb{R}_0^+$,

is called the class of current fuzzy time intervals $\widetilde{r}$.

It is not difficult to verify that $\mathscr{F}I^*(T)$ is a subclass of the space of extended fuzzy current time intervals $\widetilde{\mathscr{F}I}^*(T) \subset \widetilde{\mathscr{B}}_2^*$.

Analogously, we introduce the definition of the class $\widetilde{\mathscr{F}I}_*(T)$, which is a complement to $\widetilde{\mathscr{F}I}^*(T)$, i.e.,

$$\widetilde{r} \in \widetilde{\mathscr{F}I}_*(T) \subset \widetilde{\mathscr{B}}_{2*} \Leftrightarrow \overline{\widetilde{r}} \in \widetilde{\mathscr{F}I}^*(T) \subset \widetilde{\mathscr{B}}_2^*.$$

Now let us consider the algebraic structures of the classes of extremal fuzzy time intervals $\langle \mathscr{F}I^*(T), \mathscr{F}I_*(T) \rangle$.

First we will consider $\widetilde{\mathscr{F}I}^*(T)$. We introduce a partial ordering in $\widetilde{\mathscr{F}I}^*(T)$: If $\widetilde{r}_1, \widetilde{r}_2 \in \widetilde{\mathscr{F}I}^*(T)$, then

$$\widetilde{r}_1 \preceq \widetilde{r}_2 \Leftrightarrow \forall \tau \in T \;\; \mu_{\widetilde{r}_1}(\tau) \leq \mu_{\widetilde{r}_2}(\tau). \tag{21}$$

On the semilattice $\{\widetilde{\mathscr{F}I}^*(T), \preceq\}$ we introduce the algebraic sum operation $\widetilde{r}_1 \overset{*}{\oplus} \widetilde{r}_2$ [28]:

$$\mu_{\widetilde{r}_1 \overset{*}{\oplus} \widetilde{r}_2}(\tau) \overset{\Delta}{=} \wedge \left\{ \mu_{\widetilde{r}_1}(\tau_1) \vee \mu_{\widetilde{r}_2}(\tau_2) \mid \tau_1, \tau_2 \in T, \;\; \tau_1 + \tau_2 = \tau \right\}. \tag{22}$$

It is not difficult to verify that the structure $\{\widetilde{\mathscr{F}I}^*(T), \preceq, \overset{*}{\oplus}\}$ is a partially ordered commutative semigroup.

Let us construct, in $\widetilde{\mathscr{F}I}^*(T)$, a monotonically increasing recurrent sequence of fuzzy time intervals

$$\widetilde{r}_n = \widetilde{r}_{n-1} \overset{*}{\oplus} \Delta\widetilde{r}, \quad n \geq 1, \tag{23}$$

where $\widetilde{r}_0$, $\Delta\widetilde{r} \in \mathscr{F}I^*(T)$ are respectively the initial and the stepwise fuzzy time interval. (Here $\widetilde{r}_0 \equiv \widetilde{\varnothing}$). We obtain

$$\widetilde{r}_1 \preceq \widetilde{r}_2 \preceq \cdots.$$

The partial ordering $\preceq$ in $\widetilde{\mathscr{F}I}^*(T)$ induces in $\widetilde{\mathscr{F}}I_*(T)$ another partial ordering $\succeq$ (conjugate to $\preceq$).

If $\widetilde{r}_1, \widetilde{r}_2 \in \widetilde{\mathscr{F}}I_*(T)$, then

$$\widetilde{r}_1 \succeq \widetilde{r}_2 \Leftrightarrow \overline{\overline{\widetilde{r}}}_1 \preceq \overline{\overline{\widetilde{r}}}_2 \Leftrightarrow \forall\tau \in T \,:\, \mu_{\widetilde{r}_1}(\tau) \geq \mu_{\widetilde{r}_2}(\tau). \tag{24}$$

The algebraic sum operation $\overset{*}{\oplus}$ in $\widetilde{\mathscr{F}I}^*(T)$ induces in $\widetilde{\mathscr{F}}I_*(T)$ another operation (conjugate to $\overset{*}{\oplus}$) $\underset{*}{\oplus}$:

$$\forall\widetilde{r}_1, \widetilde{r}_2 \in \widetilde{\mathscr{F}}I_*(T) \,:\, \widetilde{r}_1 \underset{*}{\oplus} \widetilde{r}_2 = \overline{\overline{\widetilde{r}_1}} \overset{*}{\oplus} \overline{\overline{\widetilde{r}}}_2 \tag{25}$$

or, $\forall\tau \in T$,

$$\mu_{\widetilde{r}_1 \underset{*}{\oplus} \widetilde{r}_2}(\tau) = 1 - \mu_{\underline{\underline{*}}}(\tau) = \vee\left\{\mu_{\widetilde{r}_1}(\tau_1) \wedge \mu_{\widetilde{r}_2}(\tau_2) \mid \tau_1, \tau_2 \in T,\ \tau_1 + \tau_2 = \tau\right\} \tag{26}$$

Then the monotonically increasing sequence of current fuzzy intervals from the class $\widetilde{\mathscr{F}I}^*(T)$

$$\widetilde{r}_1 \preceq \widetilde{r}_2 \preceq \cdots$$

induces, in $\widetilde{\mathscr{F}}I_*(T)$, a monotonically decreasing sequence of future fuzzy intervals

$$\overline{\overline{\widetilde{r}}}_1 \succeq \overline{\overline{\widetilde{r}}}_2 \succeq \cdots$$

defined recurrently as

$$\overline{\overline{\widetilde{r}}}_n = \overline{\overline{\widetilde{r}}}_{n-1} \overset{*}{\oplus} \overline{\widetilde{\Delta r}},$$

where $\overline{\overline{\widetilde{r}}}_0 = 1_T$ and $\overline{\widetilde{\Delta r}} \in \widetilde{\mathscr{F}}I_*(T)$ are respectively the initial fuzzy interval and the stepwise fuzzy time interval.

On $\widetilde{\mathscr{F}}I_*(T)$, the induced structure $\{\mathscr{F}I_*(T), \succeq, \underset{*}{\oplus}\}$ is a partially ordered commutative semigroup.

We call the pair of structures

$$\langle\{\widetilde{\mathscr{F}I}^*(T), \preceq, \overset{*}{\oplus}\}, \{\widetilde{\mathscr{F}}I_*(T), \succeq, \underset{*}{\oplus}\}\rangle \tag{27}$$

an extremal partially ordered commutative semigroup.

To conclude the subsection, we would like to note that

1) the extremal structure (27) of current and future fuzzy time intervals is the subject which will be used in next sections.

2) In the time flow process information (data) obtained by expert measurements is incomplete. The polar characteristics of such information are imprecision and uncertainty. The imprecision degree of the obtained information defines extremal fuzzy time moments, while the uncertainty degree defines algebraic structures represented in form (27).

2.4 Examples of Construction of Extremal Fuzzy Time Intervals

Example 1. Consider the extremal measurable Borel space of second kind $(T, \mathscr{B}_{2*}, \mathscr{B}_2^*)$. Let $f : T \to T$ be some monotonically nondecreasing, left continuous function such that $f(0) = 0$, $f(+\infty) = +\infty$. It is not difficult to verify that $\forall \tau \geq 0$

$$g_T^*([0;t)) \stackrel{\Delta}{=} \frac{f(t)}{1+f(t)} \tag{28}$$

is the upper fuzzy measure on $\mathscr{B}_2^*$, and its extremal fuzzy measure on $\mathscr{B}_{2*}$ is the lower fuzzy measure

$$g_{T*}([t;+\infty)) = \frac{1}{1+f(t)}. \tag{29}$$

Now, for the current fuzzy time interval we consider the extension $\widetilde{g}_T^* \; \forall \widetilde{r} \in \widetilde{\mathscr{F}I}^*(T) \subset \widetilde{\mathscr{B}}_2^*$:

$$\widetilde{g}_T^*(\widetilde{r}) = \int_T^* \mu_{\widetilde{r}}(t) \circ g_T^*(\cdot) = \bigwedge_{0\leq\alpha\leq 1} [\alpha \vee g_T^*([\widetilde{r}]_\alpha)]$$

$$= \bigwedge_{0\leq\alpha\leq 1} [\alpha \vee g_T^*([0;t_\alpha))] = \bigwedge_{0\leq\alpha\leq 1} \left[\alpha \vee \frac{f(t_\alpha)}{1+f(t_\alpha)}\right],$$

where

$$t_\alpha = \vee \left\{ t \geq 0 \mid \mu_{\widetilde{r}}(t) \leq \alpha \leq \mu_{\widetilde{r}}(t^+) \right\},$$

and calculate the extension $\widetilde{g}_{T*} \; \forall \widetilde{r} \in \widetilde{\mathscr{F}I}_*(T) \subset \widetilde{\mathscr{B}}_{2*}$ as follows:

$$\widetilde{g}_{T*}(\widetilde{r}) = \int_{T*} \mu_{\widetilde{r}}(t) \circ g_{T*}(\cdot) = \bigvee_{0\leq\alpha\leq 1} [\alpha \wedge g_{T*}([t_\alpha;+\infty))] = \bigvee_{0\leq\alpha\leq 1} \left[\alpha \wedge \frac{1}{1+f(t_\alpha)}\right],$$

where

$$t_\alpha = \wedge \{ t \geq 0 \mid \mu_{\widetilde{r}}(t-) \leq \alpha \leq \mu_{\widetilde{r}}(t) \}.$$

Thus we have constructed the space of extended extremal fuzzy measures $(T, \widetilde{\mathscr{B}}_{2*}, \widetilde{\mathscr{B}}_2^*, \widetilde{g}_{T*}, \widetilde{g}_T^*)$.

Now, let us consider the problem of construction of extremal fuzzy time intervals. If $\widetilde{r} \in \widetilde{\mathscr{F}I}^{*}(T)$, then, by virtue of formula (20), g_T^* is assumed to be a fuzzy measure on $\mathscr{B}_1^*$, while the fuzzy interval $(\widetilde{r} \in \widetilde{\mathscr{B}}_1^*)$ is assumed to be known:

$$\mu_{\widetilde{r}}(t) = \overset{*}{\underset{[0;t]}{\int\!\!\!\!\!\!\diagup}} \mu_{\widetilde{i}}(s) \circ g_T^*, \quad \forall t \geq 0.$$

Then

$$\mu_{\widetilde{r}}(t) = \overset{*}{\underset{T}{\int\!\!\!\!\!\!\diagup}} I_{(t;+\infty)}(s) \vee \mu_{\widetilde{i}}(s) \circ g_T^*(\cdot)$$
$$= \underset{0\leq\alpha\leq 1}{\wedge} [\alpha \vee g_T^*((t;+\infty) \cup (t_\alpha;+\infty))] = \underset{0\leq\alpha\leq 1}{\wedge} [\alpha \vee g_T^*(s_{t,\alpha};+\infty)],$$

where

$$s_{t,\alpha} = t \wedge t_\alpha, \quad t_\alpha = \wedge \left\{ t \geq 0 \mid \mu_{\widetilde{i}}(t) \leq \alpha \leq \mu_{\widetilde{i}}(t^+) \right\}.$$

If in the role of g_T^* we take $\forall t \geq 0$

$$g^*(t;+\infty)) = \frac{1}{1+f(t)},$$

where $f(t)$ is a monotonically nondecreasing, left continuous function $f : T \to T$, $f(0) = 0$, $f(+\infty) = +\infty$, then

$$\mu_{\widetilde{r}}(t) = \underset{0<\alpha\leq 1}{\wedge} \left[\alpha \vee \frac{1}{1+f(s_{t,\alpha})} \right].$$

If $\widetilde{r} \in \widetilde{\mathscr{F}I}_*(T)$, then, analogously, we construct

$$\mu_{\widetilde{r}}(t) = \overset{*}{\underset{(t;+\infty)}{\int\!\!\!\!\!\!\diagup}} \mu_{\widetilde{i}}(s) \circ g_{T*}(\cdot).$$

In that case

$$\mu_{\widetilde{r}}(t) = \underset{0<\alpha\leq 1}{\vee} [\alpha \wedge g_{T*}([0;s_{t,\alpha}])],$$

where

$$s_{t,\alpha} = t \wedge t_\alpha, \quad t_\alpha = \vee \left\{ t \geq 0 \mid \mu_{\widetilde{i}}(t) \leq \alpha \leq \mu_{\widetilde{i}}(t^+) \right\}$$

or

$$\mu_{\widetilde{r}}(t) = 1 - \mu_{\overline{\widetilde{r}}}(t) = \underset{0\leq\alpha\leq 1}{\vee} \left[\alpha \wedge \frac{f(s_{t,\alpha})}{1+f(s_{t,\alpha})} \right].$$

Example 2. Let g_T^* be an upper possibilistic measure on $\mathscr{B}_2^*$, i.e., $\exists f^* : T \to [0;1]$ is a left continuous, monotonically nondecreasing function such that $f(0) = 0$, $f(+\infty) = 1$, and $\forall [0;t) \in \mathscr{B}_2^*$

$$g_T^*([0;t)) = \underset{0<s<t}{\vee} f(s) = f(t).$$

Then $\forall \widetilde{r} \in \widetilde{\mathscr{F}I}^{*}(T)$

$$\mu_{\widetilde{r}}(t) = \underset{0<\alpha\le 1}{\wedge} [\alpha \vee f^{*}(s_{t,\alpha})],$$

where

$$s_{t,\alpha} = t \wedge t_{\alpha}, \quad t_{\alpha} = \wedge \{t \ge 0 \mid \mu_{\widetilde{r}}(t) \le \alpha \le \mu_{t}(t^{+})\}.$$

Example 3. Let g_T^* be an upper λ-fuzzy measure [44] on $\mathscr{B}_2^*$ ($g_T^* \equiv g_\lambda^*$, $-1 \le \lambda \le 0$), i.e., $\forall [0;t) \in \mathscr{B}_2^*$

$$g_\lambda^*([0;t)) = \frac{1 - f^*(t)}{1 + \lambda f^*(t)},$$

where f^* is a distribution function of the measure g_λ^*, $f^* : T \to [0;1]$ is a left continuous, monotonically nondecreasing function, $f^*(0) = 0$, $f^*(+\infty) = 1$. Then

$$g_{T*}([t;+\infty)) = g_{\lambda *}([t;+\infty)) = \frac{f^*(t)(1-\lambda)}{1 + \lambda f^*(t)}$$

and $\forall \widetilde{r} \in \widetilde{\mathscr{F}I}_{*}(T)$

$$\mu_{\widetilde{r}}(t) = \underset{0\le\alpha\le 1}{\vee} \left[\alpha \wedge \frac{f^*(s_{t,\alpha})(1-\lambda)}{1 + \lambda f^*(s_{t,\alpha})}\right],$$

while $\forall \widetilde{r} \in \widetilde{\mathscr{F}I}^{*}(T)$

$$\mu_{\widetilde{r}}(t) = \underset{0\le\alpha\le 1}{\wedge} \left[\alpha \vee \frac{1 - f^*(s_{t,\alpha})}{1 + \lambda f^*(s_{t,\alpha})}\right],$$

where $s_{t,\alpha}$ is defined from Example 2.

Example 4. It is natural to introduce a fuzzy time interval $\widetilde{r} \in \mathscr{B}_2^*$ such that the kernel of $\widetilde{r}$ would coincide with the interval $[0;\tau]$.

Let us define the upper fuzzy time interval as follows. For $\forall \tau \ge 0$, $\widetilde{r}_\tau \in \mathscr{B}_2^*$:

$$\mu_{\widetilde{r}_\tau}(t) = \begin{cases} 1, & 0 \le t \le \tau, \\ g_{T*}([t;+\infty)) \vee g_T^*([0;\tau)), & t \ge \tau. \end{cases}$$

If $g_T^* : \mathscr{B}_1^* \to [0;1]$:

$$g_T^*([0;t)) = \frac{f(t)}{1 + f(t)},$$

as in Example 1, then

$$\mu_{\widetilde{r}_\tau}(t) = \begin{cases} 1, & 0 \le t \le \tau, \\ \left(\frac{f(t)}{1+f(t)}\right) \wedge \left(\frac{f(\tau)}{1+f(\tau)}\right), & t > \tau. \end{cases}$$

If $0 \le \alpha \le 1$, then the solution of the equation

$$\alpha = \frac{1}{1+f(t)} \vee \frac{f(\tau)}{1+f(\tau)}$$

with respect to t is denoted by $t_{\tau,\alpha}$. If $t_{\tau,1} = \tau$, $t_{\tau,0} = \infty$, then

$$g_T^*(\widetilde{r}_\tau) = \int\limits_T^* \mu_{\widetilde{r}_\tau}(t) \circ g_T^*(\cdot) = \bigwedge_{0<\alpha\le 1} \left[\alpha \vee \frac{t_{\tau,\alpha}}{1+t_{\tau,\alpha}}\right].$$

Note that if $\tau_1 < \tau_2$, then $\widetilde{r}_{\tau_1} \succeq \widetilde{r}_{\tau_2}$.

This example makes it possible to construct parametrically some sequence of extremal fuzzy intervals.

3 Description of a General Model of an Extremal Fuzzy Continuous Dynamic System (EFCDS)

Following the system approach of modeling complex systems [20] we propose the following: the time structure of fuzzy dynamic systems is represented by some space of extended extremal fuzzy measures

$$\langle T, \widetilde{\mathscr{F}} I_*(T), \widetilde{\mathscr{F}} I^*(T), \widetilde{g}_{T*}, \widetilde{g}_T^* \rangle, \quad T = \mathbb{R}_0^*, \tag{30}$$

and structure (1), where $\widetilde{g}_{T*}$ and $\widetilde{g}_T^*$ are some extremal fuzzy measures on $\widetilde{\mathscr{B}}_{T*} \equiv \widetilde{\mathscr{B}}_{2*}$ and $\widetilde{\mathscr{B}}_T^* \equiv \widetilde{\mathscr{B}}_2^*$, respectively (see Subsection 2.1).

Let us start describing objects of a fuzzy dynamic system. Let X $(X \ne \varnothing)$ be the set of states of some system to be investigated. Let $(X, \mathscr{B}, g)$ be the space of a fuzzy measure on the measurable space $(X, \mathscr{B})$, where $\mathscr{B}$ is a σ-algebra in X.

Let U $(U \ne \varnothing)$ be the set of all admissible controls (of external factors) acting on the system. Assume that controls are subjected to restrictions of uncertain character in the form of some space of a fuzzy measure $(U, \mathscr{B}_U, g_U)$, where $\mathscr{B}_U$ is the measurable space of controls, while the fuzzy measure g_U describes the restrictions imposed on controls.

Let Y $(Y \ne \varnothing)$ be the set of output states of the system under consideration, and $(Y, \mathscr{B}_Y, g_Y)$ be the space of a fuzzy measure, which describes a fuzzy distribution of output values of the system. Note that as usual Y is some transformation of the set of states of X.

Now let us consider the Cartesian product $X \times T$ and the space of extended composition extremal fuzzy measures (Subsection 2.2 and [32])

$$\left(X \times T, \widetilde{\mathscr{B} \otimes \mathscr{B}_{T*}}, \widetilde{\mathscr{B} \otimes \mathscr{B}_T^*}, \widetilde{g \otimes g_{T*}}, \widetilde{g \otimes g_T^*}\right),$$

which is induced by the spaces $(X, \mathscr{B}, \mathscr{B}, g, g)$ and $(T, \mathscr{B}_{T*}, \mathscr{B}_T^*, g_{T*}, g_T^*)$.

Definition 20. a) A lower measurable binary fuzzy relation $\widetilde{Q}_* \in \widetilde{\mathscr{B} \otimes \mathscr{B}_{T*}}$ is called a future fuzzy process on the measurable states of the system (i.e., $\mu_{\widetilde{Q}_*}(x,t)$ is a lower measurable function).

b) An upper measurable binary fuzzy relation $\widetilde{Q}^* \in \widetilde{\mathscr{B} \otimes \mathscr{B}^*_T}$ is called a current fuzzy process on the measurable states of the system (i.e., $\mu_{\widetilde{Q}^*}(x,t)$ is an upper measurable function).

c) A pair $(\widetilde{Q}_*, \widetilde{Q}^*)$ of lower and upper measurable binary fuzzy relations is called an extremal fuzzy process on the measurable states of the system (i.e., $Q^* \in \widetilde{\mathscr{B} \otimes \mathscr{B}^*_T}$ and $Q_* \in \widetilde{\mathscr{B} \otimes \mathscr{B}_{T*}}$).

d) An extremal fuzzy process (EFP) is said to be ergodic if there exist the limits $\forall x \in X$, $\lim\limits_{t\to\infty} \mu_{\widetilde{Q}^*}(x,t) \equiv \mu_{\widetilde{A}^*}(x)$, $\lim\limits_{t\to\infty} \mu_{\widetilde{Q}_*}(x,t) \equiv \mu_{\widetilde{A}_*}(x)$, and the limit fuzzy sets $\widetilde{A}^*$ and $\widetilde{A}_*$ are measurable $\widetilde{A}^*, \widetilde{A}_* \in \widetilde{\mathscr{B}}$.

Note that (see Subsection 2.2) $\forall \tau \in T$, $\forall x \in X$

$E_{\widetilde{Q}_*}(x,\cdot) \in \widetilde{\mathscr{B}}_{T*}$ is a future fuzzy time interval,

$E_{\widetilde{Q}^*}(x,\cdot) \in \widetilde{\mathscr{B}}^*_T$ is a current fuzzy time interval,

$E_{\widetilde{Q}_*}(\cdot,\tau) \in \widetilde{\mathscr{B}}$ is a fuzzy state of the system, which is "measurable" in the future fuzzy time interval $\widetilde{[\tau, +\infty)}$,

$E_{\widetilde{Q}^*}(\cdot,\tau) \in \widetilde{\mathscr{B}}$ is a fuzzy state of the system, which is "measurable" in the current fuzzy time interval $\widetilde{[0,\tau)}$.

It is obvious that model "measurements" of the states of the system at a real time moment $\tau > 0$ are understood as defining pairs of measurable fuzzy sets $E_{\widetilde{Q}_*}(\cdot,\tau)$, $E_{\widetilde{Q}^*}(\cdot,\tau) \in \widetilde{\mathscr{B}}$.

For all $x \in X$, $E_{\widetilde{Q}^*}(x,\cdot)$ and $E_{\widetilde{Q}_*}(x,\cdot)$ are a current fuzzy and a future fuzzy time intervals, respectively, in which the state $x \in X$ of the system is measured.

The family of fuzzy sets $\{E_{\widetilde{Q}_*}(\cdot,\tau)\}_{\tau\geq 0}$ from $\widetilde{\mathscr{B}}$ is called the trajectory of a future fuzzy process, and the family of fuzzy sets $\{E_{\widetilde{Q}^*}(\cdot,\tau)\}_{\tau\geq 0}$ from $\widetilde{\mathscr{B}}$ is called the trajectory of a current fuzzy process. The family of pairs of fuzzy sets $\{E_{\widetilde{Q}_*}(\cdot,\tau), E_{\widetilde{Q}^*}(\cdot,\tau)\}_{\tau\geq 0}$ is called the trajectory of an extremal fuzzy process $(\widetilde{Q}_*, \widetilde{Q}^*)$.

Let $\widetilde{\mathbb{R}}_* \subset X \times T \times Y$ be some lower measurable fuzzy relation $(\widetilde{\mathbb{R}}_* \in \widetilde{\mathscr{B} \otimes \mathscr{B}_{T*}} \otimes \widetilde{\mathscr{B}_Y})$ describing expert knowledge reflections of fuzzy states of the system on the output values of the system in future fuzzy time intervals, and $\widetilde{\mathbb{R}}^* \subset X \times T \times Y$ be some upper measurable fuzzy relation $(\widetilde{\mathbb{R}}^* \in \widetilde{\mathscr{B} \otimes \mathscr{B}^*_T} \otimes \widetilde{\mathscr{B}_Y})$ describing expert knowledge reflections of fuzzy states of the system on the output values of the system in current fuzzy time intervals.

Definition 21. a) A lower measurable relation $\widetilde{\mathbb{R}}_* \in \widetilde{\mathscr{B} \otimes \mathscr{B}_{T*}} \otimes \widetilde{\mathscr{B}_Y}$ is called a future fuzzy process of expert knowledge reflection of states of the system in future fuzzy time intervals.

b) An upper measurable relation $(\widetilde{\mathbb{R}}^* \in \widetilde{\mathscr{B} \otimes \mathscr{B}_T^*} \otimes \widetilde{\mathscr{B}_Y})$ is called a current fuzzy process of expert knowledge reflection of states of the system in current fuzzy time intervals.

c) A pair $(\widetilde{\mathbb{R}}_*, \widetilde{\mathbb{R}}^*)$ is called an extremal fuzzy process of expert knowledge reflection of states of the system in extremal fuzzy time intervals.

Let $\widetilde{\rho}_* \in (\widetilde{\mathscr{B} \otimes \mathscr{B}_{T*}}) \otimes (\widetilde{\mathscr{B}_U \otimes \mathscr{B}_{T*}}) \otimes \widetilde{\mathscr{B}}$ be some lower measurable fuzzy relation in the Cartesian product $(X \times T) \times (U \times T) \times X$, which describes system state transformations in time with control taken into account:

$$(X \times T) \times (U \times T) \to X.$$

This relation is a future fuzzy transition operator describing the dynamics of the system or, in other words, system state transformations in future fuzzy time intervals.

Let $\widetilde{\rho}^* \in (\widetilde{\mathscr{B} \otimes \mathscr{B}_T^*}) \otimes (\widetilde{\mathscr{B}_U \otimes \mathscr{B}_{T*}}) \otimes \widetilde{\mathscr{B}}$ be some upper measurable fuzzy relation in the Cartesian product $(X \times T) \times (U \times T) \times X$, which describes system state transformations in time with control taken into account:

$$(X \times T) \times (U \times T) \to X.$$

This relation is a current fuzzy transition operator describing the dynamics of the system or, in other words, system state transformations in current fuzzy time intervals.

We call $\widetilde{\rho}_*$ the fuzzy lower transition operator describing the system state dynamics, and $\widetilde{\rho}^*$ the fuzzy upper transition operator describing the system state dynamics. The pair $(\widetilde{\rho}_*, \widetilde{\rho}^*)$ is called the transition operator describing the system state dynamics in extremal fuzzy time intervals.

Let $\widetilde{u}^* \subset U \times T$ be some upper measurable fuzzy binary relation from $\widetilde{\mathscr{B}_U \otimes \mathscr{B}_T^*}$, which describes the action of external factors (controls) on the system in future fuzzy time intervals, and $\widetilde{u}_* \subset U \times T$ be some lower measurable fuzzy binary relation from $\widetilde{\mathscr{B}_U \otimes \mathscr{B}_T^*}$, which describes the action of external factors (controls) on the system in current fuzzy time intervals.

Definition 22. a) A fuzzy binary relation $\widetilde{u}^* \in \widetilde{\mathscr{B} \otimes \mathscr{B}_T^*}$ is called a current fuzzy control process.

b) A fuzzy binary relation $\widetilde{u}_* \in \widetilde{\mathscr{B} \otimes \mathscr{B}_{T*}}$ is called a future fuzzy control process.

c) A pair $(\widetilde{u}_*, \widetilde{u}^*)$ is called an extremal fuzzy control process.

Definition 23. a) The train

$$\left\{X, U, T, Y, \widetilde{\rho}_*, \widetilde{Q}_*, \widetilde{\mathbb{R}}_*\right\} \tag{31}$$

is called the future fuzzy dynamic system describing the dynamics of the system state in future fuzzy time intervals.

b) The train

$$\left\{X,U,T,Y,\widetilde{\rho}^*,\widetilde{Q}^*,\widetilde{\mathbb{R}}^*\right\} \tag{32}$$

is called the current fuzzy dynamic system describing the state dynamics of the system in current fuzzy time intervals.

c) The train

$$\left\{X,U,T,Y,(\widetilde{\rho}_*,\widetilde{\rho}^*),(\widetilde{Q}_*,\widetilde{Q}^*),(\widetilde{\mathbb{R}}_*,\widetilde{\mathbb{R}}^*)\right\} \tag{33}$$

is called the extremal fuzzy consitnuous dynamic system (EFCDS) describing the state dynamics of the system in extremal fuzzy time intervals.

In the sequel we will consider the case with $Y \equiv X$.

It is obvious that the EFCDS (33) describes the state dynamics of the system undergoing transformation with fuzzy uncertainty produced by observations at fuzzy time, while the extremality is due to the "measurement" of fuzzy states of the system in current and future fuzzy time intervals.

Definition 24. The system of composition equations

$$\begin{cases}\widetilde{\mathbb{R}}_* = \widetilde{\rho}_* \underset{*}{\bullet} \widetilde{Q}_*, \\ \widetilde{\mathbb{R}}^* = \widetilde{\rho}^* \overset{*}{\bullet} \widetilde{Q}^*\end{cases} \tag{34}$$

is called the system describing the state dynamics of the extremal fuzzy continuous dynamic system, where $\underset{*}{\bullet}$ and $\overset{*}{\bullet}$ are some composition operations over fuzzy relations.

Given $(\widetilde{\mathbb{R}}_*,\widetilde{\mathbb{R}}^*)$, $(\widetilde{\rho}_*,\widetilde{\rho}^*)$ and the initial fuzzy states of the system $\widetilde{A}_{0*},\widetilde{A}_0^* \in \mathscr{B}$ $(\mu_{\widetilde{A}_{0*}}(x) = \mu_{\widetilde{Q}_*}(x,0)$, $\mu_{\widetilde{A}_0^*}(x) = \mu_{\widetilde{Q}^*}(x,0)$, $\forall x \in X)$, it is important to find a solution $(\widetilde{Q}_*,\widetilde{Q}^*)$ of (34), which we call an extremal fuzzy process of system state transformation on measurable states of the system.

Below we will consider a concrete controllable fuzzy system of form (34) for the continuous case. It is obvious that in concrete EFCDS's formulas (33) and (34) model concrete complex objects with fuzzy dynamics. The finding of a system state transformation process $(\widetilde{Q}_*,\widetilde{Q}^*)$ is important when we deal with problems pertaining to optimization problem (optimal control).

In recent years, the investigation of complex dynamic systems with fuzzy uncertainty by means of the theory of fuzzy sets has been developing mainly along the following two lines:

I. Lower dynamic systems are described by composition equations in the metric or normed spaces of system states, which can be formally written in terms of fuzzy integral equations if a fuzzy measure is assumed to be a possibilistic one ([9], [10], [24], [46], [47] and so on).

II. Quite a number of studies have been devoted to the development of fuzzy integro-differential calculus with an aim of describing fuzzy dynamic systems and their control. The main feature these approaches have in common is the assumption

that the compatibility function is differentiable or integrable ([4]–[7], [10], [12], [17], [19], [23], [26], [27], [29], [30], [43] and so on), which to a certain extent facilitates the investigation of the definite class of fuzzy dynamic systems.

The instrument of extended composition fuzzy measures developed in [31] and [32], where some important properties of Sugeno lower and upper integrals and their extensions are investigated, makes it possible to study the so-called extremal fuzzy continuous dynamic systems for which:

1) a system of compositional equations for fuzzy dynamic systems is generalized in the form of system (34), where the extended Sugeno upper and lower integrals (see [31]) are used in the role of composition operations $\bullet$ and $\overset{*}{\bullet}$ (as a aggregation instrument for the EFCDS) describing the dynamics of the state of an EFCDS [33].

As known from the earlier sources of investigation of fuzzy statistics ([8], [13], [18], [22], [45] and so on) and also from our works ([31]–[36], [40]–[42]), the Sugeno integral most frequently estimates the most typical levels of compatibility of an integrable function. This is the reason for which we have chosen the Sugeno integral for the construction of extended fuzzy measures.

Systems of composition type equations [24] are a particular case of system (34), where equations are written with respect to possibility measure. The case we consider in this paper is more general since the equations are written in for any extremal fuzzy measure.

2) As different from the approach mentioned in Item II (where some processes are not integro-differentiable), in our proposed systems of equations any measurable compatibility function is integrable. However our consideration is not limited to this only class of dynamic systems.

To conclude the section, note that the compatibility functions, for which systems of equations can be written in form (34), are lower or upper measurable:

$$
\begin{aligned}
&\mu_{\widetilde{Q}^*}(x,t) : X \times T \to [0;1] \text{ is } \mathscr{B} \otimes \mathscr{B}_T^*\text{-upper measurable;}\\
&\mu_{\widetilde{Q}_*}(x,t) : X \times T \to [0;1] \text{ is } \mathscr{B} \otimes \mathscr{B}_{T*}\text{-lower measurable;}\\
&\mu_{\widetilde{\mathbb{R}}^*}(x,t,y) : X \times T \times Y \to [0;1] \text{ is } \mathscr{B} \otimes \mathscr{B}_T^* \otimes \mathscr{B}\text{-upper measurable;}\\
&\mu_{\widetilde{\mathbb{R}}_*}(x,t,y) : X \times T \times Y \to [0;1] \text{ is} \mathscr{B} \otimes \mathscr{B}_{T*} \otimes \mathscr{B}\text{-lower measurable;}\\
&\mu_{\widetilde{\rho}_*}(x_0,t_0,u,t,x) : (X \times T) \times (U \times T) \times X \to [0;1] \text{ is}\\
&\qquad (\mathscr{B} \otimes \mathscr{B}_T^*) \otimes (\mathscr{B}_U \otimes \mathscr{B}_T^*) \otimes \mathscr{B}\text{-upper measurable;}\\
&\mu_{\widetilde{\rho}^*}(x_0,t_0,u,t,x) : (X \times T) \times (U \times T) \times X \to [0;1] \text{ is}\\
&\qquad (\mathscr{B} \otimes \mathscr{B}_{T*}) \otimes (\mathscr{B}_U \otimes \mathscr{B}_{T*}) \otimes \mathscr{B}\text{-lower measurable.}
\end{aligned}
\tag{35}
$$

4 Continuous Extremal Controllable Fuzzy Process

As has been mentioned above (Subsection 2.2), in [31] we have constructed monotone structures of current fuzzy time intervals $\{\widetilde{\mathscr{F}I}^*(T), \preceq, \overset{*}{\otimes}\}$ and future fuzzy time intervals $\{\widetilde{\mathscr{F}I}_*(T), \succeq, \underset{*}{\otimes}\}$. It is obvious that the flow process of a real time

moment τ induces, in these structures, monotonically increasing and monotonically decreasing processes of current and future time intervals, respectively.

Definition 25. a) A family $\{\widetilde{r}^*_\tau\}_{\tau\geq 0}$, $\widetilde{r}^*_\tau \in \widetilde{\mathscr{B}}^*_T$, $\tau \geq 0$, of monotonically increasing sequences of upper fuzzy time intervals, i.e.,

$$\forall \tau_2 > \tau_1 \geq 0, \quad \widetilde{r}^*_{\tau_1} \preceq \widetilde{r}^*_{\tau_2}$$

is called a process of current fuzzy time intervals.

b) A family $\{\widetilde{r}_{\tau *}\}_{\tau\geq 0}$, $\widetilde{r}_{\tau *} \in \widetilde{\mathscr{B}}_{T*}$, $\tau \geq 0$, of monotonically decreasing sequences of upper fuzzy time intervals, i.e.,

$$\forall \tau_2 > \tau_1 \geq 0, \quad \widetilde{r}_{\tau_1 *} \succeq \widetilde{r}_{\tau_2 *}$$

is called a process of future fuzzy time intervals.

c) A pair of processes of future and current fuzzy time intervals $\{\widetilde{r}_{\tau *}, \widetilde{r}^*_\tau\}_{\tau\geq 0}$ is called a process of extremal fuzzy time intervals.

It obviously follows that
$\mu_{\widetilde{r}_{\tau *}}(t) : T \to [0;1]$ is $\mathscr{B}_{T*}$-lower measurable,
$\mu_{\widetilde{r}^*_\tau}(t) : T \to [0;1]$ is $\mathscr{B}^*_T$-upper measurable.

Note that a change of a real time moment $\tau > 0$ reflects model "measurements" of an extremal fuzzy process of system state transformation $(\widetilde{Q}_*, \widetilde{Q}^*)$ in extremal fuzzy time intervals $(\widetilde{r}_{\tau *}, \widetilde{r}^*_\tau)$.

Definition 26. A process of extremal fuzzy time intervals $(\widetilde{r}_{\tau *}, \widetilde{r}^*_\tau)$ is called ergodic if there exist the limits

$$\lim_{\tau\to+\infty} \widetilde{r}_{\tau *} = \widetilde{r}_{\infty *} \in \widetilde{\mathscr{B}}_{T*} \quad \left(\forall t \geq 0, \ \lim_{\tau\to+\infty} \mu_{\widetilde{r}_{\tau *}}(t) = \mu_{\widetilde{r}_{\infty *}}(t)\right),$$

$$\lim_{\tau\to+\infty} \widetilde{r}^*_\tau = \widetilde{r}^*_\infty \in \widetilde{\mathscr{B}}^*_T \quad \left(\forall t \geq 0, \ \lim_{\tau\to+\infty} \mu_{\widetilde{r}^*_\tau}(t) = \mu_{\widetilde{r}^*_\infty}(t)\right).$$

In what follows it will be assumed that there exists a relation between the measurable space of time $(T, \mathscr{B}_{T*}, \mathscr{B}^*_T)$ and the measurable space of system states $(X, \mathscr{B})$ in the form of conditional extremal fuzzy measures defined in [31]. In the considered case it is assumed that there exist conditional lower and upper fuzzy measures $g_{t*}(\cdot \mid x)$ and $g^*_t(\cdot \mid x)$, respectively, i.e., $\forall x \in X$

$$g_{t*}(\cdot \mid x) : \mathscr{B}_{T*} \to [0;1] \text{ is a lower fuzzy measure,}$$
$$g^*_t(\cdot \mid x) : \mathscr{B}^*_T \to [0;1] \text{ is an upper fuzzy measure.}$$

$g_{t*}(\cdot \mid x)$ and $g^*_t(\cdot \mid x)$ are extremal measures, while for a future fuzzy time interval $r \in \mathscr{B}_{T*}$

$$g_{t*}(r \mid \cdot) : X \to [0;1] \text{ is a } \mathscr{B}\text{-measurable function,}$$

and for a current time interval $r \in \mathscr{B}^*_T$

$$g_t^*(r \mid \cdot) : X \to [0;1] \text{ is a } \mathscr{B}\text{-measurable function.}$$

These properties also apply to extended conditional fuzzy measures $\widetilde{g}_{t*}(\cdot \mid x)$ and $\widetilde{g}_t^*(\cdot \mid x)$, i.e., $\forall x \in X, \widetilde{r}_* \in \widetilde{\mathscr{B}}_{T*}, \widetilde{r}^* \in \widetilde{\mathscr{B}}_T^*$

$$\begin{aligned}
&\widetilde{g}_{t*}(\cdot \mid x) : \widetilde{\mathscr{B}}_{T*} \to [0;1] \text{ is a lower fuzzy measure,}\\
&\widetilde{g}_t^*(\cdot \mid x) : \widetilde{\mathscr{B}}_T^* \to [0;1] \text{ is an upper fuzzy measure,}\\
&\widetilde{g}_{t*}(\widetilde{r}_* \mid \cdot) : X \to [0;1] \text{ is a } \mathscr{B}\text{-measurable function,}\\
&\widetilde{g}_t^*(\widetilde{r}^* \mid \cdot) : X \to [0;1] \text{ is a } \mathscr{B}\text{-measurable function.}
\end{aligned}$$

A relation between the spaces $(X, \mathscr{B}, g)$ and $(T, \mathscr{B}_{T*}, \mathscr{B}_T^*, g_{T*}, g_T^*)$ and their extensions through conditional measures can be represented as follows: $\forall r_* \in \mathscr{B}_{T*}$, $r^* \in \mathscr{B}_T^*, \widetilde{r}_* \in \widetilde{\mathscr{B}}_{T*}, \widetilde{r}^* \in \widetilde{\mathscr{B}}_T^*$

$$\begin{aligned}
g_{T*}(r_*) &= \int_X g_{t*}(r_* \mid x) \circ g(\cdot), \quad g_T^*(r^*) = \int_X g_t^*(r^* \mid x) \circ g(\cdot),\\
\widetilde{g}_{T*}(\widetilde{r}_*) &= \int_X \widetilde{g}_{t*}(\widetilde{r}_* \mid x) \circ g(\cdot), \quad \widetilde{g}_T^*(\widetilde{r}^*) = \int_X \widetilde{g}_t^*(\widetilde{r}^* \mid x) \circ g(\cdot),
\end{aligned} \tag{36}$$

Applying results from [31], we can write $\forall x \in X, \widetilde{r}_* \in \widetilde{\mathscr{B}}_{T*}, \widetilde{r}^* \in \widetilde{\mathscr{B}}_T^*$:

$$\begin{aligned}
\widetilde{g}_{t*}(\widetilde{r}_* \mid x) &= \int_{*T} \mu_{\widetilde{r}_*}(t) \circ g_{t*}(\cdot \mid x),\\
\widetilde{g}_t^*(\widetilde{r}^* \mid x) &= \int_T^* \mu_{\widetilde{r}^*}(t) \circ g_t^*(\cdot \mid x).
\end{aligned} \tag{37}$$

By the definition of $\widetilde{g}_{t*}(\cdot \mid x)$ and $\widetilde{g}_t^*(\cdot \mid x)$, for any lower and upper fuzzy time intervals $\widetilde{r}_* \in \widetilde{\mathscr{B}}_{T*}$ and $\widetilde{r}^* \in \widetilde{\mathscr{B}}_T^*$ there exist $\mathscr{B}$-measurable sets $\widetilde{A}_{\widetilde{r}_*} \in \widetilde{\mathscr{B}}$, $\widetilde{A}_{\widetilde{r}^*} \in \widetilde{\mathscr{B}}$ such that $\forall x \in X$

$$\mu_{\widetilde{A}_{\widetilde{r}_*}}(x) = \widetilde{g}_{t*}(\widetilde{r}_* \mid x), \quad \mu_{\widetilde{A}_{\widetilde{r}^*}}(x) = \widetilde{g}_t^*(\widetilde{r}^* \mid x). \tag{38}$$

Definition 27. The fuzzy sets $\widetilde{A}_{\widetilde{r}_*}$ and $\widetilde{A}_{\widetilde{r}^*} \in \widetilde{\mathscr{B}}$ from the extended measurable space of system states are called the expert reflections of an extremal fuzzy dynamic systems states in the extremal fuzzy time intervals $(\widetilde{r}_*, \widetilde{r}^*)$ with respect to extended extremal conditional fuzzy measures $\widetilde{g}_{t*}(\cdot \mid x)$ and $\widetilde{g}_t^*(\cdot \mid x)$.

Let us formulate a theorem that describes the ergodicity of an expert reflection process in an ergodic process of extremal fuzzy time intervals.

Theorem 4. *An ergodic process* $(\widetilde{r}_{\tau*}, \widetilde{r}_\tau^*)_{\tau \geq 0}$ *of extremal fuzzy time intervals on the measurable space of states of the system* $(X, \widetilde{\mathscr{B}})$ *induces an ergodic expert reflection process* $(\widetilde{\mathbb{R}}_*, \widetilde{\mathbb{R}}^*) \equiv (\widetilde{A}_{\widetilde{r}_{\tau*}}, \widetilde{A}_{\widetilde{r}_\tau^*})_{\tau \geq 0}$.

In this section, we consider problems of modeling EFCDS's when the control factor acts on the system or, speaking more exactly, on controllable extremal fuzzy processes.

As defined in Section 3, let U be the space of all admissible controls acting on an EFCDS in the course of its evolution. It is assumed that the restrictions on the space of control elements are of fuzzy nature: these restrictions exist in the form of a fuzzy measure on the measurable space $\mathscr{B}_U$ (the σ-algebra of subsets of U). Let $(U,\mathscr{B}_U,g_U)$ be some space of the fuzzy measure.

Let $\widetilde{u}_* \subset U\times T$ be some upper measurable binary fuzzy relation from $\widetilde{\mathscr{B}_U\otimes\mathscr{B}^*_T}$ that describes an external fuzzy action on the EFCDS in the course of current fuzzy time intervals, and $\widetilde{u}_* \subset U\times T$ be some lower measurable binary fuzzy relation from $\widetilde{\mathscr{B}_U\otimes\mathscr{B}_{T*}}$ that describes an external fuzzy action on the EFCDS in the course of future fuzzy time intervals. A pair $(\widetilde{u}_*,\widetilde{u}^*)$ is called an extremal fuzzy control (an extremal fuzzy control process), while $\widetilde{u}^*$ and $\widetilde{u}_*$ are respectively called a current fuzzy control and a future fuzzy control.

Let $\widetilde{\rho}_* \in \widetilde{\mathscr{B}\otimes\mathscr{B}}\otimes\widetilde{\mathscr{B}_U\otimes\mathscr{B}_{T*}}$ and $\widetilde{\rho}^* \in \widetilde{\mathscr{B}\otimes\mathscr{B}}\otimes\widetilde{\mathscr{B}_U\otimes\mathscr{B}^*_T}$, and $(\widetilde{\rho}_*,\widetilde{\rho}^*)$ be the operator of the EFCDS state change dynamics.

Definition 28. If $(\widetilde{r}_{\tau*},\widetilde{r}^*_\tau)_{\tau\ge0}$ is some process of extremal fuzzy time intervals, $(U,\mathscr{B}_U,g_U)$ is a space of a fuzzy measure (a space of fuzzy restrictions on controls), then a pair $(\widetilde{Q}'_*,\widetilde{Q}'^*)$ of lower and upper measurable binary fuzzy relations $(\widetilde{Q}'_* \in \widetilde{\mathscr{B}}\otimes\widetilde{\mathscr{B}\otimes\mathscr{B}_{T*}},\widetilde{Q}'^* \in \widetilde{\mathscr{B}}\otimes\widetilde{\mathscr{B}\otimes\mathscr{B}^*_T})$ is called an extremal fuzzy process of measurable states of an EFCDS in the process $(\widetilde{r}_{\tau*},\widetilde{r}^*_\tau)_{\tau\ge0}$, taking into account the fuzzy restrictions on controls $(U,\mathscr{B}_U,g_U)$: $\forall x\in X$, $u\in U$, $\tau\in T$,

$$\begin{aligned}
\mu_{\widetilde{Q}'_*}(x,u,\tau) &\overset{\Delta}{=} \int\limits_{*\,\widetilde{r}_{\tau*}}\left[\int\limits_{A_{0*}}\mu_{\widetilde{\rho}_*}(x,x',u,t)\circ g(\cdot)\right]\circ\widetilde{g}_{T*}(\cdot) \equiv \int\limits_{*\,\widetilde{r}_{\tau*}}\mu_{\widetilde{\rho}'_*}(x,u,t)\circ\widetilde{g}_{T*}(\cdot),\\
\mu_{\widetilde{Q}'^*}(x,u,\tau) &\overset{\Delta}{=} \int\limits^{*}_{\widetilde{r}^*_\tau}\left[\int\limits_{A^*_0}\mu_{\widetilde{\rho}^*}(x,x',u,t)\circ g(\cdot)\right]\circ\widetilde{g}^*_T(\cdot) \equiv \int\limits^{*}_{\widetilde{r}^*_\tau}\mu_{\widetilde{\rho}'^*}(x,u,t)\circ\widetilde{g}_{T*}(\cdot).
\end{aligned} \tag{39}$$

Definition 29. In the conditions of the action of an extremal fuzzy control process $(\widetilde{u}_*,\widetilde{u}^*)$ on an EFCDS, a pair $(\widetilde{Q}_*,\widetilde{Q}^*)$ $(\widetilde{Q}_* \in \widetilde{\mathscr{B}\otimes\mathscr{B}_{T*}}$, $\widetilde{Q}^* \in \widetilde{\mathscr{B}\otimes\mathscr{B}^*_T})$ defined as follows: $\forall(x,\tau)\in X\times T$

$$\begin{aligned}
\mu_{\widetilde{Q}_*}(x,\tau) &\overset{\Delta}{=} \int\limits_{E_{\widetilde{u}_*}(\cdot,\tau)}\mu_{\widetilde{Q}'_*}(x,u,\tau)\circ g_U(\cdot),\\
\mu_{\widetilde{Q}^*}(x,\tau) &\overset{\Delta}{=} \int\limits_{E_{\widetilde{u}^*}(\cdot,\tau)}\mu_{\widetilde{Q}'^*}(x,u,\tau)\circ g_U(\cdot)
\end{aligned} \tag{40}$$

is called an extremal fuzzy process describing the system state dynamics.

Let us present the integral representation of the process $(\widetilde{Q}_*,\widetilde{Q}^*)$ [33].

Theorem 5. *In the conditions of the action of an extremal fuzzy control process* $(\widetilde{u}_*,\widetilde{u}^*)$ *on an EFCDS with the initial extremal fuzzy state* $\langle \widetilde{A}_{0*} \equiv E_{\widetilde{Q}_*}(\cdot,\tau_0), \widetilde{A}_0^* \equiv E_{\widetilde{Q}^*}(\cdot,\tau_0)\rangle$, *the system state change dynamics is described by the extremal fuzzy process* $(\widetilde{Q}_*,\widetilde{Q}^*)$, *the integral representation of which is as follows:* $\forall x \in X,\ \tau \in T$

$$\text{a)}\quad \mu_{\widetilde{Q}_*}(x,\tau) \quad = \underset{U\times T}{\overset{*}{\int\!\!\!\int}} \left[\mu_{E_{\widetilde{u}_*}(\cdot,\tau)}(u) \wedge \mu_{E_{\widetilde{\rho}'_*(x,\cdot,\cdot)}}(u,t)\right] \circ g_U \otimes g_{E_{\widetilde{\mathbb{R}}_*}(\cdot,\tau)}(\cdot), \qquad (41)$$

where $\widetilde{g_U \otimes g_{E_{\widetilde{\mathbb{R}}_*(\cdot,\tau)}}}$ *is an extended composition lower fuzzy measure of the measures* g_U *and* $g_{E_{\widetilde{\mathbb{R}}_*}(\cdot,\tau)}$.

$$\text{b)}\quad \mu_{\widetilde{Q}^*}(x,\tau) \quad = \underset{U\times T}{\overset{*}{\int\!\!\!\int}} \left[\mu_{E_{\widetilde{u}^*}(\cdot,\tau)}(u) \vee \mu_{E_{\widetilde{\rho}'^*(x,\cdot,\cdot)}}(u,t)\right] \circ g_U \otimes g_{E_{\widetilde{\mathbb{R}}^*}(\cdot,\tau)}(\cdot), \qquad (42)$$

where $\widetilde{g_U \otimes g_{E_{\widetilde{\mathbb{R}}^*(\cdot,\tau)}}}$ *is an extended composition upper fuzzy measure of the measures* g_U *and* $g_{E_{\widetilde{\mathbb{R}}^*}(\cdot,\tau)}$.

Theorem 6. *Let* $(\widetilde{r}_{\tau*},\widetilde{r}^*_\tau)_{\tau\geq 0}$ *be some ergodic process of extremal fuzzy time intervals,* $(\widetilde{\mathbb{R}}_*,\widetilde{\mathbb{R}}^*)$ *be an extremal fuzzy reflection process induced by the process* $(\widetilde{r}_{\tau*},\widetilde{r}^*_\tau)_{\tau\geq 0}$, $(\widetilde{Q}_*,\widetilde{Q}^*)$ *be an extremal fuzzy process describing the EFCDS state dynamics, and* $(\widetilde{u}_*,\widetilde{u}^*)$ *be an extremal ergodic fuzzy control process acting on the EFCDS. Then the extremal fuzzy process* $(\widetilde{Q}_*,\widetilde{Q}^*)$ *is ergodic.*

Recalling the notion of lower and upper convergence of sequences of lower and upper measurable functions, respectively (see Subsection 2.1), and also the notion of lower and upper self-continuity of extremal fuzzy measures, we make the following statements on the ergodicity of extremal fuzzy processes.

Definition 30 (g_T-Ergodicity). We say that the fuzzy process of extremal fuzzy time intervals $(\widetilde{r}_{\tau*},\widetilde{r}^*_\tau)_{\tau\geq 0}$ is g_T-ergodic on some extremal fuzzy time intervals $\widetilde{r}_* \in \widetilde{\mathscr{B}}_{T*}$ and $\widetilde{r}^* \in \widetilde{\mathscr{B}}^*_T$, if $\exists \widetilde{r}_{*\infty} \in \widetilde{\mathscr{B}}_{T*}$ and $\widetilde{r}^*_\infty \in \widetilde{\mathscr{B}}^*_T$ extremal fuzzy time intervals such that $\forall \varepsilon > 0$

$$\begin{aligned} &\lim_{\tau\to+\infty} \widetilde{g}_{T*}\left(\widetilde{r}_* \cap \left\{t \in T \mid |\mu_{\widetilde{r}_{\tau*}}(t) - \mu_{\widetilde{r}_{*\infty}}(t)| \geq \varepsilon\right\}\right) = 0, \\ &\lim_{\tau\to+\infty} \widetilde{g}^*_T\left(\widetilde{r}^* \cup \left\{t \in T \mid |\mu_{\widetilde{r}^*_\tau}(t) - \mu_{\widetilde{r}^*_\infty}(t)| < \varepsilon\right\}\right) = 1. \end{aligned} \qquad (43)$$

Definition 31 (g_U-Ergodicity). We say that the extremal fuzzy control process $(\widetilde{u}_*\widetilde{u}^*)$ is g_U-ergodic on some fuzzy control $\widetilde{u} \in \widetilde{\mathscr{B}}_U$ if there exist extremal fuzzy controls $\widetilde{u}_\infty$ and $\widetilde{u}^\infty \in \widetilde{\mathscr{B}}_U$ such that $\forall \varepsilon > 0$ and $\forall t \in T$

$$\begin{aligned} &\lim_{\tau\to+\infty} \widetilde{g}_U\left(\widetilde{u} \cap \left\{u \in U \mid |\mu_{E_{\widetilde{u}_*}(\cdot,\tau)}(u) - \mu_{\widetilde{u}_\infty}(u)| \geq \varepsilon\right\}\right) = 0, \\ &\lim_{\tau\to+\infty} \widetilde{g}_U\left(\widetilde{u} \cup \left\{u \in U \mid |\mu_{E_{\widetilde{u}^*}(\cdot,\tau)}(u) - \mu_{\widetilde{u}^\infty}(u)| < \varepsilon\right\}\right) = 1. \end{aligned} \qquad (44)$$

Analogously to Theorem 6, we formulate the statement that the extremal fuzzy process $(\widetilde{Q}_*, \widetilde{Q}^*)$ of describing the EFCDS state change dynamics is ergodic, where the ergodicity of the processes $(\widetilde{r}_{\tau *}, \widetilde{r}^*_\tau)_{\tau \geq 0}$ and $(\widetilde{u}_*, \widetilde{u}^*)$ is replaced by the g-ergodicity.

Theorem 7. *Let the process $(\widetilde{r}_{\tau *}, \widetilde{r}^*_\tau)_{\tau \geq 0}$ be g_T-ergodic on T with limit extremal fuzzy time intervals $(\widetilde{r}_{*\infty}, \widetilde{r}^*_\infty)$, and $(\widetilde{u}_*, \widetilde{u}^*)$ be g_U-ergodic on U with limit fuzzy controls $(\widetilde{u}_\infty, \widetilde{u}^\infty)$ so that the extended fuzzy measures $\widetilde{g}_{T*}$, $\widetilde{g}^*_T$ be self-continuous. Then the extremal fuzzy process $(\widetilde{Q}_*, \widetilde{Q}^*)$ of describing the EFCDS state change dynamics is ergodic.*

To conclude the section, we say that under the action of the extremal fuzzy control process $(\widetilde{u}_*, \widetilde{u}^*)$ on the EFCDS, the extremal fuzzy process $(\widetilde{Q}_*, \widetilde{Q}^*)$ of describing the EFCDS state change dynamics is ergodic if

a) the processes $(\widetilde{r}_{\tau *}, \widetilde{r}^*_\tau)_{\tau \geq 0}$ and $(\widetilde{u}_*, \widetilde{u}^*)$ are ergodic on T and U, respectively, or

b) the processes $(\widetilde{r}_{\tau *}, \widetilde{r}^*_\tau)_{\tau \geq 0}$ and $(\widetilde{u}_*, \widetilde{u}^*)$ are g-ergodic on T and U, respectively,

the extended extremal fuzzy measures $\widetilde{g}_{T*}$ and $\widetilde{g}^*_T$ are respectively lower self-continuous and upper self-continuous, and the extended measure $\widetilde{g}_U$ is self-continuous on $\widetilde{\mathscr{B}}_U$.

Conclusions. Using the results obtained in [31]–[33] of this study, we have considered questions of fuzzy mathematical modeling of extremal fuzzy processes, where

a) we introduce the notion of an EFCDS with fuzzy uncertainty, the source of which is expert reflections on the states of EFCSD ("expert measurement") in the so-called current and future fuzzy time intervals. The general EFCDS model is described;

b) the notion of processes of expert reflection and description of the EFCDS state change dynamics are introduced. With the aid of the conditional extremal fuzzy measures $g_{t*}(\cdot \mid x)$ and $g^*_t(\cdot, \mid x)$, the extremal fuzzy expert reflection process $(\widetilde{\mathbb{R}}_*, \widetilde{\mathbb{R}}^*)$ connects the fuzzy time interval measurement process $(\widetilde{r}_{\tau *}, \widetilde{r}^*_\tau)_{\tau \geq 0}$ with the space of measurable states of the system with fuzzy distribution $(X, \mathscr{B}, g)$, while the EFCDS state description process $(\widetilde{Q}_*, \widetilde{Q}^*)$ is defined through the extremal fuzzy expert reflection process $(\widetilde{\mathbb{R}}_*, \widetilde{\mathbb{R}}^*)$, using the extended upper and lower Sugeno integrals that are considered as extremal operators describing the EFCDS state dynamics;

c) questions of the ergodicity of extremal fuzzy processes are studied. The notion of g-ergodicity is introduced, which allows one to obtain a sufficient condition for the process $(\widetilde{Q}_*, \widetilde{Q}^*)$ to be ergodic;

d) the notion of an extremal fuzzy control process $(\widetilde{u}_*, \widetilde{u}^*)$ is introduced in the case of the action of control with fuzzy restrictions in the form of the space $(U, \mathscr{B}_U, g_U)$. Models of continuous extremal controllable fuzzy processes are constructed. Questions of the ergodicity of controllable extremal fuzzy processes are studied.

5 The Fuzzy Dynamic Programming Problem

All definitions and results see in [35], [37], [39].

In alternative classical approaches to modeling and when working with the EFCDS the main accent is often placed on the assumption of fuzzyness. We will deal with fuzzy dynamic systems, where fuzzy uncertainty arises with time and time structures are monotone classes of measurable sets.

We start describing objects of a fuzzy dynamic system. Let X $(X \neq \varnothing)$ be the set of states of some system (EFCDS) to be investigated. Let $(X, \mathscr{B}, g)$ be the space of a fuzzy measure on the measurable space $(X, \mathscr{B})$, where $\mathscr{B}$ is a σ-algebra in X (fuzzy restrictions on states).

Let the time structure of fuzzy dynamic system (EFCDS) be represented by (27) and some space of extended extremal fuzzy measures

$$(T, \widetilde{\mathscr{B}}_{T*}, \widetilde{\mathscr{B}}^*_T, \widetilde{g}_{T*}, \widetilde{g}^*_T), \quad T = \mathbb{R}^*_0,$$

where $\widetilde{g}_{T*}$ and $\widetilde{g}^*_T$ are some extremal fuzzy measures on $\widetilde{\mathscr{B}}_{T*} \equiv \widetilde{\mathscr{B}}_{2*}$ and $\widetilde{\mathscr{B}}^*_T \equiv \widetilde{\mathscr{B}}^*_2$, respectively.

Let U $(U \neq \varnothing)$ be the set of all admissible controls (of external factors) acting on the EFCDS. Assume that controls are subjected to restrictions of uncertain character in the form of some space of a fuzzy measure $(U, \mathscr{B}_U, g_U)$, where $\mathscr{B}_U$ is the measurable space of controls, while the fuzzy measure g_U describes the restrictions imposed on controls.

We consider the optimization problems of EFCDS when the model of the continuous extremal controllable fuzzy process is described by the system of fuzzy integral equations ([33] and Section 4):

$$\begin{cases} \mu_{\widetilde{Q}_*}(x,\tau) = \underset{U\times T}{\overset{*}{\int\!\!\int}} \left\{ \mu_{\mathbb{E}_{\widetilde{u}_*}(\cdot,\tau)}(u) \wedge \mu_{\mathbb{E}_{\widetilde{\rho}'_*}(x,\cdot,\cdot)}(u,t) \right\} \circ \widetilde{g_U} \otimes \widetilde{g_{\mathbb{E}_{\widetilde{\mathbb{R}}_*}(\cdot,\tau)}}(\cdot), \\ \mu_{\widetilde{Q}^*}(x,\tau) = \underset{U\times T}{\overset{*}{\int\!\!\int}} \left\{ \mu_{\mathbb{E}_{\widetilde{u}^*}(\cdot,\tau)}(u) \vee \mu_{\mathbb{E}_{\widetilde{\rho}'^*}(x,\cdot,\cdot)}(u,t) \right\} \circ \widetilde{g^*_U} \otimes \widetilde{g_{\mathbb{E}_{\widetilde{\mathbb{R}}^*}(\cdot,\tau)}}(\cdot), \end{cases} \tag{45}$$

where $(\widetilde{Q}_*, \widetilde{Q}^*)$ is a fuzzy extremal process describing the system state dynamics; $(\widetilde{\mathbb{R}}_*, \widetilde{\mathbb{R}}^*)$ is an extremal fuzzy process of expert knowledge reflections in extremal fuzzy time intervals (the expert reflections on the states of EFCDS in the extremal fuzzy time intervals); $(\widetilde{\rho}_*, \widetilde{\rho}^*)$ is the transition operator of the EFCDS states; on right-hand sides of Sugeno extended lower and upper integrals the integration measures are the extremal compositional fuzzy measures extended with respect to the process $(\widetilde{\mathbb{R}}_*, \widetilde{\mathbb{R}}^*)$; μ is a symbol of a compatibility function of a fuzzy set; $\mathbb{E}$ is a symbol of projector of Galois indexing mapping.

We say that the effectiveness of EFCDS control is defined by some set of Criteria K, on which fuzzy restrictions are given for measurable subsets of K, i.e. the fuzzy measure space $(K, \mathscr{B}_K, g_K)$ (fuzzy restriction on the criteria) is defined on K [35].

Let $\widetilde{L} \in \widetilde{\mathscr{B}_K \otimes \mathscr{B}_U}$ be some fuzzy binary relation of "losses" with respect to each of the criteria $v \in K$ in the choice of control $u \in U$. Note that $\mu_{\widetilde{L}}$ is a $\mathscr{B}_K \otimes \mathscr{B}_U$-measurable compatibility function

$$\mu_{\widetilde{L}}(v,u) : K \times U \to [0,1]. \tag{46}$$

Then the complement $\overline{\widetilde{L}}$ is called the fuzzy relation of EFCDS "gain" and the values

$$\mu_{\overline{\widetilde{L}}}(v,u) = 1 - \mu_{\widetilde{L}}(v,u) \tag{47}$$

define the measure of gain in the choice of control $u \in U$ for a criterion $v \in K$.

Definition 32. a) Given all criteria, a $\mathscr{B}_U \otimes \mathscr{B}_T^*$-measurable function: $\forall (u,t) \in U \times T$

$$\mathbf{P}^K_{\widetilde{u}^*}(u,t) \stackrel{\Delta}{=} \int\limits_K \left\{ \mu_{\mathbb{E}_{\widetilde{u}^*}(\cdot,t)}(u) \vee \mu_0^*(u) \vee \mu_{\overline{\widetilde{L}}}(v,u) \right\} \circ \widetilde{g}_K^*(\cdot), \tag{48}$$

where the extended fuzzy measure $\widetilde{g}_K^* : \mathscr{B}_K \to [0,1]$ is the dual fuzzy measure of $\widetilde{g}_K$ ($\forall \widetilde{S} \in \widetilde{\mathscr{B}}_K : \widetilde{g}_K^*(\widetilde{S}) = 1 - \widetilde{g}_K(\overline{\widetilde{S}})$), is called a gain with respect to a current (upper) fuzzy control process $\widetilde{u}^* \in \widetilde{\mathscr{B}_U \otimes \mathscr{B}_T^*}$ with respect to the initial fuzzy control $\mu_{\mathbb{E}_{\widetilde{u}^*}(\cdot,\tau_0)}(u) \equiv \mu_0^*(u)$.

b) Given all criteria, a $\mathscr{B}_U \otimes \mathscr{B}_{T*}$-measurable function: $\forall (u,t) \in U \times T$

$$\mathbf{q}^K_{\widetilde{u}_*}(u,t) \stackrel{\Delta}{=} \int\limits_K \left\{ \mu_{\mathbb{E}_{\widetilde{u}_*}(\cdot,t)}(u) \wedge \mu_{0*}(u) \wedge \mu_{\widetilde{L}}(v,u) \right\} \circ \widetilde{g}_K(\cdot) \tag{49}$$

is called a loss with respect to a future (lower) fuzzy control process $\widetilde{u}_* \in \widetilde{\mathscr{B}_U \otimes \mathscr{B}_{T*}}$ with respect to the initial fuzzy control $\mu_{\mathbb{E}_{\widetilde{u}_*}(\cdot,\tau_0)}(u) \equiv \mu_{0*}(u)$.

Definition 33. a) A $\mathscr{B} \otimes \mathscr{B}_T^*$-measurable function: $\forall (u,\tau) \in U \times T$

$$I_{\widetilde{u}^*}(u,\tau) \stackrel{\Delta}{=} \int\limits_T^* \mathbf{P}^K_{\widetilde{u}^*}(u,t) \circ \widetilde{g}_{\mathbb{E}_{\widetilde{\mathbb{R}}^*}(\cdot,\tau)}(\cdot) \tag{50}$$

is called an integral current gain with respect to a current (upper) fuzzy control process $\widetilde{u}^* \in \widetilde{\mathscr{B}_U \otimes \mathscr{B}_T^*}$ on a current fuzzy time interval $\widetilde{r}_\tau^* \in \widetilde{\mathscr{B}}_T^*$.

b) A $\mathscr{B}_U \otimes \mathscr{B}_{T*}$-measurable function: $\forall (u,\tau) \in U \times T$

$$J_{\widetilde{u}_*}(u,\tau) \stackrel{\Delta}{=} \int\limits_{T*} \mathbf{q}^K_{\widetilde{u}_*}(u,t) \circ \widetilde{g}_{\mathbb{E}_{\widetilde{\mathbb{R}}_*}(\cdot,\tau)}(\cdot) \tag{51}$$

is called an integral future loss with respect to a future (upper) fuzzy control process $\widetilde{u}_* \in \widetilde{\mathscr{B}_U \otimes \mathscr{B}_{T*}}$ on a future fuzzy time interval $\widetilde{r}_{\tau*} \in \widetilde{\mathscr{B}}_{T*}$.

We have thus defined, on U, an extremal fuzzy "gain-loss" process (I_{u^*}, J_{u_*}). Further, for model (45) we will consider, in terms of (50) and (51), the problem of

formation of an optimal control (in the sense of minimization of the future loss and maximization of the current gain) of an extremal process: $\forall(u,t) \in U \times T$

$$\begin{aligned} &\overset{*}{\int\!\!\!\!\!\!\diagup}_{T} \mathbf{P}^{K}_{\widetilde{u}^*}(u,t) \circ \widetilde{g}_{\mathbb{E}_{\widetilde{\mathbb{R}}^*}(\cdot,\tau)}(\cdot) \Rightarrow \max_{\widetilde{u}^*}, \\ &\underset{*}{\int\!\!\!\!\!\!\diagup}_{T} \mathbf{q}^{K}_{\widetilde{u}_*}(u,t) \circ \widetilde{g}_{\mathbb{E}_{\widetilde{\mathbb{R}}_*}(\cdot,\tau)}(\cdot) \Rightarrow \min_{\widetilde{u}_*}. \end{aligned} \tag{52}$$

Functional equations by means of which we can define an extremal fuzzy optimal control in the sense of extremalization of criteria (52) can be written in the following form, $\forall(u,\tau') \in U \times [\tau_0,\tau]$:

$$\begin{cases} \overset{\circ}{J}_{\widetilde{u}_*}(u,\tau') = \underset{\widetilde{u}_* \in \widetilde{\mathscr{B}_U \otimes \mathscr{B}_{T*}}}{\wedge} J_{\widetilde{u}_*}(u,\tau') = \underset{\widetilde{u}_* \in \widetilde{\mathscr{B}_U \otimes \mathscr{B}_{T*}}}{\wedge} \underset{*}{\int\!\!\!\!\!\!\diagup}_{T} \mathbf{q}^{K}_{\widetilde{u}_*}(u,t) \circ \widetilde{g}_{\mathbb{E}_{\widetilde{\mathbb{R}}_*}(\cdot,\tau')}(\cdot), \\ \overset{\circ}{I}_{\widetilde{u}^*}(u,\tau') = \underset{\widetilde{u}_* \in \widetilde{\mathscr{B}_U \otimes \mathscr{B}^*_T}}{\vee} I_{\widetilde{u}^*}(u,\tau') = \underset{\widetilde{u}^* \in \widetilde{\mathscr{B}_U \otimes \mathscr{B}^*_T}}{\vee} \overset{*}{\int\!\!\!\!\!\!\diagup}_{T} \mathbf{P}^{K}_{\widetilde{u}^*}(u,t) \circ \widetilde{g}_{\mathbb{E}_{\widetilde{\mathbb{R}}^*}(\cdot,\tau')}(\cdot), \end{cases} \tag{53}$$

with the initial control conditions

$$\mathbb{E}_{\overset{\circ}{\widetilde{u}_*}}(\cdot,\tau_0) \equiv \widetilde{u}_{0*} \in \mathscr{B}_U, \quad \mathbb{E}_{\overset{\circ}{\widetilde{u}^*}}(\cdot,\tau_0) \equiv \widetilde{u}^*_0 \in \mathscr{B}_U \tag{54}$$

and the EFCDS initial states $\mathbb{E}_{\widetilde{Q}_*}(\cdot,\tau_0)$ and $\mathbb{E}_{\widetilde{Q}^*}(\cdot,\tau_0)$.

Definition 34. An extremal fuzzy control process $(\overset{\circ}{\widetilde{u}}_*, \overset{\circ}{\widetilde{u}}^*)$, $\tau_0 \le \tau' \le \tau$, with the initial conditions (54) is called an optimal for EFCDS (45) in the sense of Bellman's optimality principle if criterion (53) is satisfied.

The following theorem which gives the optimality condition (an analogue of Bellman's equation [1]) is valid.

Theorem 8. *Let a EFCDS be described by system (45). Then an extremal fuzzy control process $(\overset{\circ}{\widetilde{u}}_*, \overset{\circ}{\widetilde{u}}^*)$, $\tau_0 \le \tau' \le \tau$, is optimal in the sense of criterion (53) if and only if the following inequalities are fulfilled:* $\forall(u,\tau') \in U \times [\tau_0,\tau]$

$$\begin{cases} \overset{\circ}{J}_{\widetilde{u}_*}(u,\tau') \le \left(\underset{K}{\int\!\!\!\!\!\!\diagup} \mu_{\widetilde{L}}(v,u) \circ \widetilde{g}_K(\cdot) \right) \wedge \mu_{\mathbb{E}_{\overset{\circ}{\widetilde{u}_*}}(\cdot,\tau_0)}(u), \\ \overset{\circ}{I}_{\widetilde{u}^*}(u,\tau') \ge \left(\underset{K}{\int\!\!\!\!\!\!\diagup} \mu_{\overline{\widetilde{L}}}(v,u) \circ \widetilde{g}^*_K(\cdot) \right) \vee \mu_{\mathbb{E}_{\overset{\circ}{\widetilde{u}^*}}(\cdot,\tau_0)}(u); \end{cases} \tag{55}$$

Theorem 9. *An extremal fuzzy optimal control process $(\overset{\circ}{\widetilde{u}}_*, \overset{\circ}{\widetilde{u}}^*)$ for the EFCDS (45) in the sense of criterion (53) not depending on a EFCDS state can be defined by the following system of fuzzy-integral equations:* $\forall(u,\tau') \in U \times [\tau_0,\tau]$

$$\begin{cases} \mu_{\overset{\circ}{u}_*}(u,\tau') = \mu_{\overset{\circ}{u}_*}(u,\tau_0) \wedge \left(\int\!\!\!\!\!\!\diagup_K \mu_{\widetilde{L}}(v,u) \circ \widetilde{g}_K(\cdot) \right) \wedge \widetilde{g}_{\mathbb{E}_{\widetilde{\mathbb{R}}_*}(\cdot,\Delta(\tau_0,\tau'))}(T), \\ \mu_{\overset{\circ}{u}^*}(u,\tau') = \mu_{\overset{\circ}{u}^*}(u,\tau_0) \vee \left(\int\!\!\!\!\!\!\diagup_K \mu_{\widetilde{L}}(v,u) \circ \widetilde{g}^*_K(\cdot) \right) \vee \widetilde{g}_{\mathbb{E}_{\widetilde{\mathbb{R}}^*}(\cdot,\Delta(\tau_0,\tau'))}(T). \end{cases} \tag{56}$$

Remark 2. Expressions in (56) of an extremal optimal fuzzy control process $(\overset{\circ}{u}_*, \overset{\circ}{u}^*)$, $\tau_0 \le \tau' \le \tau$, are a variant of the solution of inequalities (55), but this fuzzy-integral representation of an optimal control gives a good analogue of the solution of the problem of stochastic dynamic programming, where the expression of an optimal control contains "direct" analogues to (56): $\int\!\!\!\!\!\!\diagup_K \mu_{\widetilde{L}}(v,u) \circ g_K(\cdot)$ is the Bellman functional which is an analogue of the kernel in the representation of a stochastic optimal control or, more exactly, an analogue of the signal of a stochastic model or its deterministic part, while the values of the extended fuzzy measures $\widetilde{g}_{\mathbb{E}_{\widetilde{\mathbb{R}}_*}(\cdot,\Delta(\tau_0,\tau'))}(T)$ and $\widetilde{g}_{\mathbb{E}_{\widetilde{\mathbb{R}}^*}(\cdot,\Delta(\tau_0,\tau'))}(T)$ are analogues of stochastic measure in the representation of stochastic optimal controls.

The case where a fuzzy control of EFCDS depends not only on time $\tau' \in [0,\tau]$ but also on a EFCDS state $x \in X$ is also studied but is omitted here.

5.1 Example

Let the set of EFCDS states be finite, $X = \{1,2,3,4\}$; $g^* : 2^X \to [0,1]$ be the possibility measure with the possibility distribution on X

$$\Pi(i) \overset{\Delta}{=} \frac{i}{4}, \quad i = 1,2,3,4 \quad \left(\forall B \in 2^X : g^*(A) = \bigvee_{i \in A} \pi(i) \right).$$

Let the EFCDS be subjected to the influence of an external control factor with the finite set $U = \{u_1, u_2\}$ (for example, $u_1 \overset{\Delta}{=}$"+1", $u_2 \overset{\Delta}{=}$"−1"). Let the uniform probability distribution play the role of the fuzzy measure $g_U : 2^U \to [0,1]$, i.e. $g_U(\{u_1\}) = g_U(\{u_2\}) = \frac{1}{2}$. The two-element set $K = \{v_1, v_2\}$ is taken as the set of chosen criteria, while the uniform probability distribution $g_K(\{v_1\}) = g_K(\{v_2\}) = \frac{1}{2}$ is considered as playing the role of the fuzzy measure $g_K : 2^K \to [0,1]$. Thus we have the fuzzy measure spaces $(X, 2^X, g)$, $(K, 2^K, g_K)$ and $(U, 2^U, g_U)$. The dual measure g^* on 2^X is the necessity measure $g(A) = 1 - \bigvee_{i \notin A} \pi(i)$. Since the fuzzy measures g_U and g_K are the probability ones, we know they are autodual and

$$g^*_U = g_U, \quad g^*_K = g_K.$$

It is assumed that the initial moment of EFCDS observation is $\tau_0 \equiv 0$. Let the initial extremal fuzzy distributions of an optimal control be

$$\mu_{\widetilde{\widetilde{u}}_*}(u_1,0)=\frac{1}{2}=\mu_{\widetilde{\widetilde{u}}^*}(u_1,0);\ \mu_{\widetilde{\widetilde{u}}_*}(u_2,0)=\frac{1}{4}=\mu_{\widetilde{\widetilde{u}}^*}(u_2,0).$$

Let the binary fuzzy loss relation $\widetilde{L}$ on $U\times K$ be defined as follows:

$$\mu_{\widetilde{L}}(u_1,v_1)=\mu_{\widetilde{L}}(u_2,v_2)=\frac{1}{2},\ \mu_{\widetilde{L}}(u_1,v_2)=\mu_{\widetilde{L}}(u_2,v_1)=\frac{1}{4}.$$

The distributions of extremal fuzzy time intervals are given as

$$\mu_{\widetilde{r}_{\tau*}}(t)=\begin{cases}0, & 0\le t\le\tau,\\ 1-\frac{\tau}{t}, & t>\tau,\end{cases}\qquad \mu_{\widetilde{r}^*_{\tau}}(t)=\begin{cases}1, & 0\le t<\tau,\\ \frac{\tau}{t}, & t\ge\tau.\end{cases}\tag{57}$$

Let the initial distribution ($\tau_0\equiv 0$) of the EFCDS state description process look like

$$\widetilde{A}_{0*}\sim\begin{pmatrix}1&2&3&4\\ \frac{1}{4}&\frac{1}{4}&\frac{1}{2}&\frac{1}{2}\end{pmatrix},\quad \widetilde{A}^*_0\sim\begin{pmatrix}1&2&3&4\\ \frac{1}{2}&\frac{1}{2}&\frac{1}{2}&\frac{1}{2}\end{pmatrix}.\tag{58}$$

We consider the example of the space $(T,\widetilde{\mathscr{B}}_{T*},\widetilde{\mathscr{B}}^*_T,\widetilde{g}^*_T,\widetilde{g}^*_T)$ where

$$\begin{aligned}&g_{T*}([t,+\infty))\overset{\Delta}{=}\frac{1}{1+t},\quad [t,+\infty)\in\mathscr{B}_{T*},\\ &g^*_T([0,t))\overset{\Delta}{=}\frac{t}{1+t},\quad [0,t)\in\mathscr{B}^*_T,\ \ t>0.\end{aligned}\tag{59}$$

Further, we introduce the conditional fuzzy measures on $\mathscr{B}_{T*}$ and $\mathscr{B}^*_T$ with respect to the set $X=\{1,2,3,4\}$:

$$\begin{aligned}&g_{t*}(r_{\tau*}\mid i)=\frac{1}{1+i\tau},\ \text{ where }\ i\in X,\ \ r_{\tau*}\in\mathscr{B}_{T*},\\ &g^*_t(r^*_\tau\mid i)=\frac{i\tau}{1+i\tau},\ \text{ where }\ i\in X,\ \ r^*_\tau\in\mathscr{B}^*_T.\end{aligned}\tag{60}$$

Thus the EFCDS state description process can be represented as follows:

$$\begin{cases}\mu_{\widetilde{Q}_*}(x,\tau)=\underset{U\times T}{\overset{*}{\bigwedge\!\!\!\bigwedge}}\left\{\mu_{\mathbb{E}_{\widetilde{u}_*}(\cdot,\tau)}(u)\wedge\mu_{\mathbb{E}_{\widetilde{\rho}'_*(x,\cdot,\cdot)}}(u,t)\right\}\circ\widetilde{g_U}\otimes\widetilde{g_{\mathbb{E}_{\widetilde{\mathbb{R}}_*}(\cdot,\tau)}}(\cdot),\\ \mu_{\widetilde{Q}^*}(x,\tau)=\underset{U\times T}{\overset{*}{\bigvee\!\!\!\bigvee}}\left\{\mu_{\mathbb{E}_{\widetilde{u}^*}(\cdot,\tau)}(u)\vee\mu_{\mathbb{E}_{\widetilde{\rho}'^*(x,\cdot,\cdot)}}(u,t)\right\}\circ\widetilde{g_U}\otimes\widetilde{g_{\mathbb{E}_{\widetilde{\mathbb{R}}^*}(\cdot,\tau)}}(\cdot),\end{cases}\tag{61}$$

where $\widetilde{A}_{0*}\equiv\mathbb{E}_{\widetilde{Q}_*}(\cdot,0)$, $\widetilde{A}^*_0\equiv\mathbb{E}_{\widetilde{Q}^*}(\cdot,0)$, $(\widetilde{\mathbb{R}}_*,\widetilde{\mathbb{R}}^*)$ is the extremal fuzzy reflection process, $\forall(x,\tau)\in X\times T$, $\forall(x,t)\in U\times T$:

$$\begin{cases}\mu_{\widetilde{\mathbb{R}}_*}(x,\tau)\overset{\Delta}{=}\widetilde{g}_{t*}(\widetilde{r}_{\tau*}\mid x)=\mu_{\widetilde{A}_{\tau*}}(x),\\ \mu_{\widetilde{\mathbb{R}}^*}(x,\tau)\overset{\Delta}{=}\widetilde{g}^*_t(\widetilde{r}^*_\tau\mid x)=\mu_{\widetilde{A}^*_\tau}(x),\end{cases}\tag{62}$$

and

$$\begin{cases} \mu_{\widetilde{\rho}'_*}(x,u,t) \overset{\Delta}{=} \int\limits_{\widetilde{A}_{0*}} \mu_{\widetilde{\rho}_*}(x,u,x',t) \circ \widetilde{g}(\cdot), \\ \mu_{\widetilde{\rho}'^*}(x,u,t) \overset{\Delta}{=} \int\limits_{\widetilde{A}^*_0} \mu_{\widetilde{\rho}^*}(x,u,x',t) \circ \widetilde{g}^*(\cdot), \end{cases} \tag{63}$$

where $\widetilde{A}_{\tau *} \in \widetilde{\mathscr{B}}$ and $\widetilde{A}^*_\tau \in \mathscr{B}$ are expert reflections on the EFCDS states in the fuzzy extremal intervals $\widetilde{r}_{\tau *} \in \widetilde{\mathscr{B}}_{T*}$ and $\widetilde{r}^*_\tau \in \widetilde{\mathscr{B}}^*_T$, respectively; $(\widetilde{\rho}_*, \widetilde{\rho}^*)$ is the EFCDS transition operator (see [34]). As known the operator $(\widetilde{\rho}'_*, \widetilde{\rho}'^*)$ is restored from the experimental-expert knowledge base on the EFCDS so that if we fix some admissible extremal control process $(\widetilde{u}_*, \widetilde{u}^*)$ (including an optimal control too), then, using the calculation procedure for Sugeno extremal integrals [34], we can write expressions for the process $(\widetilde{Q}_*, \widetilde{Q}^*)$. However we pursue a different aim here: using EFCDS data, we are to construct the extremal optimal control process $(\overset{\circ}{\widetilde{u}}_*, \overset{\circ}{\widetilde{u}}^*)$.

Since the sets X, U, K are finite, it is not difficult to check that the conditions (55) of existence of an optimal extremal control process are satisfied. By virtue of the results of Theorems 8 and 9, we can write one of the variants for an extremal optimal fuzzy control process as follows: $\forall (u,\tau) \in (X,T)$

$$\begin{cases} \mu_{\overset{\circ}{\widetilde{u}}_*}(u,\tau) = \mu_{\overset{\circ}{\widetilde{u}}_*}(u,0) \wedge \left(\int\limits_K \mu_{\widetilde{L}}(u,v) \circ \widetilde{g}_K(\cdot) \right) \wedge \widetilde{g}_{\mathbb{E}_{\widetilde{\mathbb{R}}_*}(\cdot,\tau)}(T), \\ \mu_{\overset{\circ}{\widetilde{u}^*}}(u,\tau) = \mu_{\overset{\circ}{\widetilde{u}^*}}(u,0) \vee \left(\int\limits_K \mu_{\widetilde{\overline{L}}}(u,v) \circ \widetilde{g}^*_K(\cdot) \right) \vee \widetilde{g}_{\mathbb{E}_{\widetilde{\mathbb{R}}^*}(\cdot,\tau)}(T), \end{cases} \tag{64}$$

where $u \in \{\text{“}+1\text{”}, \text{“}-1\text{”}\}$, $v \in \{v_1, v_2\}$; $\mu_{\overset{\circ}{\widetilde{u}_*}}(u,0)$ and $\mu_{\overset{\circ}{\widetilde{u}^*}}(u,0)$ are already defined, while the extended extremal fuzzy measures are defined in the form:

$$\begin{cases} \widetilde{g}_{\mathbb{E}_{\widetilde{\mathbb{R}}_*}(\cdot,\tau)}(T) = \int\limits_T \mu_{\widetilde{r}_{\tau *}}(t) \circ \widetilde{g}_{T*}(\cdot) \overset{\Delta}{=} \int\limits_T \mu_{\widetilde{r}_{\tau *}}(t) \circ \int\limits_X g_{t*}(\cdot \mid x) \circ g(\cdot), \\ \widetilde{g}_{\mathbb{E}_{\widetilde{\mathbb{R}}^*}(\cdot,\tau)}(T) = \int\limits_T{}^* \mu_{\widetilde{r}^*_\tau}(t) \circ \widetilde{g}^*_T(\cdot) \overset{\Delta}{=} \int\limits_T{}^* \mu_{\widetilde{r}^*_\tau}(t) \circ \int\limits_X g^*_t(\cdot \mid x) \circ g^*(\cdot). \end{cases} \tag{65}$$

Now we are to calculate the Sugeno integrals in formulas (64) and the values of extremal fuzzy measures (65).

Let us calculate the values of $\oint\limits_K \mu_{\widetilde{L}}(u,v)\circ\widetilde{g}_K(\cdot)$:

1) $u=u_1\equiv$"$+1$":

$$\oint\limits_K \mu_{\widetilde{L}}(u_1,v)\circ\widetilde{g}_K(\cdot)=\bigwedge_{0<\alpha\le 1}\left\{\alpha\vee g_K(v\in K\mid \mu_{\widetilde{L}}(u_1,v)\ge\alpha\right\}$$
$$=\left[\bigwedge_{0\le\alpha\le\frac14}(\alpha\vee g_K(K))\right]\wedge\left[\bigwedge_{\frac14\le\alpha\le\frac12}(\alpha\vee g_K(\{v_2\}))\right]$$
$$\wedge\left[\bigwedge_{\frac12<\alpha\le 1}(\alpha\vee g_K(\varnothing))\right]=1\wedge\frac12\wedge\frac12=\frac12.$$

2) $u=u_2=$"-1":

$$\oint\limits_K \mu_{\widetilde{L}}(u_2,v)\circ\widetilde{g}_K(\cdot)=\bigwedge_{0<\alpha\le 1}\left\{\alpha\vee g_K(v\in K\mid \mu_{\widetilde{L}}(u_2,v)\ge\alpha\right\}$$
$$=\left[\bigwedge_{0\le\alpha\le\frac14}(\alpha\vee g_K(K))\right]\wedge\left[\bigwedge_{\frac14\le\alpha\le\frac12}(\alpha\vee g_K(\{v_1\}))\right]\wedge\left[\bigwedge_{\frac12<\alpha\le 1}(\alpha\vee g_K(\varnothing))\right]$$
$$=1\wedge\left[\bigwedge_{\frac14\le\alpha\le\frac12}\left(\alpha\vee\frac12\right)\right]\wedge\left[\bigwedge_{\frac12\le\alpha<1}(\alpha)\right]=1\wedge\frac12\wedge\frac12=\frac12.$$

Since

$$\oint\limits_K \mu_{\overline{\widetilde{L}}}(u,v)\circ\widetilde{g}^*_K(\cdot)=1-\oint\limits_K \mu_{\widetilde{L}}(u,v)\circ\widetilde{g}_K(\cdot),$$

we have

$$\oint\limits_K \mu_{\overline{\widetilde{L}}}(u_1,v)\circ\widetilde{g}^*_K(\cdot)=\oint\limits_K \mu_{\overline{\widetilde{L}}}(u_2,v)\circ\widetilde{g}_K(\cdot)=\frac12.$$

Therefore $\forall\tau>0$

$$\mu_{\widetilde{\widetilde{u}}_*}(u_1,\tau)=\frac12\wedge\frac12\wedge\widetilde{g}_{\mathbb{E}_{\widetilde{\mathbb{R}}_*}(\cdot,\tau)}(T)=\frac12\wedge\widetilde{g}_{\mathbb{E}_{\widetilde{\mathbb{R}}_*}(\cdot,\tau)}(T),$$
$$\mu_{\widetilde{\widetilde{u}}_*}(u_2,\tau)=\frac14\wedge\frac12\wedge\widetilde{g}_{\mathbb{E}_{\widetilde{\mathbb{R}}_*}(\cdot,\tau)}(T)=\frac14\wedge\widetilde{g}_{\mathbb{E}_{\widetilde{\mathbb{R}}_*}(\cdot,\tau)}(T),$$
$$\mu_{\widetilde{\widetilde{u}}^*}(u_1,\tau)=\frac12\vee\frac12\vee\widetilde{g}_{\mathbb{E}_{\widetilde{\mathbb{R}}^*}(\cdot,\tau)}(T)=\frac12\vee\widetilde{g}_{\mathbb{E}_{\widetilde{\mathbb{R}}^*}(\cdot,\tau)}(T),$$
$$\mu_{\widetilde{\widetilde{u}}^*}(u_2,\tau)=\frac14\vee\frac12\vee\widetilde{g}_{\mathbb{E}_{\widetilde{\mathbb{R}}^*}(\cdot,\tau)}(T)=\frac12\vee\widetilde{g}_{\mathbb{E}_{\widetilde{\mathbb{R}}^*}(\cdot,\tau)}(T).$$

Now we are to calculate the values of the so-called extremal fuzzy "white noise" (65):

$$\widetilde{g}_{\mathbb{E}_{\widetilde{\mathbb{R}}_*}(\cdot,\tau)}(T)=\oint\limits^*_T \mu_{\widetilde{r}_{\tau*}}(t)\circ\oint\limits_X g_{t*}(\cdot\mid x)\circ g(\cdot)$$
$$=\bigvee_{0<\alpha\le 1}\{\alpha\wedge\widetilde{g}_{T*}([\widetilde{r}_{\tau*}]_{\overline{\alpha}})\}=\bigvee_{0<\alpha\le 1}\left\{\alpha\wedge\oint\limits_X\widetilde{g}_{t*}([\widetilde{r}_{\tau*}]_{\overline{\alpha}}\mid x)\circ g(\cdot)\right\}.$$

From (57) we obtain the expression for an α-cut for $\widetilde{r}_{\tau *}$:

$$[\widetilde{r}_{\tau *}]_{\overline{\alpha}} = \left\{ \begin{matrix} T & \text{if } \alpha = 0, \\ \left[\frac{\tau}{1-\alpha}, +\infty\right) & \text{if } 0 < \alpha < 1, \\ \varnothing & \text{if } \alpha = 1, \end{matrix} \right\} \in \mathscr{B}_{T*}.$$

Now (60) implies

$$\widetilde{g}_{t*}([\widetilde{r}_{\tau *}]_{\overline{\alpha}} \mid i) = \begin{cases} 1 & \text{if } \alpha = 0, \\ \frac{1}{1+i\frac{\tau}{1-\alpha}} & \text{if } 0 < \alpha < 1, \quad \forall i \in X. \\ \varnothing & \text{if } \alpha = 1, \end{cases}$$

and

$$\oint_X \widetilde{g}_t([\widetilde{r}_{\tau *}]_{\overline{\alpha}} \mid i) \circ g(\cdot) = \bigvee_{0<\beta\le 1} \left\{ \beta \wedge g\left(\left\{ i \in X \mid \frac{1}{1+i\frac{\tau}{1-\alpha}} \ge \beta \right\} \right) \right\}.$$

It is not difficult to verify that $(0 < \alpha < 1,\ \tau > 0)$

$$\left\{ i \in X \mid \frac{1}{1+i\frac{\tau}{1-\alpha}} \ge \beta \right\} = \begin{cases} \varnothing & \text{if } 1 \ge \beta > \frac{1-\alpha}{1-\alpha+\tau}, \\ \{1\} & \text{if } \frac{1-\alpha}{1-\alpha+\tau} \ge \beta > \frac{1-\alpha}{1-\alpha+2\tau}, \\ \{1,2\} & \text{if } \frac{1-\alpha}{1-\alpha+2\tau} \ge \beta > \frac{1-\alpha}{1-\alpha+3\tau}, \\ \{1,2,3\} & \text{if } \frac{1-\alpha}{1-\alpha+3\tau} \ge \beta > \frac{1-\alpha}{1-\alpha+4\tau}, \\ X & \text{if } \frac{1-\alpha}{1-\alpha+4\tau} \ge \beta > 0. \end{cases}$$

Denote $B_0 \equiv \left(\frac{1-\alpha}{1-\alpha+\tau}; 1\right]$, $B_1 \equiv \left(\frac{1-\alpha}{1-\alpha+2\tau}; \frac{1-\alpha}{1-\alpha+\tau}\right]$, $B_2 \equiv \left(\frac{1-\alpha}{1-\alpha+3\tau}; \frac{1-\alpha}{1-\alpha+2\tau}\right]$, $B_3 \equiv \left(\frac{1-\alpha}{1-\alpha+4\tau}; \frac{1-\alpha}{1-\alpha+3\tau}\right]$, $B_4 \equiv \left(0; \frac{1-\alpha}{1-\alpha+4\tau}\right]$.

Then

$$\oint_X \widetilde{g}_{t*}([\widetilde{r}_{\tau *}]_{\overline{\alpha}} \mid x) \circ g(\cdot) = \left[\bigvee_{\beta\in B_0} (\beta \wedge g(\varnothing)) \right] \vee \left[\bigvee_{\beta\in B_1} (\beta \wedge g(\{1\}) \right]$$

$$\vee \left[\bigvee_{\beta\in B_2} (\beta \vee g(\{1,2\})) \right] \vee \left[\bigvee_{\beta\in B_3} (\beta \vee g(\{1,2,3\})) \right] \vee \left[\bigvee_{\beta\in B_4} (\beta \wedge g(X)) \right]$$

$$= 0 \vee \left[\bigvee_{\beta\in B_1} (\beta \wedge 0) \right] \vee \left[\bigvee_{\beta\in B_2} (\beta \wedge 0) \right] \vee \left[\bigvee_{\beta\in B_3} (\beta \wedge 0) \right]$$

$$\vee \left[\bigvee_{\beta\in B_4} (\beta \wedge 1) \right] = \bigvee_{\beta\in B_4} \beta = \frac{1-\alpha}{1-\alpha+4\tau}.$$

We finally obtain

$$\widetilde{g}_{\mathbb{E}_{\widetilde{\mathbb{R}}_*}(\cdot,\tau)}(T) = \bigvee_{0<\alpha<1} \left\{ \alpha \wedge \oint_X \widetilde{g}_{t*}([\widetilde{r}_{\tau *}]_{\overline{\alpha}} \mid x) \circ g(\cdot) \right\} = \bigvee_{0<\alpha<1} \left\{ \alpha \wedge \frac{1-\alpha}{1-\alpha+4\tau} \right\}.$$

After studying the function in the braces with respect to α, we can continue calculations:

$$\widetilde{g}_{\mathbb{E}_{\widetilde{\mathbb{R}}_*}(\cdot,\tau)}(T)=\begin{cases}\bigvee\limits_{0<\alpha<1}\{\alpha\}=1 & \text{if } 0<\tau\le 1,\\ \bigvee\limits_{\alpha\in[1;2\tau-1-2\sqrt{\tau(\tau-1)}]}\{\alpha\} & \\ \quad =2\tau-1-2\sqrt{\tau(\tau-1)} & \text{if } \tau>1.\end{cases}$$

Since $\widetilde{g}_{\mathbb{E}_{\widetilde{\mathbb{R}}_*}(\cdot,\tau)}(\cdot)$ and $\widetilde{g}_{\mathbb{E}_{\widetilde{\mathbb{R}}^*}(\cdot,\tau)}(\cdot)$ are extended extremal measures, we have

$$\widetilde{g}_{\mathbb{E}_{\widetilde{\mathbb{R}}^*}(\cdot,\tau)}(T)=\begin{cases}0 & \text{if } 0<\tau\le 1,\\ 2+2\sqrt{\tau(\tau-1)} & \text{if } \tau>1.\end{cases}$$

For an optimal control we obtain the following expressions:

$$\mu_{\widetilde{u}_*}(u_1,\tau)=\begin{cases}\frac{1}{2}, & 0<\tau\le 1,\\ \frac{1}{2}\wedge(2\tau-1-2\sqrt{\tau(\tau-1)}), & \tau>1,\end{cases}$$

$$\mu_{\widetilde{u}_*}(u_2,\tau)=\begin{cases}\frac{1}{4}, & 0<\tau\le 1,\\ \frac{1}{4}\wedge(2\tau-1-2\sqrt{\tau(\tau-1)}), & \tau>1,\end{cases}$$

$$\mu_{\widetilde{u}^*}(u_1,\tau)=\begin{cases}\frac{1}{2}, & 0<\tau\le 1,\\ \frac{1}{2}\vee(2+2\sqrt{\tau(\tau-1)}-2\tau), & \tau>1\end{cases}=\mu_{\widetilde{u}^*}(u_2,\tau).$$

Note that when $\tau\to+\infty$ a current description process of fuzzy time intervals extends unlimitedly, while a future description process of fuzzy time intervals vanishes. The latter fact is reflected in the expressions for the fuzzy optimal extremal controls:

$$\begin{cases}\lim\limits_{\tau\to\infty}\mu_{\widetilde{u}^*}(u,\tau)\to 1, & u\in U=\{u_1,u_2\},\\ \lim\limits_{\tau\to\infty}\mu_{\widetilde{u}_*}(u,\tau)\to 0, & u\in U=\{u_1,u_2\}.\end{cases}$$

i.e. the uncertainty for a current fuzzy control process vanishes, while a future fuzzy optimal control process is not considered.

We have thereby finished the consideration of the example.

6 Conclusion

Using the results presented in the papers [31]–[39], we have considered questions of the fuzzy optimization of extremal processes, where:

a) the basic properties of Sugeno's type extremal fuzzy measure and several variants of its representations are considered;

b) the notions of extremal fuzzy time moments and intervals are introduced and their monotone algebraic structures are defined. The dualization of a time structure

forms the most important part of the fuzzy instrument of modeling and optimization of extremal fuzzy continuous dynamic systems;

c) we introduce the notion of an EFCDS with fuzzy uncertainty, the source of which is "fuzzy measurement" ("expert reflections" on the states of EFCDS) of the system state in the so-called current and future fuzzy time intervals. The general EFCDS model is described;

d) the notion of processes of expert reflection and description of the EFCDS state change dynamics are introduced. With the aid of the conditional extremal expert reflection measures $g_{t*}(\cdot \mid x)$ and $g_t^*(\cdot, \mid x)$, the extremal fuzzy reflection process $(\widetilde{\mathbb{R}}_*, \widetilde{\mathbb{R}}^*)$ connects the fuzzy time interval measurement process $(\widetilde{r}_{\tau *}, \widetilde{r}_\tau^*)_{\tau \geq 0}$ with the space of measurable states of the system with fuzzy distribution $(X, \mathscr{B}, g)$, while the EFCDS state description process $(\widetilde{Q}_*, \widetilde{Q}^*)$ is defined through the extremal fuzzy reflection process $(\widetilde{\mathbb{R}}_*, \widetilde{\mathbb{R}}^*)$, using the extended upper and lower Sugeno integrals that are considered as extremal operators describing the EFCDS state dynamics;

e) consideration is given to the continuous case of extremal fuzzy processes. Questions of the ergodicity of extremal fuzzy processes are studied. The notion of g-ergodicity is introduced, which allows one to obtain a sufficient condition for the process $(\widetilde{Q}_*, \widetilde{Q}^*)$ to be ergodic;

f) the notion of an extremal fuzzy control process $(\widetilde{u}_*, \widetilde{u}^*)$ is introduced in the case of the action of control with fuzzy restrictions in the form of the space $(U, \mathscr{B}_U, g_U)$. Models of continuous extremal controllable fuzzy processes are constructed. Questions of the ergodicity of controllable extremal fuzzy processes are studied;

g) problems of optimization of a continuous controllable extremal fuzzy process are considered using R. Bellman's optimality principle. An extremal fuzzy "gain-loss" process is defined, which plays the role of Bellman's function in the classical variant of the dynamic programming problem. Theorems 8 and 9 allow one to write variants of an optimal control for the EFCDS;

h) a practical example is given to illustrate the results obtained.

References

1. Bellman, R.E.: Dynamic Programming. Princeton University Press, Princeton (1947)
2. Bellman, R.E., Esogbue, A.O.: Fuzzy Dynamic Programming and Its Extensions. In: Fuzzy Sets Secision Analysis. Stud. Management Sci., vol. 20, pp. 147–167. North-Holland, Amsterdam (1984)
3. Bellman, R.E., Zadeh, L.A.: Decision-making in a fuzzy environment. Manage Sci. Ser. B. 17, 141–164 (1970)
4. Buckley, J.J., Feuring, J.: Fuzzy differential equations. Fuzzy Sets and Systems 110, 43–54 (2000)
5. Castillo, O., Melin, P.: Soft computing for Control of Non-Linear Dynamic Systems. Studies in Fuzziness and Soft Computing, vol. 73. Physica-Verlag, Wulzburg (2001)
6. Diamond, P., Kloeden, P.: Metric Spaces of Fuzzy Sets: Theory and Applications. World Scientific, Singapore (1994)
7. Ding, Z., Ma, M., Kandel, A.: On the observativity of fuzzy dynamical control systems (I). Fuzzy Sets and Systems 95, 53–65 (1998)
8. Dubois, D., Prade, H.: Possibility Theory. Plenum Press, New York (1988)

9. Dumitrescu, D., Haloiu, C., Dumitrescu, A.: Generators of Fuzzy Dynamical Systems. Fuzzy Sets and Systems 113, 447–451 (2000)
10. Esogbue, A.O., Fedrizzi, M., Kacprzyk, J.: Fuzzy dynamic programming with stochastic systems. In: Combining Fuzzy Imprecision with Probabilistic Uncertainty in Decision Making. Lecture Notes in Econom. and Math. Systems, vol. 310, pp. 266–285. Springer, Berlin (1988)
11. Esogbue, A.O., Kacprzyk, J.: Fuzzy Dynamic Programming. In: Fuzzy Sets in Decision Analysis, Operations Research and Statistics. Handb. Fuzzy Sets Ser., vol. 1, pp. 281–307. Kluwer Acad. Publ., Boston (1998)
12. Feng, Y.: Fuzzy stochastic differential systems. Fuzzy Sets and Systems 115, 351–363 (2000)
13. Freidman, M., Henne, M., Kandel, A.: Most typical values for fuzzy sets. Fuzzy Sets and Systems 87, 27–37 (1997)
14. Grabisch, M.: Fuzzy integral in multicriteria decision making. Fuzzy sets and Systems 69, 279–298 (1995)
15. Grabisch, M., Murofushi, T., Sugeno, M. (eds.): Fuzzy Measures and Integrals. Theory and Applications. Studies in Fuzziness and Soft Computing, vol. 40. Physica-Verlag, Heidelberg (2000)
16. Klir, G.J.: Fuzzy Measure Theory. Plenum Press, New York (1992)
17. Kaleva, O.: Fuzzy differential equations. Fuzzy Sets and Systems 24, 301–317 (1987)
18. Kandel, A.: Fuzzy statistics and forecast evaluation. IEEE Trans. on Systems, Man and Cybernetics, SMG-8 5, 396–401 (1978)
19. Kandel, A., Byatt, W.J.: Fussy processes. Fuzzy Sets and Systems 4, 117–152 (1980)
20. Klir, G.J.: Architecture of Systems Problem Solving. Plenum Press, New York (1985)
21. Klir, G.J., Folger, T.A.: Fuzzy Sets, Uncertainty and Information. Prentice Hale, England (1988)
22. Klir, G.J., Wang, Z., Harmanec, D.: Construction of fuzzy measures in expert systems. Fuzzy Sets and Systems 92, 251–264 (1997)
23. Kloeden, P.E.: Fuzzy dynamical systems. Fuzzy Sets and Systems 7(3), 275–296 (1982)
24. Kurano, M., Yasuda, M., Nakagami, J., Yoshida, Y.: A fuzzy relational equation in dynamic fuzzy systems. Fuzzy Sets and Systems 101, 439–443 (1999)
25. Liu, B.: Toward fuzzy optimization without mathematical ambiguity. Fuzzy Optim. Decis. Mak. 1(1), 43–63 (2002)
26. Melin, P., Castillo, O.: Modelling, simulation and control of non-linear dynamical systems. In: An Intelligent Approach Using Soft Computing and Fractal Theory. With 1 IBM-PC floppy disk (3.5 inch; HD). Numerical Insights, vol. 2. Taylor & Francis, Ltd., London (2002)
27. Pack, J.Y., Han, H.K.: Fuzzy differential equations. Fuzzy Sets and Systems 110, 63–77 (2000)
28. Pap, E.: Null-additive Set Functions. In: Mathematics and its Applications, vol. 337. Kluwer Academic Publishers Group, Dordrecht (1995)
29. Piegat, A.: Fuzzy modeling and control. Studies in Fuzziness and Soft Computing, vol. 69. Physica-Verlag, Heidelberg (2001)
30. Ruan, D., Kacprzyk, J., Fedrizzi, M.: Soft Computing for Risk Evaluation and Management. In: Applications in Technology, Environment and Finance. Physica-Verlag, Heidelberg (2001)
31. Sirbiladze, G.: Modeling of extremal fuzzy dynamic systems. I. Extended extremal fuzzy measures. Int. J. Gen. Syst. 34(2), 107–138 (2005)
32. Sirbiladze, G.: Modeling of extremal fuzzy dynamic systems. II. Extended extremal fuzzy measures on composition products of measurable spaces. Int. J. Gen. Syst. 34(2), 139–167 (2005)

33. Sirbiladze, G.: Modeling of extremal fuzzy dynamic systems. III. Modeling of extremal and controllable extremal fuzzy processes. Int. J. Gen. Syst. 34(2), 169–198 (2005)
34. Sirbiladze, G.: Modeling of extremal fuzzy dynamic systems. IV. Identification of fuzzy-integral models of extremal fuzzy processes. Int. J. Gen. Syst. 35(4), 435–459 (2006)
35. Sirbiladze, G.: Modelling of extremal fuzzy dynamic systems. V. Optimization of continuous controllable extremal fuzzy processes and the choice of decisions. Int. J. Gen. Syst. 35(5), 529–554 (2006)
36. Sirbiladze, G.: Modeling of extremal fuzzy dynamic systems. VI. Problems of states estimation (filtration) of extremal fuzzy process. Int. J. Gen. Syst. 36(1), 19–58 (2007)
37. Sirbiladze, G.: On fuzzy optimal controls in the weakly structurable continuous dynamic systems (WSCDS). New Math. Nat. Comput. 4(1), 41–60 (2008)
38. Sirbiladze, G.: Transformation theorems for extended lower and upper Sugeno integrals. Mat. Zametki 83(3), 439–460 (2008) (Russian)
39. Sirbiladze, G.: Bellman's Optimality Principle in the Weakly Structurable Dynamic Systems. In: 9th WSEAS Interneational Conference on Fuzzy Systems (FS 2008), Advanced Topics on Fuzzy Systems, Sofia, Bulgaria, May 2-4 (2008)
40. Sirbiladze, G., Gachechiladze, T.: Restored fuzzy measures in expert decision-making. Inform. Sci. 169(1/2), 71–95 (2005)
41. Sirbiladze, G., Sikharulidze, A.: Weighted fuzzy averages in fuzzy environment. I. Insufficient expert data and fuzzy averages. Internat. J. Uncertain. Fuzziness Knowledge-Based Systems 11(2), 139–157 (2003)
42. Sirbiladze, G., Sikharulidze, A.: Weighted fuzzy averages in fuzzy environment. II. Generalized weighted fuzzy expected values in fuzzy environment. Internat. J. Uncertain. Fuzziness Knowledge-Based Systems 11(2), 159–172 (2003)
43. Srivastava, P., Khare, M., Srivastava, Y.K.: Fuzzy dynamical systems-inverse and direct spectra. Fuzzy Sets and Systems 113, 439–445 (2000)
44. Sugeno, M.: Theory of fuzzy integrals and its applications. Ph.D. Thesis of Tokyo Insitute of Technology (1974)
45. Teodorescu, H.-N., Kandel, A., Schneider, M.: Fuzzy modeling and dynamics (preface). Fuzzy Sets and Systems 106, 1–2 (1999)
46. Yoshida, Y.: Markov chains with a transition possibility measure and fuzzy dynamic programming. Fuzzy Sets and Systems 66, 39–57 (1994)
47. Yoshida, Y. (ed.): Dynamical Aspects in Fuzzy Decision Making. Studies in Fuzziness and Soft Computing, vol. 73. Physica-Verlag, Wurzburg (2001)
48. Yoshida, Y., Yasuda, M., Nakagami, J., Kyrano, M.: A limit theorem in dynamic fuzzy systems with a monotone property. Fuzzy Sets and Systems 94, 109–119 (1998)

Vaguely Motivated Cooperation

Milan Mareš

Abstract. The transferable utility (TU) cooperative games are used as an effective mathematical representation of cooperation and coalitions forming. This contribution deals with a modified form of such games in which the expected pay-offs of coalitions are known only vaguely, where the vagueness is modelled by means of fuzzy quantities and some other fuzzy set theoretical concepts. Such games were investigated in [8] and in some other papers. Their cores and Shapley values were analyzed and some of their basic properties were shown. This contribution is to extend that analysis, namely from the point of view of the motivation of players to cooperate in coalitions, as well as the relation between the willingness to cooperate and the ability to find the conditions under that the cooperation can be percepted as fair.

Keywords: Cooperative game, TU-game, Fuzzy characteristic function, Fuzzy Shapley value, Willingness for cooperation.

1 Introduction

The concepts of coalition and bargaining, introducing the cooperative behaviour of players into games of strategy, appear in the game theory since its very beginning [14] and they form its significant component in many fundamental works (see, e. g., [7] or [15]). The coalition forming is, essentially, based on the expectations of further development of the game. It regards both – the structure of realized coalitions, as well as their presumed incomes. The expectations are mostly rather vague than stochastic, where the vagueness follows mostly from the subjectivity existing in the estimations and evaluations of the acceptability of particular potential results of bargaining. The

Milan Mareš
ÚTIA AV ČR, Pod Vodárenskou věží 4, 182 08 Praha 8, Czech Republic
e-mail: mares@utia.cas.cz

W.A. Lodwick & J. Kacprzyk (Eds.): Fuzzy Optimization, STUDFUZZ 254, pp. 271–283.
springerlink.com

theory of fuzzy set (where the seminal work is [17]) offered the game theory effective tools for mathematical processing of vagueness, mentioned above.

First, the attention was focused on the coalitions forming and its modifications influenced by the vagueness. Fuzzy coalitions defined as fuzzy subsets of the all-players set, allow the parallel participation of some player in several coalitions (see, e. g. [1, 2, 3]) and this model is investigated till now (e. g., [5, 11, 12, 13]). Other demonstration of vagueness in the cooperative games, i. e., the uncertainty regarding the expected incomes of coalitions and its distribution among their members, was investigated rather later. It was briefly mentioned in [8], and more thoroughly analyzed in [9], and recently this model is developed in several other papers.

The aim of this paper is to contribute to the understanding of methodical tools used in [9], and to interpret the conclusions derived there. Our main attention is oriented to the phenomenon of forming cooperation from the point of view of the players' motivation under uncertainty on the expected pay-offs.

The process of negotiation on the eventual cooperation includes two principal periods. Each of them is connected with specific level of motivation, and also its aim reflects different tightness of the accepted agreement.

— The first period covers the stage of a non-cooperative game of strategy, in which the players recognize the advantages of the cooperative actions coordinated with other players. The external attribute of this period is the forming of coalitions based on the knowledge of their expected gains. The classical deterministic game theory has developed the concept of core whose non-emptiness indicates the convenience of the universal coalition of all players.
— The second period includes rather more empathy among all players. The core, following from (in some sense given) expected incomes of coalitions, is very rarely a one-element set. If it is not empty then it offers many potential distributions of the total income of coalition among its members. The choice of the very distribution is the second period of the cooperative negotiations. In distinction from the first period, this choice does not immediately follow from the formal properties of the coalitional pay-offs. It is necessary to combine their values with some, more or less subjective, idea of rightful rates of particular players on the total profit of the coalition. As every player has, in too many cases, his own idea of such justice, there has to be an external authority, either some "judge" not belonging among the players, or some general, commonly acceptable rule, re-distributing the profit. The classical theory of cooperative games with transferable utility offers the Shapley value (see [16, 15, 7]) as such rule possessing acceptable and rational properties of the individual pay-offs.

The periods of negotiation mentioned above are, in the deterministic game model, solved for a long time, already. The theory of fuzzy cooperative games in the form with fuzzy characteristic function analyses these concepts and

methods especially in [9]. Nevertheless, some problems regarding the mutual relation between fuzzy core and fuzzy Shapley value were passed or, at most, only registered without more thorough discussion. In this paper, we aim to contribute to their analysis by methodological comments and several general results dealing with the two periods of negotiation, mentioned above.

The following sections are organized as follows. The next Section 2 summarizes the basic concepts which are dealt in the rest of this contribution, and including the notions of fuzzy core and fuzzy Shapley value. The general conclusions following from this analysis and regarding the motivation of player to the cooperation are presented in Section 3. The last Section 4 includes a conclusive remark.

2 The Models

The basic elements of both models of cooperative game with transferable utility analyzed in this paper – the deterministic one and its fuzzy extension – are briefly recollected in this section. The deterministic case is well known from the classical literature (see, e. g., [7] or [15]), meanwhile its fuzzification was suggested in [9]. The subsection dealing with the deterministic model is completed by the concepts of core and Shapley value. Their fuzzified counterparts are analyzed in the last subsection of this section.

2.1 TU Cooperative Game – Deterministic Case

Let us recollect, first, the fundamental definitions of the model of *cooperative game with transferable utility* (TU-game).

In the whole paper, we denote by R the set of all real numbers.

If M is a set, then we denote in the following sections by $\mathscr{P}(M)$ the set of all subsets of M (the potential set of M).

The TU-game is defined as a pair (I, v), where $I = \{1, 2, \ldots, n\}$ is a non-empty and finite *set of players* and $v : \mathscr{P}(I) \to R$ such that $v(\emptyset) = 0$ is called *characteristic function* of the game.

Every real-valued vector $(x_i)_{i \in I} \in R^n$ such that $x_1 + x_2 + \cdots + x_n \leq v(I)$ is called an *imputation* in the TU-game (I, v). The basic solution concept in such game is the set of imputations $C \subset R^n$ called the *core* of the game and such that

$$C = \left\{ x \in R^n : \sum_{i \in I} x_i \leq v(I) \text{ and for all } K \in \mathscr{P}(I), \sum_{i \in K} x_i \geq v(K) \right\}. \quad (1)$$

It is evident that the coalition I of all players can be effectively formed in a TU-game, only if its core C is non-empty. Moreover, the non-emptiness of core is the single information which the players need to recognize that the coalitional cooperation over complete set I is desirable.

Comment 1. The players need not any exogenous authority to conclude if the cooperation covering complete set I is possible, i. e., if it can be useful for all of them. The non-emptiness of the core follows from the definitoric elements of the game, I and v, and it does demand any other assumption or rule.

Anyhow, the construction of the core C itself does not mean that the negotiations are finished. If $C \neq \emptyset$ then it usually contains more than one imputation and each of the players can have his own idea of the core imputations which is the most righteous one. The structure of the given game itself, i. e., the pair (I, v) and its knowledge, does not guarantee the objective choice of a distribution of the value $v(I)$ among the players, respected and accepted by all of them.

Comment 2. The critical moment of the negotiation is the step from the retrieval of the core and its reduction on one single imputation. In other words, the players themselves are able to recognize the necessity of cooperation but they are not able to agree spontaneously with one single imputation distributing the common profit among them.

Hence, it is inevitable to include an additional element of the game, an arbiter, who decides which distribution of the profit is righteous.

In the practical negotiation, the arbiter can be a person whose authority is confirmed by all agents (players). But it can be an abstract scheme, too, accepted by all players even before the negotiations process. The game theoretical models usually do with a set of principles respected and accepted by the players. These principles were formulated in [16] and the distribution of profits forms a real-valued vector $t = (t_i)_{i \in I}$ called a vector of *Shapley values* (see, e. g., [15], [7], [9]). The principles mentioned above are as follows.

- The values t_i, $i \in I$, do not depend on the ordering of players.
- Vector of Shapley values $(t_i)_{i \in I}$ is to be an imputation, such that

$$\sum_{i=1}^{n} t_i = v(I).$$

- If (I, v_1), (I, v_2) are two TU-games over the set of players I and $(t_i(v_1))_{i \in I}$, $(t_i(v_2))_{i \in I}$ are vector of Shapley values, respectively, if $(I, v_1 + v_2)$ is a TU-game such that for each $K \in \mathscr{P}(I)$,

$$(v_1 + v_2)(K) = v_1(K) + v_2(K),$$

  and if $(t_i(v_1 + v_2))_{i \in I}$ is the vector of Shapley values for $(I, v_1 + v_2)$ then

$$t_i(v_1 + v_2) = t_i(v_1) + t_i(v_2) \quad \text{for all } i \in I.$$

Note that the non-emptiness of Core is not demanded. Shapley (see, [16]) has constructed an effective formula for the evaluation of t_i, $i = 1, \ldots, n$, namely, if for every $K \in \mathscr{P}(I)$, k is the number of players in K, then

$$t_i = \sum_{K \in \mathscr{P}(I)} \frac{(n-k)!(k-1)!}{n!}(v(K) - v(K - \{i\})), \quad i \in I. \tag{2}$$

Let us note that formula (2) defines a vector of Shapley values $(t_i)_{i \in I}$ fulfilling the above conditions even if the core C of the game (I, v) is empty, under the assumption that for any K, $L \subset I$ such that $K \cap L = \emptyset$, the inequality

$$v(K \cup L) \geq v(K) + v(L) \tag{3}$$

holds. Of course, in such case $(t_i)_{i \in I} \notin C$. If $C \neq \emptyset$ then $(t_i)_{i \in I} \in C$,

2.2 *Fuzzy Quantities*

The above deterministic TU-game model is well know, relatively simple, but its correspondence with real cooperative behaviour is rather limited by the latent assumption that the values $v(K)$, $K \subset I$, are deterministic real numbers. Such precise knowledge preceding the proper realization of the game appears too optimistic. This discrepancy can be avoided by using the concepts of fuzzy quantities theory.

In the rest of this paper, if M is a set then we denote by $\mathscr{F}(M)$ the class of all *fuzzy subsets* of M (cf. [17]).

If $A \in \mathscr{F}(M)$ then $\mu_A : M \to [0, 1]$ is the *membership* function of A. Any $a \in \mathscr{F}(R)$ with $\mu_a R \to [0, 1]$ such that

$$\mu_a(x_a) = 1 \quad \text{for at least one } x_a \in R, \tag{4}$$

$$\text{there exist } x_1,\ x_2 \in R \text{ such that } x_1 \leq x_a \leq x_2 \tag{5}$$

$$\text{and } \ \mu_a(x) = 0 \quad \text{for all } x \notin [x_1, x_2],$$

is called a *fuzzy quantity*. Each real number x_a fulfilling (4) is called a *modal value* of a. The set of all fuzzy quantities will be denoted by $\mathscr{F}^*$. As shown, e. g., in [4, 8] and many other works, it is possible to define algebraic operations over $\mathscr{F}^*$, using so called *extension principle*. In this paper, we use two of algebraic operations over fuzzy quantities. Let us consider a, $b \in \mathscr{F}^*$ and $r \in R$, then the *sum* $a \oplus b$ and *crisp product* $r \cdot a$ are fuzzy quantities, too. Their membership functions are

$$\mu_{a \oplus b}(x) = \sup_{y \in R} [\min(\mu_a(y), \mu_b(x - y))], \quad \text{for } x \in R, \tag{6}$$

$$\mu_{r \cdot a}(x) = \mu_a(x/r) \quad \text{if } r \neq 0, \text{ and} \tag{7}$$

$$\mu_{0 \cdot a}(0) = 1, \quad \mu_{0 \cdot a}(x) = 0 \quad \text{for } x \neq 0.$$

There exist numerous approaches to the ordering relation between fuzzy quantities (see, e. g., [6]). Here we use the one of them which is defined as a fuzzy relation $\geq$ with membership function $\nu_{\geq} : \mathscr{F}^* \times \mathscr{F}^* \to [0,1]$, where for $a,\, b \in \mathscr{F}^*$

$$\nu_{\geq}(a,b) = \sup\left[\min\left(\mu_a(x), \mu_b(y)\right) : x,\, y \in R,\, x \geq y\right] \tag{8}$$

is the possibility that $a \geq b$.

2.3 Fuzzy Extension of a TU-Game

As we have mentioned above, we consider here the fuzzification of the characteristic function v. If for every coalition $K \in \mathscr{P}(I)$ there exists a fuzzy quantity $w(K) \in \mathscr{F}^*$ such that $v(K)$ is a modal value of $w(K)$, then we say that the pair (I,w) is a *fuzzy extension* of the TU-game (I,v), and we call w the *fuzzy characteristic function* of (I,w).

It is not difficult to define the set of *fuzzy imputations* in (I,w) as a fuzzy subset $\mathscr{V}$ of R^n with membership function $\mu_{\mathscr{V}} : R^n \to [0,1]$, where

$$\mu_{\mathscr{V}}\left((x_1,\ldots,x_n)\right) = \nu_{\geq}\left(w(I), \sum_{i=1}^{n} x_i\right), \quad (x_1,\ldots,x_n) \in R^n, \tag{9}$$

where (8) was used.

Similarly, the *fuzzy core* of (I,w) is a fuzzy subset of R^n (see [9]) denoted by C_w and with membership function $\mu_C : R^m \to [0,1]$, where for any $\boldsymbol{x} = (x_1,\ldots,x_n) \in R^n$

$$\mu_C(\boldsymbol{x}) = \min\left[\mu_{\mathscr{V}}(\boldsymbol{x}), \min\left(\mu_{\geq}\left(\sum_{i\in K} x_i, w(K)\right) : K \in \mathscr{P}(I)\right).\right] \tag{10}$$

Remark 1. Previous definition, together with (8) immediately mean that if $\mathscr{V}$ is a fuzzy imputation and C_w is a core of a fuzzy extension (I,w) of some TU-game (I,v), and if $\alpha \geq \beta$, $\alpha,\, \beta \in [0,1]$ then

$$\left\{\boldsymbol{x} \in R^n : \mu_{\mathscr{V}}(\boldsymbol{x}) = \alpha\right\} \subset \left\{\boldsymbol{x} \in R^n : \mu_{\mathscr{V}}(\boldsymbol{x}) = \beta\right\},$$

and

$$\left\{\boldsymbol{x} \in R^n : \mu_C(\boldsymbol{x}) = \alpha\right\} \subset \left\{\boldsymbol{x} \in R^n : \mu_C(\boldsymbol{x}) = \beta\right\}.$$

Comment 3. Analogously to the deterministic concept of imputation as an accessible distribution of profit among all players, the fuzzy imputation represents the accessibility of profit distribution structured by uncertainty connected with particular incomes expected by the coalition of all players I.

Comment 4. Similarly, the fuzzy core represents the distributions of global profit among all players, which cannot be effectively protested by any coalition, and which is structured by uncertainty connected with particular incomes expected by the coalition of all players.

There exists one concept, more, whose deterministic form we know and which turns into its fuzzy counterpart if a fuzzy extension of a TU-game is considered. Namely, the Shapley value (2). Its fuzzification can be constructed in two ways.

The first way consists in the passive application of formula (2) where fuzzy quantities $w(K)$, $K \in \mathscr{P}(I)$ are used instead of the crisp values $v(K)$. This method was analyzed in [9] and it is evident that it results into fuzzy quantities for every Shapley value. More precisely, let $i \in I$ and let us numerate the coalitions from $\mathscr{P}(I)$ as

$$\{K_0, K_1, K_2, \ldots, K_N\}, \quad \text{where } K_0 = \emptyset \text{ and } N = 2^n - 1. \tag{11}$$

Then the fuzzy quantity $s_i \in \mathscr{F}^*$ with $\mu_{s(i)}$, defined by

$$\begin{aligned} s_i = \frac{(n-k_1)!(k_1-1)!}{n!}\,(w(K_1) \oplus (-1 \cdot w(K_1 - \{i\}))) \oplus \cdots \\ \cdots \oplus \frac{(n-k_N)!(k_N-1)!}{n!}\,(w(K_N) + (-1 \cdot w(K_N - \{i\}))) \end{aligned} \tag{12}$$

can be considered for the i-th component of the vector of *fuzzy Shapley values*. Here, k_j is the number of members of the coalition K_j and all k_j, $j = 1, \ldots, N$, are crisp numbers. It means that operations used in (12) are fully characterized by (6) and (7) and their properties are analyzed, e. g., in [8] and [4] and recollected also in [9]. This method, however lucid it is, displays one significant discrepancy. Namely, if $i \notin K$ for some $K \in \mathscr{P}(I)$ then in the deterministic case $v(K) - v(K - \{i\}) = 0$, and formula (2) deals with coalitions including i, only. As shown in [8], this conclusion is not correct in the case of fuzzy extension (I, w) of (I, v). If $i \notin K$ then $w(K) \oplus (-1 \cdot w(K - \{i\})) = w(K) \oplus (-1 \cdot w(K)) = a(i, K)$, where $a(i, K)$ is a fuzzy quantity from $\mathscr{F}^*$ with at least one modal value equal to 0,

$$\mu_{a(i,K)}(0) = 1,$$

and with symmetric membership function where

$$\mu_{a(i,K)}(x) = \mu_{a(i,K)}(-x) \quad \text{for all } x \in R.$$

Usually, except very special cases with degenerated fuzziness, $a(i, K)$ used in (12) extends the uncertainty of the resulting fuzzy value $s_i \in \mathscr{F}^*$ and in this sense it influences the stability of eventually achieved results of negotiation. Namely, it symmetrically increases the extent of uncertainty connected with s_i.

This, in certain sense formal, discrepancy can be avoided if we limit the summation in (12) on the coalitions from $\mathscr{P}(I)$ for which $i \in K$. More formally, we may define a fuzzy quantity $q_i \in \mathscr{F}^*$ with membership function $\mu_{q(i)} : R \to [0,1]$ by means of modified (12)

$$q_i = \frac{(n-k_1)!(k_1-1)!}{n!} \cdot \sigma(i,k_1) \cdot (w(K_1) \oplus (-1 \cdot w(K_1 - \{i\}))) \oplus \cdots \qquad (13)$$
$$\cdots \oplus \frac{(n-k_1 N)!(k_N-1)!}{n!} \cdot \sigma(i,k_N) \cdot (w(K_N) \oplus (-1 \cdot w(K_N - \{i\}))),$$

where

$$\sigma(i,K_j) = 1 \quad \text{iff } i \in K_j, \qquad \sigma(i,K_j) = 0 \quad \text{iff } i \notin K_j,\ j = 1,\ldots,N$$

and where the notations used in (12) are preserved.

The second way of constructing *fuzzy Shapley value* is based on the general extension principle, as well. For every $i \in I$ and for every K_j, $j = 0,1,\ldots,N$, we denote $\mu_j : R \to [0,1]$ the membership function of fuzzy quantity $w(K_j)$. Then we define fuzzy quantity $u_i \in \mathscr{F}^*$ with membership function $\mu_{u(i)} : R \to [0,1]$ by means of

$$\mu_{u(i)}(x) = \sup \Big[\min\left(\mu_1(y_1), \mu_2(y_2), \ldots, \mu_n(y_N)\right) : \qquad (14)$$
$$y_1,\ldots,y_N \in R,\ x = \sum_{j=1,\ldots,N} \frac{(n-k_j)!(k_j-1)!}{n!} \left(y_j - y_{\ell_j}\right) \Big],$$

where n and k_j are interpreted in agreement with (12) and for every $j = 1,\ldots,N$, $K_{\ell_j} - \{i\}$.

The fuzzy number u_i is the i-th component of *fuzzy Shapley value* of (I,w).

Remark 2. It is easy to see that for $K_j \in \mathscr{P}(I)$ and $i \in I$ such that $i \notin K$ then $y_j - y_{\ell_j}$ and, consequently, the relevant element of the sum in (13) vanishes.

Lemma 1. Let (I,w) be fuzzy extension of a TU-game (I,v), let for any $i \in I$, $s_i \in \mathscr{F}^*$ be defined by (12), $q_i \in \mathscr{F}^*$ be defined by (13), and let $u_i \in \mathscr{F}^*$ be defined by (14). Then

$$\mu_{u(i)}(x) = \mu_{q(i)}(x) \quad \text{for all } i \in I,\ x \in R$$

and there exist b, $d \in \mathscr{F}^*$, such that

$$\mu_b(0) = \mu_d(0) = 1, \qquad \mu_b(x) = \mu_b(-x), \qquad \mu_d(x) = \mu_d(-x), \qquad x \in R,$$

and

$$u_i \oplus b = s_i \oplus d.$$

Proof. The first statement,

$$\mu_{u(i)}(x) = \mu_{q(i)}(x)$$

for all $i \in I$ and $x \in R$ follows from (13) and (14), immediately, as for any $K_j \in \mathscr{P}(I)$ such that $i \notin K_j$, $\sigma(K_j, i) = 0$ and $\mu_j(0) = 1$. In this sense, the values of μ_j do not influence the value of $\mu_{u_i}(x)$.

The second statement, namely the additive equivalence of u_i and s_i in the sense of [8], follows from (13) and (12). Namely, $s_i = q_i \oplus b$, where $b \in \mathscr{F}^*$, and b is the sum of fuzzy quantities

$$\frac{(n-k_j)!(k_j-1)!}{n!}\ (w(K_j) \oplus (-1 \cdot w(K_j)))$$

for those K_j for which $i \notin K_j$. Then each of such fuzzy quantities is symmetric, i. e., they fulfil the properties formulated in the proved statement. It means that their sum b is symmetric, as well (see [8]), and the second statement is proven. □

If (I, v) is a TU-game and (I, w) its fuzzy extension, then the fundamental fuzzy solution concepts of (I, w) are fuzzy extensions of their crisp counterparts in (I, v). It is not difficult to formulate this conclusion by means of the following statements.

Theorem 1. Let (I, v) be a TU-game and (I, w) its fuzzy extension. If C and C_w are the core of (I, v) and fuzzy core of (I, w), respectively, then C_w is fuzzy extension of C. It means that for any $\boldsymbol{x} \in R^n$

$$\mu_C(\boldsymbol{x}) = 1 \quad \text{iff } \boldsymbol{x} \in C.$$

Proof. Let $\boldsymbol{x} = (x_1, x_2, \ldots, x_n) \in C$. Then

$$\begin{aligned} &\nu_{\geq}\left(\sum_{i\in K} x_i, w(K)\right) = 1 \quad \text{for all } K \in \mathscr{P}(I), \\ &\text{and} \quad \nu_{\geq}\left(w(I), \sum_{i=1}^{n} x_i\right) = 1 \end{aligned} \tag{15}$$

as follows from (8) and from the fact that each fuzzy imputation is a fuzzy extension from some crisp imputation. Consequently, $\mu_C(\boldsymbol{x}) = 1$.

Let, on the other hand, $\mu_C(\boldsymbol{x}) = 1$. Then, due to (10), all membership values in (15) are necessarily equal to 1, which immediately implies that $\boldsymbol{x} \in C$. □

Theorem 2. Let (I, v) be a TU-game and (I, w) its fuzzy extension. Let $\boldsymbol{t} = (t_1, t_2, \ldots, t_n) \in R^n$ be (crisp) Shapley value of (I, v) defined by (2). Then the vector of fuzzy quantities $\boldsymbol{s} = (s_i)_{i \in I}$ defined by (12), vector of fuzzy numbers $\boldsymbol{u} = (u_i)_{i \in I}$ defined by (14), and vector of fuzzy quantities $\boldsymbol{q} = (q_i)_{i \in I}$ defined by (15) are vectors of fuzzy extensions of t_i, $i = 1, 2, \ldots, n$, respectively.

Proof. Formulas (2) and (14) immediately imply that $u_i(t_i) = 1$ for all $i \in I$ and, consequently, fuzzy quantity u_i is a fuzzy extension of t_i for all $i = 1, 2, \ldots, n$. The first statements of Lemma 1 means that the above result is true for fuzzy quantities q_i and crisp Shapley values t_i, $i \in I$, as well. Moreover, formula

$$u_i \oplus b = s_i \oplus d$$

where $\mu_b(0) = \mu_d(0) = 1$, used in the second statement of Lemma 1, implies that if some $x \in R$ is a modal value of u_i then it is a modal value of s_i, as well (see (6)). Hence, t_i is a modal value of s_i for all $i \in I$, and the statement of the theorem is true. □

Remark 3. The relevant definitions immediately imply that if for all $j = 0, 1, \ldots, N$

$$\mu_j(v(k_j)) = 1, \qquad \mu_j(x) = 0 \quad \text{for all } x \neq v(K_j),$$

then C_w is identical with the crisp core C of (I, v), and all fuzzy Shapley values $(u_i)_{i \in I}$, $(q_j)_{i \in I}$ and $(s_i)_{i \in I}$ are equal and identical with crisp Shapley values $(t_i)_{i \in I}$ of (I, v).

3 Vague Willingness to Cooperation

After introducing or remembering the main concepts of interest, i. e., the fuzzy extensions of cooperative game with transferable utility, its core and Shapley value, we aim to transfer the ideas of crisp cooperation model into their fuzzy counterparts. Let us summarize the fundamental knowledge achieved in the deterministic theory:

- The non-emptiness of the core suffices to the recognition that rational cooperation is the optimal behaviour of players.
- But it does not suffice to identify which cooperation (partition of the common profit of I) is the rational one. Identification of this rationality demands the acceptance of an additional rule, the value (Shapley value) of the game.

The general properties of the value are summarized in subsection 2.1.

The main purpose of this section is to discuss the validity of the previous, rather methodological, consequences for the case of the fuzzy extension of a TU-game.

The vagueness of the expected pay-offs, i. e., the substitution of the crisp numbers $v(K)$ by fuzzy quantities $w(K)$ for all $K \in \mathscr{P}(I)$, may appear like a degradation of the conditions under which the players form the coalitions – our traditional thinking percepts the uncertainty as a discrepancy, in general. In fact, the existence of vagueness in the expectation of coalitional profits, enriches the analysis of the game and its structure. The fuzzy extension (I, w) of (I, v), with many levels of possibilities regarding the players' expectations, is not only much more realistic but also much more effective in the process of forming the most rational cooperative behaviour.

The fuzzy core is a lucid demonstration of the above rule. If (I, v) and (I, w) are a deterministic TU-game and its fuzzy extension, respectively, then the core C and fuzzy core C_w respect analogous relation. The fuzzy core C_w does not grind the crisp willingness for cooperation based on the deterministic expectations of profit but, on the contrary, it enlarges the potential possibilities of agreement by the (usually quite wide) class of not completely sure but possible variants. If we consider the fact that the "fully deterministic" expectations of profit made before the realization of the game cannot be as doubtless as they appear to be the fuzzy extension of the cooperation model is more realistic (and more precise) than the crisp one.

More formally, let us consider the fuzzy core C_w of a fuzzy extension (I, w) of TU-game (I, v). Then we may define the number

$$m_C = \sup\left(\mu_C(\boldsymbol{x}) : \boldsymbol{x} \in R^n\right), \tag{1}$$

which we call the *cooperative potential* of (I, w).

Remark 4. Evidently, $0 \leq m_C \leq 1$, and if (I, w), (I, w') are two fuzzy extensions of (I, v) such that $\mu_w(\boldsymbol{x}) \geq \mu_{w'}(\boldsymbol{x})$ for all $\boldsymbol{x} \in R^n$ and if m_C, m'_C are their cooperative potentials, respectively, then $m_C \geq m'_C$.

The cooperative potential can be accepted for the measure of ability of the players in I to accept the global all-players' coalition. The previous remark stresses the obvious fact that the more the fuzzy extension of (I, w) differs from its crisp base, the higher is the possibility that the players find a common agreement.

The question to be answered about the fuzzy Shapley value is rather different. Namely, it is important to know if, and in which way, the fuzzy Shapley value respects the general demands on values, formulated in subsection 2.1.

Here, we focus our attention on the fuzzy Shapley values defined by (14) and denoted by $u_i \in \mathscr{F}^*$. Due to Lemma 1, we know that its properties are identical with the properties of $q_i \in \mathscr{F}^*$ (defined by (13)) and in some sense equivalent with the properties of $s_i \in \mathscr{F}^*$. This is valid for all $i \in I$.

Remark 5. As follows from (14), immediately, the membership functions $\mu_{u(i)}$ for $i = 1, 2, \ldots, n$, are independent on the ordering of their computation.

Lemma 2. The modal values $t_1, t_2, \ldots, t_n$, $v(I)$ of the fuzzy quantities $u_1, u_2, \ldots, u_n$, $w(I)$ fulfil the equality

$$\sum_{i \in I} t_i = v(I).$$

Proof. The statement follows from (14) and (2), immediately. □

Lemma 3. Let (I, v_1) and (I, v_2) be TU-games and (I, w_1), (I, w_2) be their fuzzy extensions, respectively. Let $(I, w_1 \oplus w_2)$ be a fuzzy cooperative game such that for every $K \in \mathscr{P}(I)$

$$(w_1 \oplus w_2)(K) = w_1(K) \oplus w_2(K),$$

and, finally, let $(I, v_1 + v_2)$ be a TU-game such that for every $K \in \mathscr{P}(K)$

$$(v_1 + v_2)(K) = v_1(K) + v_2(K).$$

Then $(I, w_1 + w_2)$ is a fuzzy extension of $(I, w_1 \oplus w_2)$, modal values $(v_1 + v_2)(K)$ of $(w_1 \oplus w_2)(K)$, $K \in \mathscr{P}(I)$.

Proof. The statement follows from the definition of fuzzy extension of TU-game, and from the assumptions of this lemma. □

Lemma 4. Under the assumptions of Lemma 3 let us denote by t_i, $i = 1, 2, \ldots, n$, the Shapley values of (I, v). Then for every $i \in I$, t_i is a modal value of $u_i \in \mathscr{F}^*$.

Proof. The statement follows from (13) and (2), immediately. For every $K \in \mathscr{P}(I)$, $\mu_w(v(K)) = 1$ and, consequently, $\mu_{u(i)}(t_i) = 1$, $i = 1, 2, \ldots, n$. □

Note that the validity of the general properties of fuzzy Shapley values is in more detailed way investigated in [9], Chapter 9.

Comment 5. The vagueness included in the concept of fuzzy extension of a TU-game influences also the validity of the general principles connected with the concept of the value, especially of the Shapley value. Relative generality of the fuzzy characteristic function implies also rather free formal structure of fuzzy Shapley value, and the guaranteed fulfillment of the basic properties of value for the modal values of its fuzzified form, only.

4 Conclusive Remark

The previous brief analysis of the fuzzified *TU*-games where the fuzzification regards the characteristic function and concepts derived from it, allows to formulate the following heuristic conclusion.

The vagueness of expected pay-offs, which is natural, rather influences the formal properties of the core and Shapley value, but it does not violate their very important functions – namely, to indicate the motivation of player for cooperation, and to show an acceptable distribution of the common profit among cooperating players. Or course, the vagueness of expectations causes certain vagueness of the concepts of core and Shapley value, but this vagueness does not limit the information hidden in the core, and it rather modifies than limits similar information in the Shapley value.

In other words the fuzzification of pay-offs, in principle, does not significantly influence the ability and willingness of the players to cooperate.

References

1. Aubin, J.P.: Fuzzy Games. Report MRC–1480. Math. Res. Centre Univ. of Wisconsin (1973)
2. Butnariu, D.: Solution concepts for n-person fuzzy games. In: Gupta, M.M., Radage, R.K., Yager, R.R. (eds.) Advances in Fuzzy Set Theory and Applications. North-Holland, Amsterdam (1979)
3. Butnariu, D., Klement, E.P.: Triangular Norm–Based Measures and Games With Fuzzy Coalitions. Kluwer, Dordrecht (1993)
4. Dubois, D., Kerre, E.E., Mesiar, R., Prade, H.: Fuzzy interval analysis. In: Fundamentals of Fuzzy Sets, pp. 483–581. Kluwer, Dordrecht (2000)
5. Grabish, M., Nguyen, H.T., Walker, E.A.: Fundamentals of Uncertainty Calculi With Applications to Fuzzy Inference. Kluwer, Dordrecht (1995)
6. Kerre, E.E., Wang, X.: Reasonable properties for the ordering of fuzzy quantities. Part I, Part II. Fuzzy Sets and Systems 118, 373–385, 387–405 (2001)
7. Luce, R.D., Raiffa, H.: Games and Decisions. J. Wiley and Sons, London (1957)
8. Mareš, M.: Computations Over Fuzzy Quantities. CRC-Press, Boca Raton (1994)
9. Mareš, M.: Fuzzy Cooperative Games. Physica-Verlag, Heidelberg (2001)
10. Mareš, M.: Fuzzy components of cooperative market. In: Batyrshin, I., Kacprzyk, J. (eds.) Perception Based Data Mining and Decision Making in Economics and Finance, pp. 209–239. Springer, Heidelberg (2007)
11. Mareš, M., Vlach, M.: Fuzzy coalitions as fuzzy classes of crisp coalitions. In: Transactions of 7th Japan–Czech Seminar on Data Analysis under Uncertainty, pp. 32–34. Awaji Yumebutai Center, Awaji (2004)
12. Mareš, M., Vlach, M.: Alternative representation of fuzzy coalitions. Int. J. of Uncertainty, Fuzziness and Knowledge-Based Systems (in print)
13. Mareš, M., Vlach, M.: Disjointness of fuzzy coalitions. Kybernetika 44(3), 416–429 (2008)
14. von Neumann, J., Morgenstern, O.: Theory of Games and Economic Behaviour. Princeton Univ. Press, Princeton (1953)
15. Rosenműller: The Theory of Games and Markets. North Holland, Amsterdam (1982)
16. Shapley, L.S.: The value for n-prson games. In: Contributions to the Theory of Games II. Annals of Math. Studies, vol. 28, pp. 307–317 (1953)
17. Zadeh, L.A.: Fuzzy sets. Information and Control 8(3), 338–353 (1965)

Part IV
Fuzzy Network and Combinatorial Optimization

Computing Min-Max Regret Solutions in Possibilistic Combinatorial Optimization Problems

Adam Kasperski and Paweł Zieliński

Abstract. In this chapter we discuss a wide class of combinatorial optimization problems with a linear sum and a bottleneck cost function. We first investigate the case when the weights in the problem are modeled as closed intervals. We show how the notion of optimality can be extended by using a concept of a deviation interval. In order to choose a solution we adopt a robust approach. We seek a solution that minimizes the maximal regret, that is the maximal deviation from optimum over all weight realizations, called scenarios, which may occur. We then explore the case in which the weights are specified as fuzzy intervals. We show that under fuzzy weights the problem has an interpretation consistent with possibility theory. Namely, fuzzy weights induce a possibility distribution over the scenario set and the possibility and necessity measures can be used to extend the optimality evaluation and the min-max regret approach.

1 Introduction

In many optimization problems we seek an object composed of elements of a given set to achieve some goal. For instance, in a wide class of network problems the element set consists of all edges of a given graph and we seek an optimal path, spanning tree, cut, matching etc. in this graph. A comprehensive review of various problems of this type can be found in [1, 30, 35]. While describing a particular system we often meet some parameters associated with the elements whose values are not

Adam Kasperski
Wrocław University of Technology, Institute of Industrial Engineering and Management, Wybrzeże Wyspiańskiego 27, 50-370, Wrocław, Poland
e-mail: adam.kasperski@pwr.wroc.pl

Paweł Zieliński
Wrocław University of Technology, Institute of Mathematics and Computer Science, Wybrzeże Wyspiańskiego 27, 50-370, Wrocław, Poland
e-mail: pawel.zielinski@pwr.wroc.pl

W.A. Lodwick & J. Kacprzyk (Eds.): Fuzzy Optimization, STUDFUZZ 254, pp. 287–312.
springerlink.com

precisely known. For instance, in a traffic network the traveling times between distinct points are rarely known in advance and this uncertainty must be taken into account while choosing a path in this network. In order to model the risk connected with imprecise parameters a stochastic approach can be adopted (see. e.g. [21]). For every parameter a probability distribution for its values is specified and, typically, the expected cost of a solution is minimized. The stochastic approach has several drawbacks. Namely, it may be hard or expensive to estimate the probability distribution for an unknown parameter. Also, the obtained solution may be not reasonable if it is used only once because it may be poor under the worst parameter realization which may occur.

An approach which has received an increasing attention in the recent years is the one of *robust optimization*. The idea of robust approach is to find a solution that hedges against the worst realizations of parameters which may occur. A good introduction to robust optimization can be found in a book [29]. For this class of problems a part of the input is a *scenario set*, which contains all realizations of the parameters, called *scenarios*, which may occur. No probability distribution over the scenario set is specified. Then a solution is computed, which minimizes a given criterion under the worst scenario. One of the most popular methods of defining the scenario set is to specify for every parameter a closed interval, which contains all its possible values. The scenario set is then the Cartesian product of all the uncertainty intervals. In order to choose a solution a *maximal regret* criterion can be used. The maximal regret is the maximal difference between the cost of a solution and the optimum over all scenarios. It was first suggested as a criterion for choosing a decision under uncertainty by Savage [39]. A deep discussion on the maximal regret can also be found in a book [31].

The min-max regret approach to combinatorial optimization problems with interval data has attracted a considerable attention recently. A recent survey of the known results in this area can be found in [2, 22]. It turns out that the complexity of the min-max regret problem strongly depends on the choice of the cost function in its deterministic version. Under a *bottleneck cost* the min-max regret problem is polynomially solvable if only the deterministic problem is polynomially solvable [7]. However, under a more popular *linear sum cost*, the min-max regret versions of the shortest path [7, 24, 42], the minimum spanning tree [6, 7], the minimum assignment [3] and the minimum $s-t$ cut [4] turned out to be NP-hard. A polynomial algorithm is known for the min-max regret selecting items problem [8, 12], which is a special case of the 0-1 Knapsack with unit capacities of all items. Some approximation algorithms for the class of min-max regret problems with the linear sum cost can be found in [23, 26].

In this chapter we show how the known min-max regret approach can be extended. The key idea is to model the imprecision using *fuzzy intervals*. A fuzzy interval can be seen as a monotone, under inclusion, family of closed intervals parametrized by the value of $\lambda \in [0,1]$. It is also a fuzzy set in the space of reals, whose membership function is a *possibility distribution* for the values of an unknown quantity. A description of possibility theory can be found in a book [14], where one can also find some methods of obtaining possibility distributions from

the possessed knowledge. Fuzzy intervals allow us to define a possibility distribution over the scenario set. So now, for every scenario we can assign a real number from the interval [0,1], which says us what is the possibility that this scenario will occur. In order to choose a solution we can adopt an elegant concept proposed for fuzzy linear programming in [19, 20]. It turns out that this solution method can be viewed as a direct extension of the min-max regret approach to the fuzzy case, which additionally has a clear possibilistic interpretation. Furthermore, the fuzzy combinatorial optimization problems are easier to solve than fuzzy linear programming described in [20].

This chapter is organized as follows. First, in Section 2, we recall a formulation of a combinatorial optimization problem with deterministic weights. We describe the problems with two types of cost functions, namely the bottleneck and the linear sum ones. We also introduce the concept of a deviation, which is a distance of a solution (element) from optimality. The concept of deviation will play a central role in our analysis. In Section 3, we discuss the combinatorial optimization problems with interval weights. By extending the concept of deviation we show how the optimality of solutions and elements can be characterized and how to choose a solution. We seek a solution that minimizes the maximal regret, that is the largest deviation which may occur for this solution. We present all known complexity results for the interval problems. In Section 4, we investigate the combinatorial optimization problems with fuzzy weights. We first recall some basic notions of possibility theory. We then show how the concept of scenario set can be extended by defining a possibility distribution over all scenarios. We also introduce the concept of a fuzzy deviation and show how to characterize the optimality of solutions and elements, using possibility and necessity measures. Finally, we adopt a method of choosing a solution under fuzzy weights and we construct several methods of computing this solution.

2 Deterministic Combinatorial Optimization Problems

In this section we briefly recall a formulation of a general combinatorial optimization problem. Let $E = \{e_1, \ldots, e_n\}$ be a finite set of elements and let $\Phi \subseteq 2^E$ be a set of subsets of E called a set of *feasible solutions*. For every element $e \in E$ there is a nonnegative weight w_e, which expresses a single parameter associated with e such as cost, time, length etc. We will use $F(X)$ to denote a cost of solution $X \in \Phi$. Two types of the cost function are widely used, namely a *linear sum cost* $F(X) = \sum_{e \in X} w_e$ and a *bottleneck cost* $F(X) = \max_{e \in X} w_e$. The deterministic combinatorial optimization problem P is the following one:

$$\text{P:} \quad \min_{X \in \Phi} F(X), \tag{1}$$

where $F(X)$ is either the linear sum or the bottleneck cost. So, an instance of the problem is specified by a triple $(E, \Phi, \mathbf{w})$, where $\mathbf{w}$ is a vector of element weights.

The formulation (1) encompasses a large variety of problems. In the important class of *network problems*, E is a set of edges of a given directed or undirected

graph $G = (V, E)$ and Φ consists of all subsets of the edges that form some objects in G such as paths, spanning trees, matchings, cuts etc. In general (1) includes the problems, which can be formulated as 0-1 programming ones. To see this, we need to associate a binary variable $x_i \in \{0, 1\}$ with every element $e_i \in E$ and describe Φ using a system of constraints involving the binary variables. Notice that some of the problems are polynomially solvable while the other ones are NP-hard. In this chapter we will assume that P is polynomially solvable. A description of such problems with both linear sum and bottleneck cost can be found for instance in books [30, 35] and in papers [10, 17, 36, 37].

In theory and practice the class of *matroidal problems* is of great importance. Recall that a *matroid* is a pair $(E, \mathscr{I})$, where E is a nonempty element set and $\mathscr{I}$ is a set of subsets of E such that $\mathscr{I}$ is closed under inclusion (if $A \in \mathscr{I}$ and $B \subseteq A$ then $B \in \mathscr{I}$) and fulfills the so-called *growth property* (if $A, B \in \mathscr{I}$ and $|A| < |B|$ then there is $e \in B \setminus A$ such that $A \cup \{e\} \in \mathscr{I}$). The maximal under inclusion elements in $\mathscr{I}$ are called *bases*. In a matroidal problem the set of feasible solutions Φ consists of all bases of a given matroid. Perhaps, the best known example of a matroidal problem is the minimum spanning tree, where E is a set of edges of a given undirected graph and $\mathscr{I}$ consists of all subsets of the edges that form acyclic subgraphs of G. Then $(E, \mathscr{I})$ is called a *graphic matroid* and its base is a spanning tree of G, so Φ contains all spanning trees of G. Another important example is the minimum selecting items problem. In this problem, E is a set of items and $\mathscr{I}$ consists of all subsets of E, whose cardinalities are less than or equal to a given number p. The system $(E, \mathscr{I})$ is the so-called *uniform matroid* and X is a base of this matroid if and only if $|X| = p$. In this case Φ contains all subsets of E, whose cardinalities are precisely p. We will see in the next sections that the particular structure of matroidal problems sometimes allows us to design efficient algorithms under uncertainty.

In the approach presented in this chapter a central role will be played by the concept of a deviation. A *deviation of solution* $X \in \Phi$ is defined as follows:

$$\delta_X = F(X) - \min_{Y \in \Phi} F(Y).$$

Hence deviation δ_X expresses a "distance" of X from optimum. Obviously X is optimal if and only if $\delta_X = 0$. A similar concept can be introduced for elements. Let $\Phi_f \subseteq \Phi$ be the set of all feasible solutions that contain element f. Then a *deviation of element* $f \in E$ is defined as follows:

$$\delta_f = \min_{Y \in \Phi_f} F(Y) - \min_{Y \in \Phi} F(Y).$$

We call element f *optimal* if $\delta_f = 0$. In other words, f is optimal if and only if it is a part of an optimal solution. The solution (element) deviation gives us an information how far from optimality a solution (element) is.

3 Combinatorial Optimization Problems with Interval Weights

In practice, precise values of the element weights in a combinatorial optimization problem may be unknown. In this section we discuss perhaps the simplest uncertainty representation, where for every unknown weight a closed interval containing all its possible values is specified. We extend the concept of deviation and we show how the optimality of a given solution or element can be characterized and how to choose a solution under interval weights.

3.1 Scenario Set

Assume that we only know that the value of the weight w_e of element $e \in E$ will fall within a closed interval $W_e = [\underline{w}_e, \overline{w}_e]$. Notice that a precise weight w_e can be modeled as a *degenerate* interval such that $\underline{w}_e = \overline{w}_e$. We assume that there is no probability distribution in W_e, $e \in E$, and all weights are unrelated, that is the value of every weight does not depend on the values of the remaining weights. A vector $S = (s_e)_{e \in E}$ such that $s_e \in W_e$ for all $e \in E$ is called a *scenario* and it represents the state of the world in which $w_e = s_e$ for all $e \in E$. A *scenario set* Γ is formed by the Cartesian product of all the uncertainty intervals, namely $\Gamma = \times_{e \in E} W_e$. Notice that our assumptions imply that for any two scenarios S_1 and S_2 it is not possible to say which one is more likely to happen. In other words, there is no probability distribution in scenario set Γ.

Among the scenarios an important role is played by the *extreme* ones, where all weights take the lower or upper bounds in their uncertainty intervals, i.e. the scenarios from the set $\times_{e \in E} \{\underline{w}_e, \overline{w}_e\}$. Let $A \subseteq E$ be a subset of the elements. In scenario S_A^+ all elements $e \in A$ have weights $\overline{w}_e$ and all the remaining elements have weights $\underline{w}_e$. In the symmetric scenario S_A^- all elements $e \in A$ have weights $\underline{w}_e$ and all the remaining elements have weights $\overline{w}_e$.

Under the interval uncertainty representation the cost of solution X depends on scenario $S \in \Gamma$ an we will denote it as $F(X,S)$. Of course, $F(X,S)$ is either the linear sum cost $\sum_{e \in X} s_e$ or the bottleneck cost $\max_{e \in X} s_e$. We will use $F^*(S)$ to denote the cost of an optimal solution under scenario S. In order to obtain $F^*(S)$ we must solve the deterministic problem (II) under the weight realization specified by scenario S. Now the solution and element deviations also depend on scenario S and we will denote them as $\delta_X(S)$ and $\delta_f(S)$, respectively.

3.2 Deviation Interval and Optimality Evaluation

Recall that in the deterministic case a deviation gives a full characterization of optimality. In the interval case the optimality can be fully characterized by the so-called *deviation interval*. For a given solution $X \in \Phi$ we define $\Delta_X = [\underline{\delta}_X, \overline{\delta}_X]$,

where $\underline{\delta}_X = \min_{S\in\Gamma} \delta_X(S)$ and $\overline{\delta}_X = \max_{S\in\Gamma} \delta_X(S)$. The quantity $\overline{\delta}_X$ is called in literature the *maximal regret* of X [29] and it expresses the largest distance of X from optimality. Similarly, for a given element $f \in E$ we have $\Delta_f = [\underline{\delta}_f, \overline{\delta}_f]$, where $\underline{\delta}_f = \min_{S\in\Gamma} \delta_f(S)$ and $\underline{\delta}_f = \max_{S\in\Gamma} \delta_f(S)$.

The intervals Δ_X and Δ_f contain all values of solution and element deviations which may occur and allow us to give the following optimality characterization. We say that a solution X is *possibly optimal* if $\underline{\delta}_X = 0$ and it is *necessarily optimal* if $\overline{\delta}_X = 0$. Obviously, solution X is possibly optimal if and only if it is optimal under some scenario $S \in \Gamma$ and it is necessarily optimal if and only if it is optimal under all scenarios $S \in \Gamma$. Exactly the same optimality characterization can be given for the elements. So, we can also introduce the possibly and necessarily optimal elements using deviation intervals of elements. It is easy to check that every possibly (necessarily) optimal solution is composed of possibly (necessarily) optimal elements. However, the converse statement is not true since it is not difficult to give an example of a solution composed of possibly (necessarily) optimal elements, which is not possibly (necessarily) optimal (see [25]).

3.3 Choosing a Solution under Interval Weights

Now an important question arises which solution should be chosen under interval weights. One can simply choose a possibly optimal one. This can be done by computing an optimal solution under any particular scenario $S \in \Gamma$. This choice is optimistic because we need to assume that a good scenario will occur. However, the quality of the solution may be very poor if a bad scenario will realize. One can also try to compute a necessarily optimal solution. Indeed, such a solution is an ideal choice but, contrary to the possibly optimal solutions, it rarely exists. In other words, the necessary optimality is too strong criterion. We thus can see that in order to choose a solution, a compromise between the possible and necessary optimality is required. This compromise is achieved by computing a solution that minimizes the maximal regret $\overline{\delta}_X$, that is the largest deviation (a distance to optimality) over all scenarios. So, under the interval uncertainty representation we focus on the following optimization problem:

$$\min_{X\in\Phi} \overline{\delta}_X. \tag{2}$$

An optimal solution to (2) is called an *optimal min-max regret solution*. We get immediately that every necessarily optimal solution is an optimal min-max regret one (but the converse statement is not true). In the next two sections we will show that every optimal min-max regret solution is possibly optimal. Hence it fulfills the minimum requirement of being optimal under some scenario. In consequence, the deviation interval of an optimal min-max regret solution is of the form $[0, \overline{\delta}_X]$, where $\overline{\delta}_X$ is the smallest among all $X \in \Phi$.

3.4 Computational Properties of the Interval-Valued Problem

In this section we focus on the computational properties of problem P with interval weights. We will show that the complexity of computing deviation intervals and min-max regret solutions strongly depends on the choice of the cost function.

3.4.1 Problems with Linear Sum Cost

In this section we discuss the case when $F(X,S) = \sum_{e \in X} s_e$, so we consider a problem with the linear sum cost function. The following proposition results directly from the definition of the cost function:

Proposition 1. *For any solution $X \in \Phi$ it holds $\underline{\delta}_X = \delta_X(S_X^-)$ and $\overline{\delta}_X = \delta_X(S_X^+)$.*

If the deterministic problem P is polynomially solvable, then the deviation interval Δ_X for a given solution X can be computed in polynomial time. Hence we can also characterize efficiently the optimality of X and compute its maximal regret. This is very important property of this class of problems. It is worth pointing out that for the linear programming problem with interval objective function coefficients, computing the maximal regret of a given solution is NP-hard [9]. Proposition 1 implies the following result:

Proposition 2. *Every optimal min-max regret solution X is possibly optimal and it is composed of possibly optimal elements.*

Proof. Suppose, by contradiction, that an optimal min-max regret solution X is not possibly optimal. Then, by Proposition 1, $\underline{\delta}_X = \delta_X(S_X^-) = F(X,S_X^-) - F^*(S_X^-) > 0$. Let $Y \in \Phi$ be an optimal solution under S_X^-. Hence $F(Y,S_X^-) < F(X,S_X^-)$. Using the definition of the linear sum cost function we can see that $F(Y,S) < F(X,S)$ for all scenarios $S \in \Gamma$, so $\delta_Y(S) < \delta_X(S)$ for all $S \in \Gamma$. Finally, using again Proposition 1, we get $\overline{\delta}_Y = \delta_Y(S_Y^+) < \delta_X(S_Y^+) \leq \overline{\delta}_X$, which contradicts the assumption that X is an optimal min-max regret solution. Since X is possibly optimal it must be composed of possibly optimal elements. □

We know that a necessarily optimal solution X, i.e. such that $\overline{\delta}_X = 0$, is an optimal min-max regret one. Sometimes such a solution may exist and it can be detected by using the following result:

Theorem 1 ([23]). *Let Y be an optimal solution under scenario S such that $s_e = \frac{1}{2}(\underline{w}_e + \overline{w}_e)$ for all $e \in E$. Then there is a necessarily optimal solution if and only if Y is necessarily optimal.*

So, if problem P is polynomially solvable, then we can detect in polynomial time a necessarily optimal solution if it exists. There is also a general link between necessarily optimal elements and optimal min-max regret solutions.

Theorem 2 ([28]). *If all weight intervals are nondegenerate, then there is an optimal min-max regret solution which contains all necessarily optimal elements.*

The assumption that all weight intervals are nondegenerate is crucial. To see this, consider the minimum spanning tree problem in a connected graph $G = (V, E)$. Assume that $W_e = [1, 1]$ for all $e \in E$. Of course, every element (edge) $e \in E$ is necessarily optimal but all elements do not even form a feasible solution. If there are some degenerate weights, then it can only be shown that for every necessarily optimal element f there is an optimal min-max regret solution that contains f [28].

Let us now focus on computing the deviation interval Δ_f for a given element $f \in E$. Unfortunately, this problem is much harder than computing a solution deviation interval. It is not difficult to show that $\delta_f(S)$ attains minimum and maximum in some extreme scenarios [22]. However, computing these scenarios is not trivial and algorithms for performing this task are known only for some special cases of problem P. A general result can be proven for matroidal problems:

Theorem 3 ([25]). *If* P *is a matroidal problem, then for any element $f \in E$ it holds $\underline{\delta}_f = \delta_f(S^-_{\{f\}})$ and $\overline{\delta}_f = \delta_f(S^+_{\{f\}})$.*

If P is not a matroidal problem, then computing Δ_f may be NP-hard. Specifically, if P is the shortest path, the minimum assignment or the minimum s-t cut, then computing $\underline{\delta}_f$ for a given element f is NP-hard [28]. Furthermore, for these problems even deciding whether $\underline{\delta}_f \leq 0$ is NP-complete, so the problem of asserting the possible optimality of a given element is computationally intractable. This result also means that the lower bound of an element deviation interval is hard to approximate. Interestingly, no polynomially solvable deterministic problem is known for which computing the upper bound $\overline{\delta}_f$ under interval weights is NP-hard. Apart from matroidal problems, $\overline{\delta}_f$ can be efficiently computed in the shortest path problem provided that the input graph is directed and acyclic [16].

Finally, let us focus on solving the min-max regret problem (2). Unfortunately, it turns out to be NP-hard if P is shortest path [7, 24, 42], minimum spanning tree [7, 6], minimum assignment [3] and minimum $s-t$ cut [4]. It is polynomially solvable for the minimum selecting items problem, which has a very simple combinatorial structure [7, 12]. In literature there are two general methods of solving (2). One can design a mixed integer programming model and solve it by using one of many available packages [22, 32, 40]. Alternatively, a branch and bound algorithm can be used to solve the problem [5, 33, 34]. Both techniques have appeared to be quite efficient for some problems and for a description of the results of computational tests we refer the reader to [22, 33, 34, 40].

Notice that Proposition 2 and Theorem 2 suggest a method of preprocessing a problem before solving it. Suppose that we are able to partition the set of elements into three sets, namely $E = A \cup B \cup C$, where A contains nonpossibly optimal elements, B contains necessarily optimal elements and C contains all the remaining elements (the set C contains possibly optimal elements and elements whose status is unknown). According to Proposition 2, we can remove all elements in A from E without violating optimal min-max regret solutions. Similarly, according to Theorem 2, under nondegenerate weights we can automatically add all elements from B to the constructed solution (if there are some degenerate weights, then we can add

a single element from B to the constructed solution). This may significantly reduce the problem size and speed up determining of an optimal min-max regret solution.

3.4.2 Problems with Bottleneck Cost

In this section we discuss a problem with the bottleneck cost $F(X,S) = \max_{e\in X} s_e$. The following theorem suggests a method of computing the deviation interval of a specified solution:

Theorem 4. *For any solution X it holds*

$$\underline{\delta}_X = \max\{0, \max_{e\in X} \underline{w}_e - F^*(S_E^+)\}, \tag{3}$$

$$\overline{\delta}_X = \max_{e\in X} \max\{0, \overline{w}_e - F^*(S_{\{e\}}^+)\}. \tag{4}$$

Proof. The proof of equality (4) can be found in [7]. We prove equality (3). Let $S \in \Gamma$ be a scenario that minimizes the deviation, that is $\underline{\delta}_X = \delta_X(S) = F(X,S) - F^*(S)$. Since $\max_{e\in X} \underline{w}_e \le F(X,S)$, $F^*(S_E^+) \ge F^*(S)$ and $\underline{\delta}_X \ge 0$ it follows immediately that

$$\underline{\delta}_X \ge \max\{0, \max_{e\in X} \underline{w}_e - F^*(S_E^+)\}. \tag{5}$$

It remains to show that the inequality $\le$ also holds in (5). Let Y be an optimal solution under S_E^+ and let $g = \arg\max_{e\in Y} \overline{w}_e$. We consider two cases. (i) $\max_{e\in X} \underline{w}_e > \overline{w}_g$. Denote $h = \arg\max_{e\in X} \underline{w}_e$. Consider scenario S such that $s_e = \min\{\underline{w}_h, \overline{w}_e\}$ for all $e \in X$ and $s_e = \overline{w}_e$ for all $e \in E \setminus X$. Since $\underline{w}_h \ge \underline{w}_e$ for all $e \in X$, $S \in \Gamma$. It is easy to check that $F(X,S) = \underline{w}_h$ and $F^*(S) = F(S_E^+)$. Hence $\underline{\delta}_X \le \delta_X(S) = \max_{e\in X} \underline{w}_e - F^*(S_E^+) \le \max\{0, \max_{e\in X} \underline{w}_e - F^*(S_E^+)\}$, which together with (5) yield (3). (ii) $\max_{e\in X} \underline{w}_e \le \overline{w}_g$. Consider scenario S such that under this scenario all elements $e \in E \setminus X$ have weights $\overline{w}_e$ and all the elements $e \in X$ have weights $\min\{\overline{w}_e, \overline{w}_g\}$. Since $\underline{w}_e \le \overline{w}_g$ for all $e \in X$, $S \in \Gamma$. One can easily verify that X is optimal under S, which means that $\underline{\delta}_X = 0 \le \max\{0, \max_{e\in X} \underline{w}_e - F^*(S_E^+)\}$. This, together with (5), give (3). □

We thus can see that it is not difficult to compute the deviation interval for a given solution and to characterize its optimality, provided that the deterministic problem P is polynomially solvable. Similarly to the problems with linear sum cost function, the following proposition holds:

Proposition 3. *Every optimal min-max regret solution X is possibly optimal and it is composed of possibly optimal elements.*

Proof. If X is not possibly optimal, then by Theorem 4, we have $\max_{e\in X} \underline{w}_e > F^*(S_E^+)$. Let Y be an optimal solution under S_E^+. It is easy to verify that $F(X,S) > F(Y,S)$ under all scenarios $S \in \Gamma$. The same argument as in the proof of Proposition 1 yields $\overline{\delta}_X > \overline{\delta}_Y$, so X cannot be an optimal min-max regret solution. Of course, a possibly optimal solution is entirely composed of possibly optimal elements. □

Let us now focus on the elements. The following theorem allows us to compute a lower bound on the deviation interval of a given element. Its proof is very similar to the proof of Theorem 2.

Theorem 5. *For any element $f \in E$ it holds*

$$\underline{\delta}_f = \max\{0, \min_{X \in \Phi_f} F(X, S_E^-) - F^*(S_E^+)\}. \tag{6}$$

Using Theorem 5 we can design an efficient method of computing the quantity $\underline{\delta}_f$ for every particular problem, which is polynomially solvable. In order to compute $\min_{X \in \Phi_f} F(X, S_E^-)$ a slight modification of the algorithm for solving P is only required. Therefore, contrary to the problems with the linear sum cost, we can also characterize efficiently the possible optimality of a given element. However, a general characterization of the quantity $\overline{\delta}_f$ is unknown and it is an interesting subject of further research. Both bounds of Δ_f can be efficiently computed if P is matroidal problem. It is not difficult to prove the following result:

Proposition 4. *If* P *is a matroidal problem then* $\underline{\delta}_f = \max\{0, \underline{w}_f - F^*(S_E^+)\}$ *and* $\overline{\delta}_f = \max\{0, \overline{w}_f - F^*(S_{\{f\}}^+)\}$.

Equality (4) allows us to solve efficiently the min-max regret problem (2), provided that P is polynomially solvable. To see this let us define weights $\hat{w}_e = \max\{0, \overline{w}_e - F^*(S_{\{e\}}^+)\}$ for all $e \in E$. Then

$$\min_{X \in \Phi} \overline{\delta}_X = \min_{X \in \Phi} \max_{e \in X} \hat{w}_e$$

and the min-max regret problem reduces to solving the deterministic problem P with nonnegative real weights $\hat{w}_e$, $e \in E$. We thus get the following theorem:

Theorem 6 ([7]). *If the deterministic problem* P *can be solved in $f(n)$ time, then its min-max regret version can be solved in $O(nf(n))$ time.*

The running time $O(nf(n))$ follows from the fact that we need to solve n times the deterministic problem P to obtain weights $\hat{w}_e$ for all $e \in E$. The computations can be additionally refined and for details we refer the reader to [7]. The most important consequence of Theorem 6 is that the min-max regret version of problem P is polynomially solvable if only the deterministic problem P is polynomially solvable. So, the situation is quite different from the problems with the linear sum cost.

4 Combinatorial Optimization Problems with Fuzzy Weights

In the previous section we have described the class of problems with interval weights. It turns out that all the introduced concepts can be naturally extended without significant increasement of the problem complexity. The key idea is to use fuzzy

intervals to model the uncertain element weights and apply possibility theory to extend the concept of deviation. This section is devoted to the class of combinatorial problems with fuzzy weighs.

4.1 Basic Notions of Possibility Theory

Possibility theory offers us a framework of dealing with imprecision. A detailed description of this theory can be found in a book [14]. We now recall some of its notions, which will be used later in this section. A *fuzzy interval* $\tilde{A}$ is a fuzzy set in the space of reals whose membership function $\mu_{\tilde{A}} : \mathbb{R} \to [0,1]$ is normal, quasi concave, upper semicontinuous and has a compact support. The main property of a fuzzy interval is that all its λ-*cuts*, that is the sets $\tilde{A}^{\lambda} = \{x : \mu_{\tilde{A}}(x) \geq \lambda\}$ for $\lambda \in (0,1]$, are closed intervals. We will also denote by $\tilde{A}^0$ the smallest closed set containing the support of $\tilde{A}$. So, we can represent a fuzzy interval $\tilde{A}$ as a family of closed intervals $\tilde{A}^{\lambda} = [\underline{a}^{\lambda}, \overline{a}^{\lambda}]$ parametrized by the value of $\lambda \in [0,1]$. It is easy to see that this family is monotone, that is $\tilde{A}^{\lambda_1} \subseteq \tilde{A}_2^{\lambda}$ if $\lambda_1 \geq \lambda_2$. Having the family of λ-cuts of $\tilde{A}$, the membership function $\mu_{\tilde{A}}$ can be computed as follows:

$$\mu_{\tilde{A}}(x) = \sup\{\lambda \in [0,1] : \ x \in \tilde{A}^{\lambda}\} \tag{7}$$

and $\mu_{\tilde{A}}(x) = 0$ if $x \notin \tilde{A}^0$.

In practice the class of *trapezoidal fuzzy intervals* is commonly used (see Figure 1). Every trapezoidal fuzzy interval can be described as a quadruple $(\underline{a}, \overline{a}, \alpha, \beta)$ and can be represented by the following family of λ-cuts:

$$\tilde{A}^{\lambda} = [\underline{a} - \alpha(1-\lambda), \overline{a} + \beta(1-\lambda)], \ \lambda \in [0,1]. \tag{8}$$

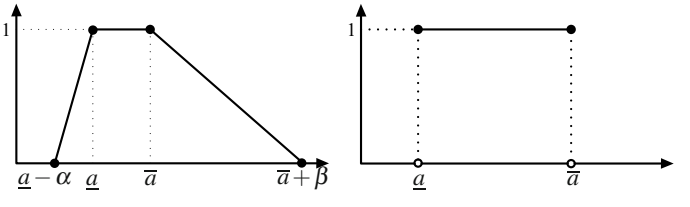


Fig. 1 Trapezoidal fuzzy interval $(\underline{a}, \overline{a}, \alpha, \beta)$ and closed interval $[\underline{a}, \overline{a}] = (\underline{a}, \overline{a}, 0, 0)$.

Notice that this representation also contains closed intervals (if $\alpha = \beta = 0$) and real numbers (if additionally $\underline{a} = \overline{a}$). We will use shorter notation (a, α, β) if $a = \underline{a} = \overline{a}$ and we will call (a, α, β) a *triangular fuzzy interval*. We also define $(\overline{a}, \beta) = (0, \overline{a}, 0, \beta)$. In order to simplify notations and discussion we will only use trapezoidal

fuzzy intervals. However, it is not difficult to extend all the introduced notions to a more general class of fuzzy intervals of the L-R type with compact support (see [14] for a descriptions of this class of fuzzy intervals).

W now give an interpretation of a fuzzy interval. Let a be a real quantity whose value is not precisely known. We associate with a a fuzzy interval $\tilde{A}$, whose membership function $\mu_{\tilde{A}}$ is a *possibility distribution* for the values of a, that is

$$\Pi(a = x) = \mu_{\tilde{A}}(x),$$

where $\Pi(a = x)$ is the possibility of the event that a will take the value of x. There are several methods of obtaining possibility distribution for an unknown quantity and their description can be found in [14]. Observe that $\tilde{A}^{\lambda}$ contains all values of a whose possibility of occurrence is not less than λ. In consequence, $\tilde{A}^0$ should contain all possible values of a and $\tilde{A}^1$ should contain the most plausible ones.

Let $\tilde{G}$ be a fuzzy set in the space of reals. Then $a \in \tilde{G}$ is a *fuzzy event* and the possibility and necessity of $a \in \tilde{G}$ are defined as follows [13]:

$$\Pi(a \in \tilde{G}) = \sup_{x \in \mathbb{R}} \min\{\mu_{\tilde{A}}(x), \mu_{\tilde{G}}(x)\}. \tag{9}$$

$$\mathrm{N}(a \in \tilde{G}) = 1 - \Pi(a \notin \tilde{G}) = 1 - \sup_{x \in \mathbb{R}} \min\{\mu_{\tilde{A}}(x), 1 - \mu_{\tilde{G}}(x)\}. \tag{10}$$

where $1 - \mu_{\tilde{G}}(x)$ is the membership function of the complement of the fuzzy set $\tilde{G}$. It is not difficult to show that if $\tilde{G} = (0, \overline{g}, 0, \beta) = (\overline{g}, \beta)$, then the following equality is true:

$$\mathrm{N}(a \in \tilde{G}) = 1 - \inf\{\lambda \in [0,1] : \overline{a}^{\lambda} \leq \overline{g}^{1-\lambda}\} \tag{11}$$

and $\mathrm{N}(a \in \tilde{G}) = 0$ if $\overline{a}^1 > \overline{g}^0$. Equality (11) is illustrated in Figure 2.

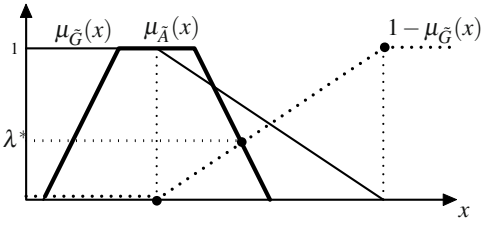


Fig. 2 $\mathrm{N}(a \in \tilde{G}) = 1 - \lambda^*$

In the next section we will show how possibility theory allows us to extend the concept of scenario set.

4.2 Fuzzy Scenario Set

Assume that the element weights in a combinatorial optimization problem are unknown real quantities w_e, $e \in E$. We associate with every weight w_e a fuzzy interval $\tilde{W}_e$. According to the interpretation given in the previous section, the membership function $\mu_{\tilde{W}_e}$ is a possibility distribution for the values of the weight w_e. Let $S = (s_e)_{e \in E} \in \mathbb{R}^n$ be a vector representing the state of the world in which $w_e = s_e$ for all $e \in E$. As in Section 3, we will call S a scenario. The possibility distributions associated with element weights induce the following possibility distribution over all scenarios $S \in \mathbb{R}^n$:

$$\pi(S) = \Pi\left(\bigwedge_{e \in E} [w_e = s_e]\right) = \min_{e \in E} \Pi(w_e = s_e) = \min_{e \in E} \mu_{\tilde{W}_e}(s_e). \tag{12}$$

Observe that $\pi(S)$ may be regarded as a membership function of a fuzzy set in $\mathbb{R}^n$. We will call this fuzzy set a *fuzzy scenario set* and $\pi(S)$ is the possibility of the event that scenario $S \in \mathbb{R}^n$ will occur. Notice that we generalize in this way scenario set Γ defined in Section 3. Indeed, under interval uncertainty representation $\pi(S) = 1$ if $S \in \Gamma$ and $\pi(S) = 0$ otherwise, so $\pi(S)$ is then a characteristic function of the set Γ. Under fuzzy weights, $\pi(S)$ may take any value in the interval $[0,1]$. Hence fuzzy weights provide us more information about the state of the world which may occur. In particular, scenario S is impossible if $\pi(S) = 0$ and we have $\pi(S) = 1$ for the most plausible scenarios. Notice that the definition of a fuzzy interval assures that $\pi(S) = 1$ for at least one scenario S.

Using (12) and the definition of λ-cut it is easily seen that for every $\lambda \in [0,1]$ the following equality holds:

$$\{S:\ \pi(S) \geq \lambda\} = \times_{e \in E} [\underline{w}_e^{\lambda}, \overline{w}_e^{\lambda}]. \tag{13}$$

So, the set of all scenarios whose possibility of occurrence is not less than λ is the Cartesian product of the interval weights being the λ-cuts of the fuzzy weights. Hence it forms a scenario set, which we will denote as Γ^{λ}. This property allows us to decompose the fuzzy problem into a family of interval problems. We will make use of this fact in the next sections.

4.3 Fuzzy Deviations

As for the problems with deterministic and interval weights, we can use the concept of deviation to characterize the optimality of solutions and elements. Recall that under interval weights deviations δ_X and δ_f fall within closed intervals Δ_X and Δ_f. Under fuzzy weights, the solution and element deviations are unknown quantities, which fall within fuzzy intervals $\tilde{\Delta}_X$ and $\tilde{\Delta}_f$. The membership functions $\mu_{\tilde{\Delta}_X}$ and $\mu_{\tilde{\Delta}_f}$

are possibility distributions for the values of δ_X and δ_f and, according to possibility theory, they are defined as follows:

$$\mu_{\tilde{\Delta}_X}(y) = \Pi(\delta_X = y) = \sup_{\{S:\, \delta_X(S)=y\}} \pi(S),$$
$$\mu_{\tilde{\Delta}_f}(y) = \Pi(\delta_f = y) = \sup_{\{S:\, \delta_f(S)=y\}} \pi(S).$$

Consider fuzzy deviation interval $\tilde{\Delta}_X$ of a given solution $X \in \Phi$. A λ-cut of $\tilde{\Delta}_X$ contains all values of the deviation of X whose possibility of occurrence is not less than λ. Hence

$$\tilde{\Delta}_X^\lambda = \{y : \mu_{\tilde{\Delta}_X}(y) \geq \lambda\} = \{\delta_X(S) :\ \pi(S) \geq \lambda, S \in \mathbb{R}^n\}.$$

But (13) implies $\tilde{\Delta}_X^\lambda = \{\delta_X(S) :\ S \in \Gamma^\lambda\} = [\underline{\delta}_X^\lambda, \overline{\delta}_X^\lambda]$, where $\underline{\delta}_X^\lambda$ minimizes and $\overline{\delta}_X^\lambda$ maximizes $\delta_X(S)$ over all $S \in \Gamma^\lambda$. We can now use the results from Sections 3.4.1 and 3.4.2 to compute the bounds of $\tilde{\Delta}_X^\lambda$. For a problem with the linear sum cost, Proposition 1 gives

$$\underline{\delta}_X^\lambda = \delta_X(S_X^{-\lambda}) = F(X, S_X^{-\lambda}) - F^*(S_X^{-\lambda}) = \sum_{e \in X} \underline{w}_e^\lambda - F^*(S_X^{-\lambda}), \tag{14}$$

$$\overline{\delta}_X^\lambda = \delta_X(S_X^{+\lambda}) = F(X, S_X^{+\lambda}) - F^*(S_X^{+\lambda}) = \sum_{e \in X} \overline{w}_e^\lambda - F^*(S_X^{+\lambda}), \tag{15}$$

where $S_X^{-\lambda}$ and $S_X^{+\lambda}$ are the corresponding extreme scenarios in Γ^λ.

Now our aim is to compute the family of cuts $\tilde{\Delta}_X^\lambda$ for $\lambda \in [0,1]$. The possibility distribution for the deviation of X can be then obtain by formula (7). Observe that it remains to compute functions $F^*(S_X^{-\lambda})$ and $F^*(S_X^{+\lambda})$ of $\lambda \in [0,1]$. This task can be performed by applying a parametric technique. Namely, we wish to compute sequences $0 = \lambda_0 < \lambda_1 < \ldots < \lambda_k = 1$ and $X_0, \ldots, X_{k-1}$ such that X_i is an optimal solution under $S_X^{+\lambda}$ or $S_X^{-\lambda}$ for $\lambda \in [\lambda_i, \lambda_{i+1}]$. Having these sequences it is easy to describe analytically functions $F^*(S_X^{-\lambda})$ or $F^*(S_X^{+\lambda})$ for $\lambda \in [0,1]$. It turns out that if $\underline{w}_e^\lambda$ and $\overline{w}_e^\lambda$ are linear functions of λ for each $e \in E$, then for some particular problems such as shortest path or minimum spanning tree their parametric counterparts can be efficiently solved (see e.g. [15, 38, 41]). In consequence, the family of intervals $\tilde{\Delta}_X^\lambda$, $\lambda \in [0,1]$, can be efficiently computed if the uncertain weights are modeled as trapezoidal fuzzy intervals. A similar reasoning applies to the problems with the bottleneck cost function. One should only use Theorem 4 to obtain the corresponding parametric problems. We now illustrate the computation of $\tilde{\Delta}_X$ by an example.

Example 1. Consider a shortest path problem shown in Figure 3a. We are given a directed graph composed of 5 arcs and we wish to find a shortest path between nodes s and t.

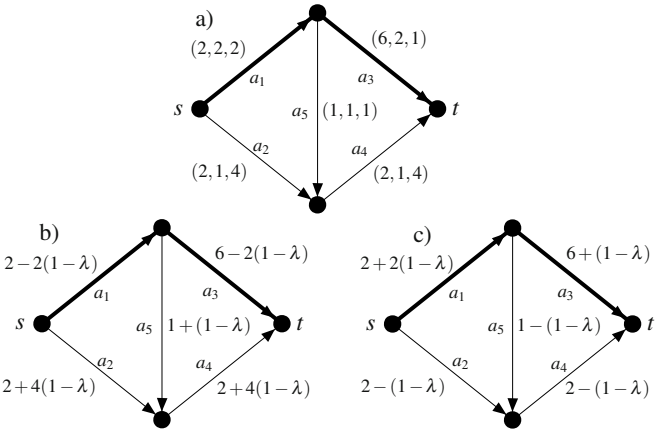


Fig. 3 a) A sample shortest path problem with fuzzy weights. b) The extreme scenario $S^{-\lambda}_{\{a_1,a_3\}}$. c) The extreme scenario $S^{+\lambda}_{\{a_1,a_3\}}$.

For every arc weight w_a, $a \in A$, a triangular fuzzy interval $\tilde{W}_a = (w_a, \alpha_a, \beta_a)$ is given. Let us examine path $X = \{a_1, a_3\}$. In Figures 3b and 3c the extreme scenarios $S_X^{-\lambda}$ and $S_X^{+\lambda}$ are shown. Notice that these scenarios are linear functions of $\lambda \in [0,1]$ obtained by formula (8). It holds $F(X, S_X^{-\lambda}) = 8 - 4(1-\lambda)$ and $F(X, S_X^{+\lambda}) = 8 + 3(1-\lambda)$ for $\lambda \in [0,1]$. Applying a parametric technique to the problem shown in Figure 3b, we get a sequence of λ's $0 < \frac{4}{7} < \frac{4}{5} < 1$ that corresponds to the sequence of optimal solutions $\{a_1,a_3\}, \{a_1,a_5,a_4\}, \{a_2,a_4\}$. That is, solution $\{a_1,a_3\}$ is optimal for $\lambda \in [0, \frac{4}{7}]$, solution $\{a_1,a_5,a_4\}$ is optimal for $\lambda \in [\frac{4}{7}, \frac{4}{5}]$ and solution $\{a_2,a_4\}$ is optimal for $\lambda \in [\frac{4}{5}, 1]$. Hence $F^*(S_X^{-\lambda})$ is a piecewise linear function, whose value is $8-4(1-\lambda)$ for $\lambda \in [0, \frac{4}{7}]$, $5+3(1-\lambda)$ for $\lambda \in [\frac{4}{7}, \frac{4}{5}]$ and $4+8(1-\lambda)$ for $\lambda \in [\frac{4}{5}, 1]$. Subtracting $F^*(S_X^{-\lambda})$ from $F(X, S_X^{-\lambda})$ yields $\underline{\delta}_X^{\lambda}$. Similarly, the function $\overline{\delta}_X^{\lambda}$ is obtained by applying the parametric technique to the problem shown in Figure 3c. The resulting functions $\underline{\delta}_X^{\lambda}$ and $\overline{\delta}_X^{\lambda}$ are presented in Figure 4. Having the bounds $\underline{\delta}_X^{\lambda}$ and $\overline{\delta}_X^{\lambda}$ for $\lambda \in [0,1]$ we can construct the possibility distribution $\mu_{\tilde{\Delta}_X}$ for the deviations of X by applying formula (7). This possibility distribution is shown in Figure 4. □

Computing fuzzy deviation of a given element is more complex. It is a direct consequence of the fact that the corresponding interval problem may be hard to solve. In other words, it may be hard to identify extreme scenarios that minimize or maximize an element deviation. We can compute the fuzzy interval $\tilde{\Delta}_f$ only for some particular problems such as matroidal ones (see e.g. [27]).

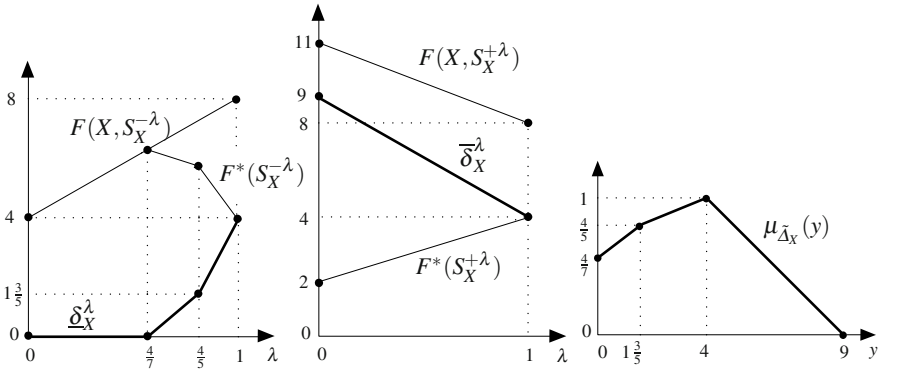


Fig. 4 Bounds $\underline{\delta}_X^{\lambda}$ and $\overline{\delta}_X^{\lambda}$ and the possibility distribution $\mu_{\tilde{\Delta}_X}(y) = \Pi(\delta_X = y)$

4.4 Degrees of Possible and Necessary Optimality

The fuzzy deviations allow us to characterize possible and necessary optimality of solutions and elements. Recall that the statement "X is optimal" is equivalent to the assertion $\delta_X = 0$. So, we can define the *degrees of possible and necessary optimality* of solution X in the following way:

$$\Pi(X \text{ is optimal}) = \Pi(\delta_X = 0) = \mu_{\tilde{\Delta}_X}(0), \tag{16}$$

$$\mathrm{N}(X \text{ is optimal}) = 1 - \Pi(\delta_X > 0) = 1 - \sup_{y>0} \mu_{\tilde{\Delta}_X}(y). \tag{17}$$

In the same way we can define the degrees of optimality of the elements. It is enough to replace X with f in (16) and (17). The following relations hold between the optimality degrees of solutions and elements:

$$\Pi(X \text{ is optimal}) \leq \max_{e \in X} \Pi(e \text{ is optimal}).$$

$$\mathrm{N}(X \text{ is optimal}) \leq \max_{e \in X} \mathrm{N}(e \text{ is optimal}).$$

Having possibility distributions $\mu_{\tilde{\Delta}_X}$ and $\mu_{\tilde{\Delta}_f}$ we can immediately compute the degrees of optimality of X and f. However, if one wishes to obtain only the optimality degrees, then the computations can be significantly simplified. Equalities (7) and (16) imply

$$\Pi(X \text{ is optimal}) = \sup\{\lambda \in [0,1] : 0 \in \tilde{\Delta}_X^{\lambda}\} = \sup\{\lambda \in [0,1] : \underline{\delta}_X^{\lambda} = 0\} \tag{18}$$

and $\Pi(X \text{ is optimal}) = 0$ if $\underline{\delta}_X^{0} > 0$. So, in order to compute the degree of possible optimality of X, we need to find the largest value of λ such that X is possibly optimal

under scenario set Γ^{λ} (which is equivalent to the condition $\underline{\delta}_X^{\lambda} = 0$). Since $\underline{\delta}_X^{\lambda}$ is nondecreasing function of λ, the standard binary search technique can be applied to perform this task. Also, the following equality is easy to establish:

$$\mathrm{N}(X \text{ is optimal}) = 1 - \inf\{\lambda \in [0,1] : \overline{\delta}_X^{\lambda} = 0\} \tag{19}$$

and $\mathrm{N}(X \text{ is optimal}) = 0$ if $\overline{\delta}_X^{1} > 0$. So, we need to find the smallest value of λ such that X is necessarily optimal under scenario set Γ^{λ} (which is equivalent to the condition $\overline{\delta}_X^{\lambda} = 0$). Because $\overline{\delta}_X^{\lambda}$ is nonincreasing function of λ, the binary search technique also solves this problem. If $f(n)$ is the time required to assert whether a given solution is possibly (necessarily) optimal in the interval problem, then its degree of possible (necessary) optimality can be computed in $O(f(n)\log \varepsilon^{-1})$ time, where $\varepsilon \in (0,1)$ is an assumed precision of calculations.

Exactly the same reasoning can be applied to the elements (we only need to replace X with f in (18) and (19)). Note, however, that the complexity of computations for an element strongly depends on the combinatorial structure of problem P.

4.5 Choosing a Solution under Fuzzy Weights

We now address the problem of choosing a solution under fuzzy weights. The degrees of optimality, introduced in the previous section, suggest us a solution method. We can choose a solution, which maximizes the degree of possible or necessary optimality. Maximizing the degree of possible optimality is trivial. There is always at least one solution $X \in \Phi$ for which the degree of possible optimality attains its maximal value equal to 1. It can be obtained by computing an optimal solution under scenario S such that $\pi(S) = 1$. On the other hand, the degree of necessary optimality of every feasible solution may be very small or even equal to 0. We thus meet the same problem as in the interval uncertainty representation - the possible optimality is too weak criterion of choosing a solution and the necessary optimality is too strong.

Now the idea is to replace the strong optimality requirement with a weaker one. Suppose that a decision maker knows his/her preference about solution deviation and expresses it using a *fuzzy goal* $\tilde{G} = (\overline{g}, \beta_g)$. So, the values of deviation in $[0, \overline{g}]$ are completely accepted, the values in $[\overline{g} + \beta_g, \infty)$ are not at all accepted and the degree of acceptance decreases linearly in the interval $[\overline{g}, \overline{g} + \beta_g]$ (the assumption that it decreases linearly is not restrictive and any decreasing function can be used to model the decision maker preferences). We can now replace the strong requirement $\delta_X = 0$ with a weaker one, namely $\delta_X \in \tilde{G}$. Recall that δ_X is an unknown quantity characterized by possibility distribution $\mu_{\tilde{\Delta}_X}$. So, $\delta_X \in \tilde{G}$ is a fuzzy event and we can compute the necessity that it holds, $\mathrm{N}(\delta_X \in \tilde{G})$, using (10):

$$\mathrm{N}(\delta_X \in \tilde{G}) = 1 - \sup_{y \in \mathbb{R}} \min\{\mu_{\tilde{\Delta}_X}(y), 1 - \mu_{\tilde{G}}(y)\}.$$

Consider again the shortest path problem from Example 1. The function $\mu_{\tilde{\Delta}_X}$ is a possibility distribution for deviation δ_X of path X. This possibility distribution is shown in Figure 5. The part of $\mu_{\tilde{\Delta}_X}$ representing the largest deviation of X is shown in bold. In Figure 5 a fuzzy goal $\tilde{G}$ and its complement $\tilde{G}^d$ are also shown. The complement $\tilde{G}^d$ expresses a degree of dissatisfaction of the values of solution deviation.

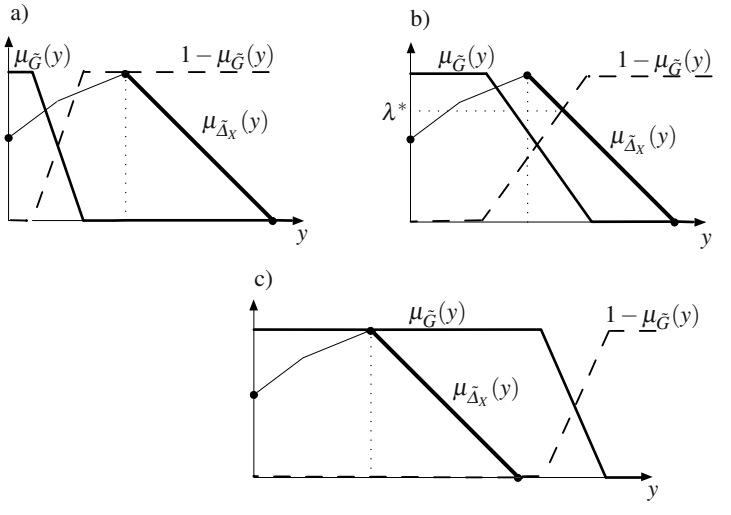


Fig. 5 Three different situations depending on the choice of fuzzy goal $\tilde{G}$: a) $\mathrm{N}(\delta_X \in \tilde{G}) = 0$, b) $\mathrm{N}(\delta_X \in \tilde{G}) = 1-\lambda^*$, c) $\mathrm{N}(\delta_X \in \tilde{G}) = 1$

Consider the case illustrated in Figure 5a. The goal $\tilde{G}$ is chosen so that the largest deviation of X is fully contained in its complement $\tilde{G}^d$. So, $\Pi(\delta_X \in \tilde{G}^d) = 1$ and $\mathrm{N}(\delta_X \in \tilde{G}^d) = 0$. In other words, with possibility equal to 1 a scenario may occur for which the deviation of X is not at all accepted. Figure 5c shows an opposite case. The goal $\tilde{G}$ is chosen so that the largest deviation of X is completely not in $\tilde{G}^d$. So, $\Pi(\delta_X \in \tilde{G}^d) = 0$ and $\mathrm{N}(\delta_X \in \tilde{G}^d) = 1$. In this case for every scenario S such that $\pi(S) > 0$ the deviation $\delta_X(S)$ is completely accepted. Clearly, this is an ideal situation. In Figure 5b a third case is shown, where the largest deviation of X is only partially contained in $\tilde{G}^d$. So, $\Pi(\delta_X \in \tilde{G}^d) = \lambda^*$ and $\mathrm{N}(\delta_X \in \tilde{G}^d) = 1-\lambda^*$. This means that for all scenarios S such that $\pi(S) \geq \lambda^*$ the degree of dissatisfaction is not greater than λ^* or, equivalently, the degree of satisfaction is not less than $1-\lambda^*$.

Now it is reasonable to choose a solution whose deviation belongs to $\tilde{G}$ with the highest confidence. This leads to the following optimization problem:

$$\max_{X \in \Phi} \mathrm{N}(\delta_X \in \tilde{G}). \tag{20}$$

An optimal solution to (20) is called a *most necessarily soft optimal solution* and it was first proposed as a solution under fuzzy weights in [20]. If we choose $\tilde{G} = (0,0)$, then we get the following special case of (20):

$$\max_{X \in \Phi} \mathrm{N}(\delta_X = 0) = \max_{X \in \Phi} \mathrm{N}(X \text{ is optimal}). \tag{21}$$

So, in (21) we seek a most necessarily optimal solution. As we will see in the next section, the problem (21) may be easier to solve than (20). Using (11) we can express the problem (20) as the following mathematical programming one:

$$\begin{array}{l} \min \lambda \\ \overline{\delta}_X^{\lambda} \leq \overline{g}^{1-\lambda} \\ X \in \Phi \\ \lambda \in [0,1] \end{array} \tag{22}$$

If λ^* is the optimal objective value of (22) and X^* is an optimal solution, then $\mathrm{N}(\delta_{X^*} \in \tilde{G}) = 1 - \lambda^*$. If (22) is infeasible, then $\mathrm{N}(\delta_X \in \tilde{G}) = 0$ for all feasible solutions X.

It is easy to check that problem (22) is a generalization of the min-max regret approach. If all $\tilde{W}_e$, $e \in E$, are closed intervals and $\tilde{G} = (0,M)$ for a sufficiently large number M, then (22) is equivalent to (2). In the next two sections we will focus on some methods of solving (22).

4.5.1 Binary Search Technique

Observe that $\overline{\delta}_X^{\lambda}$ is nonincreasing and $\overline{g}^{1-\lambda}$ is nondecreasing function of $\lambda \in [0,1]$. Therefore (22) can be solved by applying the standard binary search technique shown in Figure 6. The algorithm simply seeks a minimal value of λ in the interval [0,1], for which there is a solution $X \in \Phi$ that satisfies inequality $\overline{\delta}_X^{\lambda} \leq \overline{g}^{1-\lambda}$. The quantity $\overline{\delta}_X^{\lambda}$ is the maximal regret of solution X under scenario set Γ^{λ}. Therefore, the inequality $\overline{\delta}_X^{\lambda} \leq \overline{g}^{1-\lambda}$ is satisfied for some $X \in \Phi$ if and only if it is satisfied by an optimal min-max regret solution under Γ^{λ}. So, if we are able to solve the min-max regret problem with interval data in $f(n)$ time, then the binary search solves the fuzzy problem in $O(f(n)\log \varepsilon^{-1})$ time with a given precision $\varepsilon \in (0,1)$.

We can see now that if the min-max regret problem is polynomially solvable, then its fuzzy generalization is polynomially solvable up to a given precision ε. Notice that for the class of problems with the bottleneck cost function, it is enough that the deterministic problem is polynomially solvable (see Theorem 4). For the problems with the linear sum cost the situation is more complex since the min-max regret problem is mostly NP-hard. However, if the deterministic problem is polynomially solvable, then we can solve efficiently the special case (21), that is we can find efficiently a most necessarily optimal solution with a given precision ε. If $\tilde{G} = (0,0)$, then $\overline{g}^{\lambda} = 0$ for all $\lambda \in [0,1]$. The condition $\overline{\delta}_X^{\lambda} \leq 0$ can be efficiently

1: Find an optimal min-max regret solution X under Γ^1
2: **if** $\overline{\delta}_X^1 > \overline{g}^0$ **then return** $\emptyset$
3: $\lambda_1 \leftarrow 0.5$, $k \leftarrow 1$, $\lambda_2 \leftarrow 0$
4: **while** $|\lambda_1 - \lambda_2| < \varepsilon$ **do**
5: $\quad \lambda_2 \leftarrow \lambda_1$
6: $\quad$ Find an optimal min-max regret solution Y under Γ^{λ_1}
7: $\quad$ **if** $\overline{\delta}_Y^{\lambda_1} \leq \overline{g}^{1-\lambda_1}$ **then** $\lambda_1 \leftarrow \lambda_1 - 1/2^{k+1}$, $X \leftarrow Y$ **else** $\lambda_1 \leftarrow \lambda_1 + 1/2^{k+1}$
8: $\quad k \leftarrow k+1$
9: **end while**
10: **return** X

Fig. 6 Computing a most necessarily soft optimal solution with a given precision $\varepsilon \in (0,1)$. Algorithm returns $\emptyset$ if $\mathrm{N}(\delta_X \in \tilde{G}) = 0$ for all $X \in \Phi$.

verified for a fixed λ by using Theorem 1 because $\overline{\delta}_X^{\lambda} \leq 0$ if and only if there is a necessarily optimal solution under scenario set Γ^{λ}.

The binary search is the most general method of solving the fuzzy problem. However, it gives only an approximate solution. Furthermore, it may be not efficient for the problems with the linear sum cost function because solving $O(\log \varepsilon^{-1})$ times an NP-hard problem may be time consuming. In the next sections we show some alternative methods of finding a most necessarily soft optimal solution.

4.5.2 Parametric Technique of the Problems with Bottleneck Cost

Consider the class of problems with the bottleneck cost function. Using (4) and (11) we can express the fuzzy problem in the following way:

$$\inf \left\{ \lambda \in [0,1] : \min_{X \in \Phi} \max_{e \in X} \hat{w}_e^{\lambda} \leq \overline{g}^{1-\lambda} \right\}. \tag{23}$$

where $\hat{w}_e^{\lambda} = \max\{0, \overline{w}_e^{\lambda} - F^*(S_{\{e\}}^{+\lambda})\}$. We can obtain weights $\hat{w}_e^{\lambda}$ for all $e \in E$ using a parametric technique (see e.g. [11]). As the result we obtain another parametric bottleneck problem with weights $\hat{w}_e^{\lambda}$, $e \in E$, that is

$$\overline{\delta}^{\lambda} = \min_{X \in \Phi} \max_{e \in X} \hat{w}_e^{\lambda}. \tag{24}$$

Solving (24) we obtain sequences $0 \leq \lambda_0 \leq \lambda_1 \leq \cdots \leq \lambda_k = 1$ and $X_0, \ldots, X_{k-1}$ such that X_i is an optimal solution for $\lambda \in [\lambda_i, \lambda_{i+1}]$. Having these sequences it is easy to describe analytically function $\overline{\delta}^{\lambda}$ for $\lambda \in [0,1]$. The function $\overline{\delta}^{\lambda}$ is nonincreasing, hence from (23) we can see that in order to obtain a most necessarily soft optimal solution we must find the intersection point λ^* of $\overline{\delta}^{\lambda}$ with $\overline{g}^{1-\lambda}$. Then, if $\lambda^* \in [\lambda_i, \lambda_{i+1}]$, then X_i is a necessarily soft optimal solution. If such an intersection point does not exist, then two cases are possible - either $\overline{\delta}^1 > \overline{g}^0$ or $\overline{\delta}^0 < \overline{g}^1$. In the former

case $\mathrm{N}(\delta_X \in \tilde{G}) = 0$ for all feasible solutions X and in the latter one $\mathrm{N}(\delta_{X_0} \in \tilde{G}) = 1$ and X_0 is a necessarily soft optimal solution.

The solution procedure based on a parametric technique is more time consuming than the binary search shown in the previous section. It has, however, two important advantages. First of all, it gives an exact necessarily soft optimal solution. Furthermore, it provides a lot of additional information in the fuzzy problem. Observe that, regardless of fuzzy goal, a most necessarily soft optimal solution is always among $X_0,\ldots,X_{k-1}$. We can thus treat the set of solutions $\{X_0,\ldots,X_{k-1}\}$ as a solution of the fuzzy problem. Introducing fuzzy goal $\tilde{G}$ allows us to chose one of these solutions. One can also check easily how the solution changes when the fuzzy goal $\tilde{G}$ is changed. So, we can perform a sensitivity analysis of the obtained solution.

4.5.3 MIP Formulation for the Problems with Linear Sum Cost

In this section we show an exact method of solving (20) for the problems with the linear sum cost. Under some additional assumptions we design a mixed integer linear programming (MIP) model, which can be then solved by some available software. Let us assign a binary variable $x_i \in \{0,1\}$ to every element $e_i \in E$. This variable will indicate whether element e_i is contained in a constructed solution. Every feasible solution $X \in \Phi$ can be represented as a vector of binary variables $\boldsymbol{x} = [x_1,\ldots,x_n]$, where $x_i = 1$ if and only if $e_i \in X$. We assume that the set of feasible solutions can be described by a system of linear constraints of the form $\{\boldsymbol{x} \in \{0,1\}^n : \mathscr{A}\boldsymbol{x}^T = \boldsymbol{b}\}$, where $\mathscr{A}$ is a matrix and $\boldsymbol{b}$ is a vector of fixed coefficients. We allow also inequalities $\leq$ and $\geq$ in the constraints since they can easily be converted to equalities by adding a number of additional slack variables. In order to simplify notations we will use $\tilde{W}_i$ to denote the fuzzy interval associated with the weight of element e_i.

We will assume that the matrix $\mathscr{A}$ is *totally unimodular*. Recall that in a totally unimodular matrix the determinants of all its nonsingular square submatrices are equal to -1 or 1 (see e.g. [18]). This assumption restricts the class of considered problems. However, if the deterministic problem P is polynomially solvable, then it can often be formulated as a 0-1 linear programming problem with a totally unimodular constraints matrix. This is, for instance, the case for a wide class of network flow problems such as shortest path, minimum spanning tree, minimum assignment or minimum cut [1, 18].

Recall (see (15)) that $\overline{\delta}_X^{\lambda} = \delta_X(S_X^{+\lambda}) = F(X,S_X^{+\lambda}) - F^*(S_X^{+\lambda})$. Using the vector of binary variables $\boldsymbol{x}$ representing X, we can see that $F(X,S_X^{+\lambda}) = \sum_{i=1}^n \overline{w}_i^{\lambda} x_i$. Under scenario $S_X^{+\lambda}$ the weight of element e_i is $\overline{w}_i^{\lambda} x_i + \underline{w}_i^{\lambda}(1-x_i)$. So, $F^*(S_X^{+\lambda})$ can be expressed as follows:

$$\begin{array}{ll} \min \sum_{i=1}^{n} [\overline{w}_i^{\lambda} x_i + \underline{w}_i^{\lambda}(1-x_i)]y_i & \\ \mathscr{A}\boldsymbol{y}^T = \boldsymbol{b} & \\ y_i \in \{0,1\} & i = 1,\ldots,n \end{array} \tag{25}$$

We now use the assumption that matrix $\mathscr{A}$ is totally unimodular. Under this assumption (see e.g. [18]) we can replace constraints $y_i \in \{0,1\}$ in (25) with $0 \leq y_i \leq 1$ without changing the cost of an optimal solution to (25). As the result we get the following problem:

$$\begin{array}{ll} \min \sum_{i=1}^{n} [\overline{w}_i^{\lambda} x_i + \underline{w}_i^{\lambda}(1-x_i)]y_i & \\ \mathscr{A}\boldsymbol{y}^T = \boldsymbol{b} & \\ 0 \leq y_i \leq 1 & i = 1,\ldots,n \end{array} \tag{26}$$

We can now construct a dual model to (26). This dual model has a vector of dual variables $\boldsymbol{u}$ associated with the constraints of (26). Denote by $\phi(\boldsymbol{u})$ the objective of the dual and by $D^{\lambda}(\boldsymbol{x})$ the set of feasible dual vectors. So, the dual model is $\max_{\boldsymbol{u} \in D^{\lambda}(\boldsymbol{x})} \phi(\boldsymbol{u})$ and it is linear with respect to both $\boldsymbol{u}$ and $\boldsymbol{x}$ if λ is fixed. Now the strong duality theorem implies:

$$F^*(S_X^{+\lambda}) = \max_{\boldsymbol{u} \in D^{\lambda}(\boldsymbol{x})} \phi(\boldsymbol{u}).$$

Hence

$$\overline{\delta}_X^{\lambda} = \sum_{i=1}^{n} \overline{w}_i^{\lambda} x_i - \max_{\boldsymbol{u} \in D^{\lambda}(\boldsymbol{x})} \phi(\boldsymbol{u}),$$

which together with (22) give

$$\begin{array}{ll} \min \lambda & \\ \sum_{i=1}^{n} \overline{w}_i^{\lambda} x_i - \max_{\boldsymbol{u} \in D^{\lambda}(\boldsymbol{x})} \phi(\boldsymbol{u}) \leq \overline{g}^{1-\lambda} & \\ \mathscr{A}\boldsymbol{x}^T = \boldsymbol{b} & \\ x_i \in \{0,1\} & i = 1,\ldots,n \\ \lambda \in [0,1] & \end{array} \tag{27}$$

We can omit the maximum operator in (27) obtaining the following equivalent model:

$$\begin{array}{ll} \min \lambda & \\ \sum_{i=1}^{n} \overline{w}_i^{\lambda} x_i - \phi(\boldsymbol{u}) \leq \overline{g}^{1-\lambda} & \\ \mathscr{A}\boldsymbol{x}^T = \boldsymbol{b} & \\ \boldsymbol{u} \in D^{\lambda}(\boldsymbol{x}) & \\ x_i \in \{0,1\} & i = 1,\ldots,n \\ \lambda \in [0,1] & \end{array} \tag{28}$$

Assuming that the element weights are trapezoidal fuzzy intervals $\tilde{W}_i = (\underline{w}_i, \overline{w}_i, \alpha_i, \beta_i)$ for all $e_i \in E$, we can substitute $\overline{w}_i^{\lambda} = \overline{w}_i + \beta_i(1-\lambda)$ and $\underline{w}_i^{\lambda} = \underline{w}_i - \alpha_i(1-\lambda)$ in (28). The resulting model will be still not linear because some expressions of the form λx_i may appear. However, we can make (28) linear by replacing all such expressions with additional variables and adding some additional linear constraints. After

this modification, problem (28) will be a mixed integer linear programming one. We will illustrate this method by an example.

Example 2. Consider the following *minimum selecting items* problem. Let $E = \{e_1, \ldots, e_n\}$ be a set of items. The solution set Φ consists of all subsets X of E such that $|X| = p$, where p is a given integer. So, we wish to choose exactly p items among E. Assume that fuzzy interval $\tilde{W}_i = (\underline{w}_i, \overline{w}_i, \alpha_i, \beta_i)$ is given for every $e_i \in E$. We also fix a fuzzy goal $\tilde{G} = (\overline{g}, \beta_g)$. The binary variable $x_i \in \{0,1\}$ indicates whether item e_i is chosen or not. The solution set Φ in this problem can be described by the single constraint $x_1 + x_2 + \cdots + x_n = p$. Obviously, matrix $\mathscr{A} = [1,1,\ldots,1]$ is totally unimodular. The subproblem (26) takes the following form:

$$\begin{array}{ll} \min \sum_{i=1}^{n} [\overline{w}_i^{\lambda} x_i + \underline{w}_i^{\lambda}(1-x_i)]y_i & \\ y_1 + y_2 + \cdots + y_n = p & \\ 0 \le y_i \le 1 & i = 1,\ldots,n \end{array}$$

Assigning dual variable u_0 to the equality constraint and dual variables $u_1, \ldots, u_n$ to constraints $y_i \le 1$, $i = 1,\ldots,n$, we get the following dual model:

$$\begin{array}{ll} \max pu_0 - u_1 - \cdots - u_n & \\ u_0 - u_i \le \overline{w}_i^{\lambda} x_i + \underline{w}_i^{\lambda}(1-x_i) & \\ u_i \ge 0 & i = 1,\ldots,n \end{array}$$

Consequently, $\phi(\boldsymbol{u}) = pu_0 - u_1 - \cdots - u_n$ and set $D^{\lambda}(\boldsymbol{x})$ is described by the constraints of the dual model. We are now ready to design the model using formulation (28). We also substitute $\overline{w}_i^{\lambda} = \overline{w}_i + \beta_i(1-\lambda)$ and $\underline{w}_i^{\lambda} = \underline{w}_i - \alpha_i(1-\lambda)$. After easy computations we get

$$\begin{array}{ll} \min \lambda & \\ \sum_{i=1}^{n} (\overline{w}_i + \beta_i)x_i - \sum_{i=1}^{n} \beta_i \lambda x_i - pu_0 + \sum_{i=1}^{n} u_i \le \overline{g} + \beta_g \lambda & \\ \sum_{i=1}^{n} x_i = p & \\ u_0 - u_i \le (\overline{w}_i - \underline{w}_i + \alpha_i + \beta_i)x_i - (\alpha_i + \beta_i)\lambda x_i - \alpha_i(1-\lambda) + \underline{w}_i & i = 1,\ldots,n \\ u_i \ge 0 & i = 1,\ldots,n \\ \lambda \in [0,1] & \\ x_i \in \{0,1\} & i = 1,\ldots,n \end{array}$$

The obtained model is still not linear. We can, however, substitute $t_i = \lambda x_i$ and add additional linear constraints $t_i - x_i \le 0$, $\lambda - t_i + x_i \le 1$, $-\lambda + t_i \le 0$, $t_i \ge 0$ for all $i = 1,\ldots,n$. This assures that $t_i = \lambda$ if $x_i = 1$ and $t_i = 0$ if $x_i = 0$. The resulting final model will be a mixed integer linear programming one and can be solved by using a standard software. Of course, the same technique can be applied to other problems with totally unimodular constraints matrix. □

5 Conclusions

In this chapter we have discussed a general class of combinatorial optimization problems with fuzzy weights. We have provided an interpretation of such problems in the setting of possibility theory. The possibility and necessity measures allow us to characterize the optimality of solutions and elements and to define a solution concept. This solution concept is an adaptation of the necessary soft optimality first proposed for fuzzy linear programming. In general, every fuzzy problem boils down to solving a small number of interval problems. Every algorithm for computing a deviation interval and a min-max regret solution under interval weights can be easily adopted to solve a more general fuzzy problem. It is enough to apply a standard binary search technique. The complexity of an interval problem depends on the type of the cost function in its deterministic counterpart. In general, the problems with bottleneck cost function are easier to solve than the ones with linear sum cost.

There are some open questions concerning the approach described in this chapter. Most of them refer to the interval uncertainty representation. For instance, the problem of evaluating the necessary optimality of elements is open (its complexity is known only for some particular problems). Also, designing fast algorithms for computing optimal min-max regret solutions is an important subject of further research. For fuzzy problems, the efficiency of the MIP formulation should be investigated. Also, the parametric techniques, which allow us to compute fuzzy deviation intervals and solve the bottleneck problems should be explored more deeply. Finally, if the interval problem is NP-hard, then some heuristics and approximation algorithms for its fuzzy generalization should be designed.

Acknowledgements. The second author was partially supported by Polish Committee for Scientific Research, grant N N111 1464 33.

References

1. Ahuja, R.K., Magnanti, T.L., Orlin, J.B.: Network Flows: Theory, Algorithms, and Applications. Prentice-Hall, Englewood Cliffs (1993)
2. Aissi, H., Bazgan, C., Vanderpooten, D.: Min-max and min-max regret versions of combinatorial optimization problems: A survey. European Journal of Operational Research (2008) doi:10.1016/j.ejor.2008.09.012
3. Aissi, H., Bazgan, C., Vanderpooten, D.: Complexity of the min-max and min-max regret assignment problems. Operations Research Letters 33(6), 634–640 (2005)
4. Aissi, H., Bazgan, C., Vanderpooten, D.: Complexity of the min-max (regret) versions of cut problems. In: Deng, X., Du, D.-Z. (eds.) ISAAC 2005. LNCS, vol. 3827, pp. 789–798. Springer, Heidelberg (2005)
5. Aron, I.D., Hentenryck, P.V.: A constraint satisfaction approach to the robust spanning tree problem with interval data. In: Darwiche, A., Friedman, N. (eds.) UAI 2002, Proceedings of the 18th Conference in Uncertainty in Artificial Intelligence, University of Alberta, Edmonton, Alberta, Canada, August 1-4, pp. 18–25. Morgan Kaufmann, San Francisco (2002)

6. Aron, I.D., Hentenryck, P.V.: On the complexity of the robust spanning tree problem with interval data. Operations Research Letters 32(1), 36–40 (2004)
7. Averbakh, I.: Minmax regret solutions for minimax optimization problems with uncertainty. Operations Research Letters 27(2), 57–65 (2000)
8. Averbakh, I.: On the complexity of a class of combinatorial optimization problems with uncertainty. Mathematical Programming 90, 263–272 (2001)
9. Averbakh, I., Lebedev, V.: On the complexity of minmax regret linear programming. European Journal of Operational Research 160(1), 227–231 (2005)
10. Camerini, P.M.: The minimax spanning tree problem and some extensions. Information Processing Letters 7, 10–14 (1978)
11. Chan, T.: Finding the shortest bottleneck edge in a parametric minimum spanning tree. In: Proceedings of the 16th Annual ACM-SIAM Symposium on Discrete Algorithms, pp. 232–240 (2005)
12. Conde, E.: An improved algorithm for selecting p items with uncertain returns according to the minmax-regret criterion. Mathematical Programming 100(2), 345–353 (2004)
13. Dubois, D., Fargier, H., Galvagnon, V.: On latest starting times and floats in activity networks with ill-known durations. European Journal of Operational Research 147, 266–280 (2003)
14. Dubois, D., Prade, H.: Possibility Theory: An Approach to Computerized Processing of Uncertainty. Plenum Press (1988)
15. Fernandez-Baca, D., Slutzki, G.: Solving parametric problems on trees. Journal of Algorithms 10 (1989)
16. Fortin, J., Zieliński, P., Dubois, D., Fargier, H.: Interval analysis in scheduling. In: van Beek, P. (ed.) CP 2005. LNCS, vol. 3709, pp. 226–240. Springer, Heidelberg (2005)
17. Gabow, H.N., Tarjan, R.E.: Algorithms for two bottleneck optimization problems. Journal of Algorithms 9, 411–417 (1988)
18. Garfinkel, R.S., Nemhauser, G.L.: Integer Programming. John Wiley & Sons, Chichester (1972)
19. Inuiguchi, M.: On possibilistic/fuzzy optimization. In: Melin, P., Castillo, O., Aguilar, L.T., Kacprzyk, J., Pedrycz, W. (eds.) IFSA 2007. LNCS (LNAI), vol. 4529, pp. 351–360. Springer, Heidelberg (2007)
20. Inuiguchi, M., Sakawa, M.: Robust optimization under softness in a fuzzy linear programming problem. International Journal of Approximate Reasonning 18(1-2), 21–34 (1998)
21. Kall, P., Wallace, S.W.: Stochastic Programming. Wiley, Chichester etc. (1994)
22. Kasperski, A.: Discrete Optimization with Interval Data. Minmax Regret and Fuzzy Approach. Studies in Fuzziness and Soft Computing, vol. 228. Springer, Heidelberg (2008)
23. Kasperski, A., Zieliński, P.: An approximation algorithm for interval data minmax regret combinatorial optimization problems. Information Processing Letters 97(5), 177–180 (2006)
24. Kasperski, A., Zieliński, P.: The robust shortest path problem in series-parallel multidigraphs with interval data. Operations Research Letters 34(1), 69–76 (2006)
25. Kasperski, A., Zieliński, P.: On combinatorial optimization problems on matroids with uncertain weights. European Journal of Operational Research 177(2), 851–864 (2007)
26. Kasperski, A., Zieliński, P.: On the existence of an FPTAS for minmax regret combinatorial optimization problems with interval data. Operations Research Letters 35(4), 525–532 (2007)
27. Kasperski, A., Zieliński, P.: Using gradual numbers for solving fuzzy-valued combinatorial optimization problems. In: Melin, P., Castillo, O., Aguilar, L.T., Kacprzyk, J., Pedrycz, W. (eds.) IFSA 2007. LNCS (LNAI), vol. 4529, pp. 656–665. Springer, Heidelberg (2007)

28. Kasperski, A., Zieliński, P.: Minmax regret approach and optimality evaluation in combinatorial optimization problems with interval and fuzzy weights. European Journal of Operational Research (2009) (to appear)
29. Kouvelis, P., Yu, G.: Robust Discrete Optimization and Its Applications. Kluwer Academic Publishers, Boston (1997)
30. Lawler, E.: Combinatorial Optimization: Networks and Matroids. Saunders College Publishing, Fort Worth (1976)
31. Luce, R.D., Raiffa, H.: Games and Decisions. Introduction and Critical Survey. Dover Publications, Inc., New York (1957)
32. Montemanni, R.: A benders decomposition approach for the robust spanning tree problem with interval data. European Journal of Operational Research 174(3), 1479–1490 (2006)
33. Montemanni, R., Gambardella, L.M.: A branch and bound algorithm for the robust spanning tree problem with interval data. European Journal of Operational Research 161(3), 771–779 (2005)
34. Montemanni, R., Gambardella, L.M., Donati, A.V.: A branch and bound algorithm for the robust shortest path problem with interval data. Operations Research Letters 32(3), 225–232 (2004)
35. Papadimitriou, C.H., Steiglitz, K.: Combinatorial Optimization: Algorithms and Complexity. Dover Pub. (1998)
36. Punnen, A.P.: A linear time algorithm for the maximum capacity path problem. European Journal of Operational Research 53, 402–404 (1991)
37. Punnen, A.P.: A fast algorithm for a class of bottleneck problems. Computing 56, 397–401 (1996)
38. Raghavan, S., Ball, M.O., Trichur, V.S.: Bicriteria product design optimization: An efficient solution procedure using and/or trees. Naval Research Logistics 49, 574–592 (2002)
39. Savage, L.J.: The Foundations of Statistics. John Wiley and Sons, New York (1951)
40. Yaman, H., Karasan, O.E., Pinar, M.Ç.: The robust spanning tree problem with interval data. Operations Research Letters 29(1), 31–40 (2001)
41. Young, N., Tarjan, R., Orlin, J.: Faster parametric shortest path minimum-balance algorithms. Networks 21, 205–221 (1991)
42. Zieliński, P.: The computational complexity of the relative robust shortest path problem with interval data. European Journal of Operational Research 158(3), 570–576 (2004)

Stochastic Bottleneck Spanning Tree Problem on a Fuzzy Network

Yue Ge and Hiroaki Ishii

Abstract. This paper considers a fuzzy network version of the stochastic bottleneck spanning tree problem. Existence of each edge is not necessary certain and it is given by a certain value between 0 and 1. 1 means that it exists certainly and 0 means it does not exist. For intermediate numbers, a higher value corresponds to a higher possibility of existence. Furthermore each edge has a random cost independent to other edges. The probability that the maximum burden among these selected edges is not greater than the capacity should be not less than the fixed probability. Under the above setting, we seek a spanning tree minimizing the capacity and maximizing the minimal existence possibility among these selected edges. Since usually there is no spanning tree optimizing two objectives at a time, we derive an efficient solution procedure to obtain a set of some non-dominated spanning tree after defining non-domination of spanning trees. Finally we discuss the further research problems.

Keywords: Fuzzy network, Spanning tree problem, Random variable, Non-domination, Efficient algorithm.

1 Introduction

The minimum spanning tree problem has been well studied and until now many efficient algorithms such as [1, 2, 4, 14, 17] have been proposed. Stochastic versions of them, i.e., spanning tree problems with random edge costs are also considered [5, 6, 7, 8, 9]. The stochastic bottleneck spanning tree problem is one of them and it has been studied [7]. This paper further generalizes it towards a fuzzy network version. Though fuzzy versions of the spanning tree problem are

Yue Ge
Department of Mathematics, Harbin Institute of Technology, Harbin, China
e-mail: gyue2001@hotmail.com

Hiroaki Ishii
Graduate School of Information Science and Technology, Osaka University, Osaka, Japan
e-mail: ishii@ist.osaka-u.ac.jp

W.A. Lodwick & J. Kacprzyk (Eds.): Fuzzy Optimization, STUDFUZZ 254, pp. 313–321.
springerlink.com

considered [10, 12], and bottleneck spanning tree problems with fuzzy random edge costs are also considered [11, 13], fuzzy network version with random edge costs have not previously been considered.

Section 2 formulates our problem and transforms it into an equivalent deterministic problem. Based on the results in Section 2, Section 3 proposes a solution procedure for the problem. Finally Section 4 concludes this paper and discusses further research problems.

2 Problem Formulation

Let $N=(V,E)$ denote a fuzzy network consisting of vertex set $V=\{v_1,v_2,\cdots,v_n\}$ and edge set $E=\{e_1,e_2,\cdots,e_m\}\subset V\times V$. Each edge e_j has random cost c_j and existence possibility μ_j.

Spanning tree $T=(V,S)$ is a partial network of N satisfying the following conditions (see [3] for example).

(1) T has the same vertex set as N;
(2) $S\subseteq E$, $|S|=n-1$ where $|S|$ denotes the cardinality of the set S;
(3) T is connected.

Furthermore, T can be denoted with 0-1 variables x_1, x_2, $\cdots$, x_m as follows.

$$T:\ x_j=1,\ e_j\in S$$
$$x_j=0,\ e_j\notin S$$

Conversely, if a set of edges such as $\{e_j \mid x_j=1,\ j=1,2,\cdots,m\}$ forms a spanning tree of N with vertex set V, then $X=(x_1,x_2,\cdots,x_m)$ is also called a spanning tree of N hereafter in this paper.

For convenience sake we assume that each edge cost c_j is a random variable according to the normal distribution $N(m_j,\sigma_j^2)$ with mean m_j and variance σ_j^2 and they are mutually independent. Under above setting we consider the following bi-criteria problem P.

$$\begin{aligned} P:\ &\text{Minimize}\quad f\\ &\text{Maximize}\quad \min\{\mu_j \mid e_j\in T\}\\ &\text{subject to}\quad \Pr[\max\{c_j \mid e_j\in T\}\le f]\ge\alpha\\ &\qquad\qquad T:\ \text{spanning tree}\end{aligned}$$

The above problem reflects on the construction of a communication network that connects some cities directly or indirectly. If each communication quantity per

unit time between one city and another usually fluctuates randomly, minimizing the maximal capacity necessary for handling these quantities becomes a chance constraint $\Pr[\max\{c_j \mid e_j \in T\} \le f] \ge \alpha$ requiring the probability that the network does not overload should be not less than α (α is constant and assumed to be $1/2 < \alpha < 1$). While μ_j denotes a preference using a line e_j as a communication network and this preference is determined judging from the various factors (construction cost, security, etc.) other than its capacity.

The chance constraint $\Pr[\max\{c_j \mid e_j \in T\} \le f] \ge \alpha$ is transformed into the following deterministic equivalent condition according to [7].

$$\begin{aligned}
&\Pr[\max\{c_j \mid e_j \in T\} \le f] \ge \alpha \\
\Leftrightarrow\ &\Pr[\bigcap_{e_j \in T} \{c_j \le f\}] \ge \alpha \Leftrightarrow \prod_{e_j \in T} \Pr\{c_j \le f\} \ge \alpha \\
\Leftrightarrow\ &\prod_{e_j \in T} F(\frac{f - m_j}{\sigma_j}) \ge \alpha \Leftrightarrow \sum_{e_j \in T} \log F(\frac{f - m_j}{\sigma_j}) \ge \log \alpha \\
\Leftrightarrow\ &\sum_{j=1}^{m} \log F(\frac{f - m_j}{\sigma_j}) x_j \ge \log \alpha
\end{aligned}$$

where F is the probability distribution of the standard normal distribution $N(0,1)$ and $X = (x_1, x_2, \cdots, x_m)$ is the spanning tree corresponding to T.

Since $\min\{\mu_j \mid e_j \in T\} \ge \mu \ \Leftrightarrow\ \mu_j \ge \mu,\ e_j \in T \ \Leftrightarrow\ \mu_j x_j \ge \mu x_j,\ x_j = 0 \text{ or } 1$, therefore objective Maximize $\min\{\mu_j \mid e_j \in T\}$ is equivalent to

$$\begin{aligned}
&\text{Maximize} \quad \mu \\
&\text{subject to} \quad \mu_j x_j \ge \mu x_j,\ x_j = 0 \text{ or } 1,\ j = 1, 2, \cdots, m. \\
&\qquad\qquad X:\ \text{spanning tree}
\end{aligned}$$

Thus P is equivalent to the following deterministic problem $\bar{P}$.

$$\begin{aligned}
\bar{P}:\ &\text{Minimize} \quad f \\
&\text{Maximize} \quad \mu \\
&\text{subject to} \quad \sum_{j=1}^{m} \log F(\frac{f - m_j}{\sigma_j}) x_j \ge \log \alpha \\
&\qquad\qquad \mu_j x_j \ge \mu x_j,\ x_j = 0 \text{ or } 1,\ j = 1, 2, \cdots, m. \\
&\qquad\qquad X:\ \text{spanning tree}
\end{aligned}$$

If we fix a decision variable μ, we restrict edges that we can select, i.e., only edges whose existence possibilities are over μ are candidates to select as those of

a spanning tree. First we define sub-network $N_\mu = (V, E_\mu)$ as a crisp network consisting of vertex set V and edge set $E_\mu = \{e_j \mid \mu_j \geq \mu,\ e_j \in E\}$. This restricted version of problem $\bar{P}$ on a network $N_\mu = (V, E_\mu)$ is now denoted by $\bar{P}_\mu$. Then $\bar{P}_\mu$ is the deterministic equivalent problem of stochastic bottleneck spanning tree problem treated in [7]. But in order to make $\bar{P}_\mu$ meaningful, we must find the upper bound of μ. This bound μ^U is the optimal value of the bottleneck spanning tree problem when we consider $\mu_1,\ \mu_2,\ \cdots,\ \mu_m$ as edge costs, i.e.,

$$\mu^U = \max[\min\{\mu_j \mid e_j \in T\},\ T \text{ is a spanning tree of } N].$$

The next section shows the solution procedure for problem $\bar{P}_\mu$ by assuming $\mu \leq \mu^U$ and that for our bi-criteria problem.

3 Solution Procedure

For different fixed μ, in order to solve $\bar{P}_\mu$, we consider the following subproblem P_μ^q with parameter q $(q > 0)$, when we fix $f = q$.

$$\begin{aligned} P_\mu^q:\ &\text{Maximize} \sum_{e_j \in E_\mu} \log F\left(\frac{q - m_j}{\sigma_j}\right) x_j \\ &\text{subject to}\quad x_j = 0 \text{ or } 1,\ j \in \{j \mid e_j \in E_\mu\} \\ &\qquad\qquad X_\mu:\ \text{spanning tree corresponding to } N_\mu \end{aligned}$$

For the convenience sake we assume $E_\mu = \{e_1, e_2, \cdots, e_t\}$ and $t = \mid E_\mu \mid$ by renumbering indices of edges if necessary, accordingly we denote $X_\mu = (x_{\mu 1}, x_{\mu 2}, \cdots, x_{\mu t})$.

P_μ^q is the ordinary maximum spanning tree problem and can be solved by algorithms such as [1, 2, 4, 14, 16, 17].

Let X_μ^q denote an optimal solution of P_μ^q and Z_μ^q its value. Then we have the following property due to [7] and $E_\mu \subseteq E_{\mu'}$ for $\mu \geq \mu'$.

Property 1. Z_μ^q is an increasing function of q and decreasing function of μ.

Further let (X_μ^*, f_μ^*) denote an optimal solution of $\overline{P}_\mu$. The following theorem shows a relation between $\overline{P}_\mu$ and P_μ^q.

Theorem 1

(1) $Z_\mu^q > \log\alpha \Leftrightarrow f_\mu^* < q$;

(2) $Z_\mu^q = \log\alpha \Leftrightarrow f_\mu^* = q$;

(3) $Z_\mu^q < \log\alpha \Leftrightarrow f_\mu^* > q$.

Proof. Due to Theorem 1 in [7].

Each algorithm for an optimal spanning tree is determined by the order of edge costs, and the order of edge costs $\log F(\frac{q-m_j}{\sigma_j})$ is same as the order of $c_j(q) \triangleq (q-m_j)/\sigma_j$, since F is non-decreasing function. Define

$$g_k^\mu(q) \triangleq \max\left\{c_j(q) \mid e_j = (v_k, v_l) \in E_\mu\right\}$$

$$g_k(q) \triangleq \max\left\{c_j(q) \mid e_j = (v_k, v_l) \in E\right\}$$

for each vertex $v_k \in V$.

Then each $g_k^\mu(q)$ and $g_k(q)$ are piecewise linear functions of q because each $c_j(q)$ is linear function of q. Merging these breakpoints of $g_k^\mu(q)$ and those of $g_k(q)$, let the results be

$$-\infty = q_0^\mu < q_1^\mu < \cdots < q_{s(\mu)}^\mu < q_{s(\mu)+1}^\mu = +\infty$$

$$-\infty = q_0 < q_1 < \cdots < q_s < q_{s+1} = +\infty$$

where $s(\mu)$ and s are the numbers of different points for $g_k^\mu(q)$ and $g_k(q)$, respectively.

Note that q_1^μ, $\cdots$, $q_{s(\mu)}^\mu$ is a subsequence of q_1, $\cdots$, q_s. From above results, we have the following Algorithm 1 for $\overline{P}_\mu$. The algorithm is based on Sollin's algorithm [1, p.179] for the minimum spanning tree problem. Z_μ^q is calculated by using algorithms of [2] or [17] and all the breakpoints are calculated by Megiddo's method [15, Appendix]. In the algorithm, L_μ and U_μ denote the lower and the upper bound of f_μ^*.

Algorithm 1

Step 1. Find the first h such that $Z_{\mu}^{q_h^{\mu}} > \log\alpha$ and set $L_{\mu} \leftarrow q_{h-1}^{\mu}$, $U_{\mu} \leftarrow q_h^{\mu}$. Go to Step 2.

Step 2. From all $e_j \in E_{\mu}$ giving $g_k^{\mu}(q)$ for each $v_k \in V$ on the interval $[L_{\mu}, U_{\mu}]$, construct spanning forest $(T_1, T_2, \cdots, T_a)$ of N, where T_i is a subtree and a is their number. Go to Step 3.

Step 3. If $a = 1$, construct optimal spanning tree X_{μ}^* as follows.

$$X_{\mu}^*: \; x_{\mu j}^* = 1, \; e_j \in T_1$$
$$x_{\mu j}^* = 0, \; e_j \notin T_1$$

Then the corresponding optimal value f_{μ}^* is obtained.

$$f_{\mu}^* = \{f_{\mu}^* \mid \sum_{j=1}^{t} \log F(\frac{f_{\mu}^* - m_j}{\sigma_j}) x_{\mu j}^* = \log\alpha\}.$$

If $a \neq 1$, then go to Step 4.

Step 4. For each T_i, calculate

$$h_i^{\mu}(q) = \max\{c_j(q) \mid e_j \text{ connect } T_i \text{ with other subtree}, e_j \in E_{\mu}\}$$

and find all breakpoints of $h_1^{\mu}(q), \; \cdots, \; h_a^{\mu}(q)$ on the interval $[L_{\mu}, U_{\mu}]$. Merge them and let the result be

$$L_{\mu} = \overline{q}_0 < \overline{q}_1 < \cdots < \overline{q}_r = U_{\mu}.$$

Go to Step 5.

Step 5. Find the first $\overline{q}_b$ such that $Z_{\mu}^{\overline{q}_b} > \log\alpha$ and set $L_{\mu} \leftarrow \overline{q}_{b-1}$, $U_{\mu} \leftarrow \overline{q}_b$. Next update the forest $(T_1, T_2, \cdots, T_a)$ by adding edges giving $h_i^{\mu}(q)$ for $q \in [L_{\mu}, U_{\mu}]$ to the current forest. Return to Step 3.

Theorem 2. The above Algorithm finds an optimal solution (X_{μ}^*, f_{μ}^*) in at most $O(t \log^2 n \log\log n)$ computational time if we can solve the following equation in less than $O(t \log^2 n \log\log n)$ computational time.

$$\sum_{j=1}^{t} \log F(\frac{f_{\mu}^* - m_j}{\sigma_j}) x_{\mu j}^* = \log\alpha.$$

Proof. Due to Theorem 2 in [7].

Now we turn to solve the original bi-criteria spanning tree problem. Usually there is no spanning tree optimizing both objectives at a time and so we propose an algorithm to find some non-dominated spanning trees. First we define non-domination.

Definition 1 (Non-Dominated Spanning Tree)

For μ, μ' such that $\mu > \mu'$, if $f_{\mu}^{*} = f_{\mu'}^{*}$, we say spanning tree X_{μ}^{*} dominates spanning tree $X_{\mu'}^{*}$. If there does not exist a spanning tree dominating X_{μ}^{*}, X_{μ}^{*} is called a non-dominated spanning tree (*NDS* for short).

Next we sort $\mu_1, \mu_2, \cdots, \mu_m$ not greater than μ^U and let the result be $\mu^1 < \mu^2 < \cdots < \mu^d$, where d is the number of different values of μ.

We are now ready to describe our algorithm for problem $\overline{P}$.

Algorithm 2

Step 1. Set $\mu = \mu^1$, $NDS = \phi$ and $l = 1$. Go to Step 2.

Step 2. Solve $\overline{P}_{\mu}$ using ALGORITHM 1 and find the optimal solution $(X_{\mu}^{*}, f_{\mu}^{*})$. If X_{μ}^{*} dominates some spanning trees in *NDS,* then update *NDS* by deleting them and adding X_{μ}^{*}. Otherwise update it by adding X_{μ}^{*}. Go to Step 3.

Step 3. Set $l = l + 1$, If $l = d + 1$, then terminate and *NDS* is the set of some non-dominated spanning trees. Otherwise set $\mu = \mu^l$ and return to Step 2.

Note that L_{μ} increases as l increases and so we need check each breakpoint at most once.

Theorem 3. Algorithm 2 finds some non-dominated spanning trees for our problem in at most $O(m^2 \log^2 n \log\log n)$ computational time.

Proof. Validity is clear from the fact that it check all possibilities of μ and the validity of Algorithm 1 [7, Theorem 2]. So we only consider the computational complexity of Algorithm 2. Sorting $\mu_1, \mu_2, \cdots, \mu_m$ takes $O(m \log n)$ computational time, calculating μ^U takes $O(m \log n \log\log n)$ computational time. So preprocessing does not dominate computational time. Since d is $O(m)$, basically Algorithm 2 solves $O(m)$ number of $\overline{P}_{\mu}$ by using Algorithm 1. Therefore Algorithm 2 takes $O(m^2 \log^2 n \log\log n)$ computational time since each execution of

Algorithm 1 takes $O(m \log^2 n \log\log n)$ computational time. But usually each spanning tree problem need not be solved from scratch. So at most $O(m^2 \log^2 n \log\log n)$ computational time is needed.

4 Conclusion

This paper has investigated the stochastic bottleneck spanning tree problem on a fuzzy network and proposed an efficient algorithm for seeking non-dominated spanning trees. Refinement of the algorithm is necessary. In particular we need not solve $\overline{P}_{\mu}$ from scratch except the first one $\overline{P}_{\mu^1}$ in ALGORITHM 2. We are currently working on fuzzy network version of many other combinatorial optimization problems such as shortest path problems since in many actual situations, these problems should be considered on fuzzy network structures.

References

[1] Berge, C., Ghouila-Hour, A.: Programming, Games and Transportation Networks. Wiley, New York (1965)
[2] Cheriton, D., Tarjan, R.E.: Finding minimum spanning Trees. SIAM Journal on Computing 5, 724–742 (1976)
[3] Christofides, H.N.: Graph Theory: An Algorithmic Approach. Academic Press, New York (1975)
[4] Gabow, H.N.: A good algorithm for a smallest spanning trees with an edge degree constraint. Networks 8, 201–208 (1978)
[5] Geetha, S., Nair, K.P.K.: On stochastic spanning tree problem. Networks 23, 675–679 (1993)
[6] Ishii, H., Matsutomi, T.: Confidence regional method of stochastic spanning tree problem. Mathematical Computer Modeling 22, 77–82 (1995)
[7] Ishii, H., Nishida, T.: Stochastic bottleneck spanning tree problem. Networks 13, 443–449 (1983)
[8] Ishii, H., Shiode, S., Nishida, T.: Chance constrained spanning tree problem. Journal of the Operations Research Society of Japan 24, 147–157 (1981)
[9] Ishii, H., Shiode, S., Nishida, T., Namasuya, Y.: Stochastic spanning tree problem. Discrete Applied Mathematics 3, 263–273 (1981)
[10] Itoh, T., Ishii, H.: An approach based on necessity measure to the fuzzy spanning tree problems. Journal of the Operations Research Society of Japan 39, 247–257 (1996) (in Japanese)
[11] Katagiri, H., Ishii, H.: Chance constrained bottleneck spanning tree problem with fuzzy random edge costs. Journal of the Operation Research Society of Japan 43, 128–137 (2000)
[12] Katagiri, H., Ishii, H., Sakawa, M.: A fuzzy spanning tree problem based on possibility and necessity. Journal of Japan Society for Fuzzy Theory and Systems 12, 797–805 (2000) (in Japanese)

[13] Katagiri, H., Sakawa, M., Ishii, H.: Fuzzy random bottleneck spanning tree problems using possibility and necessity measures. European Journal of Operational Research 152, 88–95 (2004)
[14] Kruskal, J.B.: On the shortest spanning subtree of a graph and traveling salesman problem. Proceedings of American Mathematical Society 7, 48–50 (1956)
[15] Megiddo, N.: Combinatorial optimization with rational objective functions. Mathematics of Operations Research 4, 414–424 (1979)
[16] Prim, R.C.: Shortest Connection Networks and Some Generalizations. Bell System Technical Journal 36, 1389–1401 (1957)
[17] Yao, A.C.: An $O(|E|\log\log|V|)$ algorithm for finding minimum spanning trees. Information Processing Letters 4, 21–23 (1975)

The Use of Fuzzy Numbers in Practical Project Planning and Control

Dorota Kuchta

Abstract. The paper proposes how to use fuzzy numbers in project planning and control in such a way that it would have a chance to be used in practice. The method is destined for all the projects, but especially for those where in the initial phase the knowledge about the project is very incomplete and is made stepwise more precise during the project execution, also for those in which initial assumptions about the project execution are due to later changes.

Keywords: fuzzy number, earned value, project scheduling, project control.

1 Introduction

It is generally agreed upon the fact that each project is by its very nature and definition to some degree unique, and that today the project environment is so unstable, that project planning is biased by a large degree of uncertainty. Thus, it is essential to take this into account both in project planning and control, and to do this in a way that comprises all the projects, also those where the degree of uncertainty is very high, like some IT as well as research and development projects. It is generally accepted that fuzzy numbers are a good tool to express uncertainty, limited knowledge. Thus, their use in project management has been frequently proposed in the literature, however, it seems that the practical implementation of those proposals is almost non-existent. In this paper we want to propose such an application of fuzzy numbers to project management which is on one hand simple enough to be actually used, and on the other hand is comprehensive enough to be

Dorota Kuchta
Institute of Organisation and Management
Wroclaw University of Technology
Ul. Smoluchowskiego 25
50-371 Wrocław
Poland
e-mail: dorota.kuchta@pwr.wroc.pl

W.A. Lodwick & J. Kacprzyk (Eds.): Fuzzy Optimization, STUDFUZZ 254, pp. 323–339.
springerlink.com

useful for most project, also those innovative, very unique, with a high degree of risk and uncertainty

2 State of the Art of Project Control Versus Project Planning

We start by presenting what has already been done in the project management research which can be taken as the basis of our method. We will present both some elements of the use of fuzzy numbers in project management and the crisp approaches, which we want to modify applying fuzzy numbers to them.

2.1 Project Planning

We will not talk here about project planning based on crisp numbers, because we claim that such planning is in may projects simply impossible or deceiving. Many projects (examples will be presented later on) are in the planning phase unknown to such a degree, that each crisp estimate of duration and cost is not realistic and does not allow a good project management, especially risk management. In our opinion at least some projects require taking in the planning phase non-crisp information into account.

Fuzzy numbers have been extensively applied to the project planning (e.g. [2,11]). Most proposals are based on fuzzy optimization, where the objective is to minimize the project execution time, and the activities duration are given in the form of fuzzy numbers. The problem is that if the activities duration times are fuzzy, the notions like critical path, activities floats etc. are not unequivocally defined, and new notions like "the possibility degree to which this or that project path will be critical" have to be introduced, which seem to have little chance to be applied in practice. In this paper we propose not to consider them and to restrict ourselves to this part of the application of fuzzy optimization to project planning which is the simplest possible and yet useful from the practical point of view. What has to be retained of the vast spectrum of proposals of the application to fuzzy numbers to project planning is, in the author's opinion, the simplest and most appealing practical possibility: the possibility to model project activities planned duration and planned cost in the form of fuzzy numbers and to obtain all the spectrum of possible planned project durations and cost values.

Let n be thus the number of project activities. Let us assume that all the constraints imposed on the activities (like precedence relation, resource limits etc.) are known and are taken into account whenever the optimal makespan of the project is determined, let us denote the set of constraints known in the moment of planning as $\Re_0$. Let $\tilde{D}_i \ (i=1,...,n)$ be the fuzzy planned duration of each activity and $\tilde{C}_i \ (i=1,...,n)$ the fuzzy planned cost. To facilitate the application, we propose to use only trapezoidal fuzzy numbers, represented unequivocally by four parameters. Thus, each fuzzy number $\tilde{A}$ will be unequivocally represented

by four numbers $\left(a^1, a^2, a^3, a^4\right)$, where $a^1 \le a^2 \le a^3 \le a^4$ such that each interval $P^t = \left[a^2 - t\left(a^2 - a^1\right), a^3 + t\left(a^4 - a3\right)\right]$ represents possible values of the magnitude represented by $\tilde{A}$ for the uncertainty level $t \in [0,1]$. If we assume the highest level of uncertainty, i.e. t=1, we have $P^1 = \left[a^1, a^4\right]$, thus we admit a large spectrum of possible values, one of each the magnitude represented by $\tilde{A}$ will eventually take. If we assume the lowest uncertainty value, i.e. t=0, we take into account the narrow interval (which may be reduced to one single point) $P^0 = \left[a^1, a^4\right]$ as the set of possible values of the magnitude represented by $\tilde{A}$. Addition of trapezoidal fuzzy numbers can be defined as

$$\tilde{A} + \tilde{B} = \left(a^1, a^2, a^3, a^4\right) + \left(b^1, b^2, b^3, b^4\right) = \left(a^1 + b^1, a^2 + b^2, a^3 + b^3, a^4 + b^4\right)$$

With such representation, we can do calculation which would be easy to implement in the practice and will lead to information important to project managers. We can calculate the planned project fuzzy cost simply as $P\tilde{C} = \sum_{i=1}^{n} \tilde{C}_i$ and also the planned project duration. As far as the latter is concerned, the summation of $\tilde{D}_i \ (i = 1,...,n)$ is done over those activities which belong to the critical path (or one of critical paths) and the set of activities forming the critical path may change according to the crisp values of activities duration. That is why the easiest approach from the practical point of view is to choose a view values of the uncertainty level $t \in [0,1]$, let us say an increasing sequence $T = \{t_1 = 0, t_2, ..., t_{m-1}, t_m = 1\}$, and for each j=1,…,m calculate the optimistic and pessimistic project duration ($OPD\left(t_j\right)$ and $PPD\left(t_j\right)$, respectively for the given uncertainty level, using optimisation methods for determining the shortest project makespan under the given constraints. $OPD\left(t_j\right)$ will be calculated by assuming for the project activities indexed by i=1,…,n crisp durations $d_i^2 - t_j\left(d_i^2 - d_i^1\right)$, and $PPD\left(t_j\right)$ with crisp durations $d_i^3 + t_j\left(d_i^4 - d_i^3\right)$, taking into account $\Re_0$. The set of couples $\left\{\left(OPD\left(t_j\right), PPD\left(t_j\right)\right)\right\}_{j=1}^{m}$ gives the project manager an approximation of possible project durations. We will called this set of couples fuzzy project duration $P\tilde{D}$. Of course, in special cases where the critical path will be the same independently of which crisp values from the ranges predicted in the activities fuzzy duration values occur (an example will be presented further on), we can do a simple summation of the corresponding trapezoidal fuzzy numbers and get a "normal" trapezoidal fuzzy number as $P\tilde{D}$.

2.2 Project Control

The problem of project control is treated in literature almost exclusively for crisp planned duration and cost values. The use of fuzzy numbers in project control, described in [6] and [8], refers to crisp planned values and fuzzy numbers are used to judge to which degree the actual values can be treated as planned or as "good" o "bad". The only project control approach for fuzzy activities duration values, presented in [5], is very general, it has to be made more specific to order to have a chance of being applied in practice.

The method of controlling a project which has been planned (in terms of activities duration and cost) in terms of fuzzy numbers, which will be presented in this paper, will inevitably be based on the crisp approach to project control, thus a short critical review of this approach will be presented.

The basic approach for the control of project cost, but also time and scope, is called Earned Value Method. The method is presented, in its recently modified versions, among other in [1,7,9]. Its idea is as follows: We have a project with *n* activities and a set of constraints $\Re_0$ imposed on the activities, for each activity $i = 1,...n$ there is defined the corresponding work to be done W_i expressed in some physical units like meters, tons or hours, the planned cost of each unit UC_i and the planned duration of the activity D_i. Of course, the ratio W_i / D_i is the planned work efficiency for the *i*-th activity (it will be denoted as E_i), and the planned cost of the activity, denoted as C_i, will be equal to $W_i \cdot UC_i$. W_i ($i = 1,...n$) is treated as fixed, changes in this parameter mean changes in the scope of the project and is explicitly not taken into account in the Earned Value Method. Applying the normal optimisation methods which allow to find a schedule, taking into account $\Re_0$, which gives the shortest possible project duration, we know the project planned duration *PD*, and adding up the C_i of all the activities we know the project planned cost *PC*. Now, the idea of the method is that we do not treat *PD* and *PC* as fixed values. On the contrary, we acknowledge that in the planning phase we will almost certainly be wrong in their estimation and that we will be the better in estimating those two values the more the project is advanced. This is because of various reasons, linked to the newness and uncertainty inherent in each project, to human factors, communication problems, changing environment etc.

This is why once the project realisation is started, at fixed time intervals during the project execution, let us number them $l = 1,..., L$, where *L* is unknown till the actual end of the project and is the number of the last control moment, the planned duration *PD* and the planned cost *PC* of the whole project will be reestimated, recalculated. Let us use PD_0 and PC_0 as alternative symbols for *PD* and *PC*, they express the result of time and cost estimation in the planning phase. PD_l and PC_l, $l = 1,..., L$, will be the same magnitudes reestimated al the consecutive

control moments, when our knowledge about the project is usually higher than at the beginning, let it be only because of the fact that the project has already be executed to some extent which always gives some insight into the project and its environment. Normally, the greater *l*, the closer the values PD_l and PC_l should be to the actual project duration and cost, *AD* and *AC* respectively, known only once the project is finished - although there may be cases where because of some unexpected event near the end of the project this will not be true.

As far as the consecutive values of PD_l and PC_l, $l = 1,...,L$, are concerned, we can write the following obvious formulae:

$$PD_l = ADB_l + RD_l \tag{1}$$

$$PC_l = ACB_l + RC_l \tag{2}$$

where ADB_l stands for actual duration of the project before the control moment *l*, thus in fact for the time that has elapsed from the actual beginning of the project till moment *l*, RD_l stands for the remaining duration, thus the time after moment *l* which we think, according to our knowledge at moment *l*, is needed to terminate the project, ACB_l is the actual cost of the project incurred before moment *l*, and RC_l is the remaining cost, thus the cost that we think, again according to our knowledge at moment *l*, is still needed to be incurred after moment *l* in order to terminate the project.

Formulae (1) and (2) help to structure the problem, but do not facilitate the solution: although ADB_l and ACB_l at control moment l are known exactly, RD_l and RC_l are not and must be the subject of replanning and reestimation on the basis of the knowledge we possess at control moment *l*.

The various approaches to estimate RD_l and RC_l proposed in the literature ([7,9]) have the following common feature: they are based on the project "behaviour" up to moment *l*, trying to express this behaviour, the history of the past project realisation, in the form of synthetic indicators referring to the whole project, not to individual activities or even group of activities. Now we will present the idea of those indicators.

The Earned Value indicators whose aim is to show how the project has been performed up to control moment *l*, are calculated on the basis of the following values:

- BCWS(*l*) (Budgeted Cost of Work Scheduled up to moment *l*), equal to $\sum_{i=1}^{n} pp_i(l) \cdot W_i \cdot CU_i$, where $pp_i(l)$ is the portion of work W_i (expressed as positive fraction not greater than 1) that was planned to be finished before control moment *l*;

- BCWP(l) (Budgeted Cost of Work Performed up to moment l), equal to $\sum_{i=1}^{n} ap_i(l) \cdot W_i \cdot CU_i$, where $ap_i(l)$ is the portion of work W_i that has actually been done before control moment l;
- ACWP(l) (Actual Cost of Work Performed up to moment l) – the actual cost incurred in the project till the control moment l, which is taken from actual accounting documents but can be expressed in an analogous way: $\sum_{i=1}^{n} ap_i(l) \cdot W_i \cdot ACU_i$, where ACU_i is the average actual cost of performing one unit of W_i before control moment l.

Of course, we can also calculate BCWS(l,i), BCWP(l,i), ACWP(l,i) for individual activities (the corresponding formulae are then identical to the above listed ones, but with the summation symbol), but usually, for simplicity and because of lack of time synthetic values for the whole project are calculated. It should be mentioned that BCWP(l) is called earned value in control moment l and is interpreted in two ways:

- The cost we should have incurred according to plan for the work actually performed so far in the project.
- The amount of work that has been actually performed so far in the project (measured not in physical units, but in planned unitary cost, which is not an obvious measurement of work performed and is used to facilitate things, but causes some problems, which will be illustrated later on).

At the same time BCWS(l) is interpreted as the amount of work that according to plan should have been accomplished up to moment l. Again, the amount of work is measured here in monetary units, in planned unitary cost, which may be a problematic matter.

Lipke [7] introduced the notion of earned schedule at control moment l, ES(l), equal to the moment corresponding to s-th control point[1], when BCWS(s)=BCWP(l), thus to the moment when we should or should have accomplished the work we have accomplished up to moment l. Then, according to our notation, ES(l)= ADB_s.

Then the history of the project up to control moment l is estimated, using the following indicators:

- Cost Performance Index in control moment l: $CPI(l) = \frac{BCWP(l)}{ACWP(l)}$
- Schedule Performance Index in control moment l: $SPI(l) = \frac{BCWP(l)}{BCWS(l)}$

[1] For simplicity reason we assume that such an s from the set of control points exists, otherwise some insignificant modifications would have been necessary, because the moment we are interested in might fall between two control points.

- Modified Schedule Performance Index in control moment *l*: $MSPI(l) = \frac{ES(l)}{ADB_l}$.

  The interpretation of the indices is generally such that if they are smaller than 1, something is wrong with the project, if they are greater than 1, something is good with the project, value 1 means that something is as it was planned. This something is project cost in case of *CPI(l)* and the pace of project work accomplishment in case of *SPI(l)* and *MSPI(l).*

As far as formulae (1) and (2) are concerned, they are used by applying for RD_l and

RC_l the following general formulae

$$RD_l = (PD - ADB_l) \cdot DI_l \tag{3}$$

$$RC_l = (PC - ACB_l) \cdot CI_l \tag{4}$$

where DI_l and CI_l are, respectively, some duration/cost indicators equal to or based on *SPI(l)* and *MSPI(l)* (in case of DI_l) and *CPI(l)* and *SPI(l)* (in case of CI_l). The underlying philosophy is as follows: for formula (4), if so far for the work accomplished we paid twice as much as we had planned (which is expressed by *CPI(l)*), we can assume that we will pay twice as much as we have planned for the rest of the project (in this case CI_l would be equal to *CPI(l)*) , and maybe if we additionally have some delay in the amount of work accomplished, expressed by *SPI(l)*, this will have some (unfortunately, unspecified more clearly) influence on the total cots of the project (that is why sometimes CI_l is based both on *CPI(l)* and *SPI(l)*). Similarly for the project duration (formula (3)): if so far we have accomplished half of the work that has been planned (which will be shown by *SPI(l)*), or we have done so far the work which should have been ready two months ago (expressed by *MSPI(l)*), we can assume that the same degree of lateness will apply to the rest of the project.

The problem with such a philosophy is that it works only sometimes. This is said clearly in the corresponding papers, especially in [9], where the authors summarize various variants of formula (4). They equip each variant with an assumption of the type: "if the rest of the project follows the *SPI (CPI, SPI*CPI* etc.*)* pattern". Well, how should we know and what this means? There may be projects where we can say that we will probably continue to work at the same pace and at the same cost as before, but there is certainly a large class of project where such a statement would be completely impossible, and in most project at least some of yet non-started activities will certainly behave in a manner not at all connected to the behaviour of the already accomplished ones. What is more, the indicators *SPI(l), CPI(l)* and *MSPI(l)* may be deceiving, for at least four reasons:

- They are based on the information about the amount of work expressed in planned cost, thus 2 physical units planned to be made by higher paid workers mean more work than then the same 2 units planned to be made by lower paid workers, whereas in both cases the actual amount of work is the same. If we simply change the order of the two activities versus the planned order, we will get the information that we have done more (or less) work than planned, which will not be true;
- They try to combine the information concerning all the activities into one indicator, if one activity is very late (or much more expensive) versus the plan, but another one very early (or much cheaper) versus the plan, we may get the information that everything has been going on exactly as planned, because the two variances may compensate each other. The worst thing will happen if the activity with problems (late, more expensive), hidden by the situation of other activities, is the one which will influence the behaviour of the future activities (e.g. because some of the non-started activities will be performed by the same subcontractor). In such a case the information delivered by the Earned Value Method will be completely deceiving;
- They are based entirely on the project history and do not take into account any new pieces of information about the project future which where not available in the planning phase (like new prices, a different subcontractor than the one which was planned etc.)
- The Earned Value Method is considered in a complete detachment from the reactive scheduling problem – i.e. rescheduling in reaction to what has happened in the project so far, treated in many papers (e.g.[10]). It is as if rescheduling was considered as something to be avoided, because it is difficult? Not possible in practice? In our opinion the project management supporting systems offer the possibility to find a new optimal or sub-optimal schedule in each control model l (with PD_l as the objective function, whose possible values will be of interest to us) taking into account the available information about the project history up to control moment l, and each project control without reactive scheduling does not make much sense. To justify this statement, let us start with the following example:

Example 1. The project consists of three activities. They have the following characteristics: $D_1 = D_2 = 1; D_3 = 2$, $W_1 = W_2 = 10; W_3 = 20$, $CU_1 = CU_2 = CU_3 = 1$. It has been planned that the 3. activity will be executed in the 1. and 2. time unit, activity 1 in the first time unit and activity 2 in the 2. time unit, thus PD=2. The first control moment (l=1) is the end of the first time unit. It is then stated that in the first time unit we have executed and finished the 1. and 2. activity, activity 3 has not been started, the unitary cost has been as planned. It is easy to check than we have BCWS(1)=BCWP(1)=ACWP(1)=20, all the indices discussed above will be equal to 1, formula (1) will give $PD_1 = PD_0 = 2$, and in reality simple rescheduling (determining the optimal project duration under the new conditions)

would show that the shortest possible project duration according to the knowledge in the 1st control moment is 3 (the 3. activity has still to be executed entirely and its planned duration is 2 time units).

Another example justifying the statement comes from [4], where it is described in detail. The project goal is to introduce a new product into the market. The project comprises the following activities: specifying the product on the basis of the market's functional requirements, designing the product, building the prototype, testing the prototype, manufacturing the product, marketing the product to potential customers. Hardie [4] gives a good justification of the fact that the initial schedule, made at the planning moment, is bound to be changed in this type of project, and what is more, not just changed because the duration and cost of each activity may change, but also the scope (the work to be done) is very difficult to estimate at the beginning. And what is even more, Hardie [4] justifies that some activities may have to be repeated, and in the planning phase we even cannot say which of them, when, how often and to which extent. If in the testing phase an error in the design is detected, the designing activity has to be repeated, the same may happen if once the product is on the market, the customers are not satisfied with its functions, and also in the manufacturing phase, if some technical problems are detected – then the product will have to be redesigned and nobody is able to tell exactly in the planning phase whether it will be necessary and how much time and cost this will require. A very similar situation exists in most IT projects, where it is only during the project execution that the final functionality of the project product takes shape, in cooperation with the customer.

2.3 Complex Project Management – Planning and Control

In our opinion the Earned Value Method in its classical (presented in this section) form, combined with planning based on crisp activities cost and duration planned values and detached from reactive scheduling, may be useful only for a rather narrow class of projects, it what is more, it may sometimes be difficult to decide whether our project belongs to this class. For such projects as described in the previous section, we claim that:

- Planning using crisp values does not make sense, because we know that almost no value in the planning phase is known exactly, that almost each project parameter in the planning phase might be much more truthfully described in the form of fuzzy numbers. Using crisp planning values together with the classical form of Earned Value Method may to situation like the following one:

Example 2. The 1. activity has the planned cost equal to the trapezoidal fuzzy number (9,10,10,11), the 2. activity – to the trapezoidal fuzzy number (1,10,10,20). Thus, the knowledge about the cost of the 2. activity is much less precise than it is for the 1. activity. If we use crisp planning, we take for both activities the planned cost 10. The activities should be executed one after

another. If the execution of the 1. activity costs in reality 13 and the control moment falls after its termination, we will – according to the Earned Value principles – conclude that also the second activity will cost 13 (as the first one was 30% more expensive in the crisp approach, the same is assumed for the follwoingone). In the crisp approach this will give the information: cost overrun in case of both activities. In fact, a cost negative variance with respect to the plan has occurred only for the 1. activity, because value 13 falls outside the interval of our prevision for this activity. In case of activity 2, on the other hand, value 13 does not mean a negative cost variance – it was known that such a value might occur for this activity. Treating in the control moment after the termination of the 1. activity value 13 as an important information about the future, i.e. about the second activity, may be very deceiving, because first of all 13 is not a new information, we knew before the cost of the 2. activity might be 13, and what is more, we may lose from our sight another piece of information, possibly essential: if we were completely wrong in estimating the cost of the 1. activity, we may have been wrong in estimation the cost of the 2. activity, which might also fall outs the range predicted by us and be equal e.g. to 22. Everything depends of course on the specific situation, but it seems that project planning and control based on crisp estimations $C_1 = C_2 = 10, PC_0 = 20$ and $PC_1 = 26$ are of limited use in view of the actual information available in the planning phase: $\tilde{C}_1 = (9,10,10,11), \tilde{C}_2 = (1,10,10,20), P\tilde{C}_0 = (10,20,20,31)$.

- The method of project control should also be based on non-crisp values, thus we should calculate in each control moment l $P\tilde{D}_l$ and $P\tilde{C}_l$, equal respectively to $\left(pd_l^1, pd_l^2, pd_l^3, pd_l^4\right)$ and $\left(pc_l^1, pc_l^2, pc_l^3, pc_l^4\right)$, although in most cases he "degree of fuzziness", i.e. the differences $pd_l^4 - pd_l^1$ and $pc_l^4 - pc_l^1$, should get smaller the greater *l*. However, sometimes it will not be the case: sometimes, like in the product design, manufacturing and marketing project described in the previous section, we may only later during the project execution discover how little we know about a certain activity. In such a case, the $P\tilde{D}_l$ and $P\tilde{C}_l$ should reflect the present state of knowledge, even if it is "more fuzzy" than in previous control moments. The project manager should be in each stage of the project aware of the state of knowledge about the future of the project and the degree of its crispness, which is essential for a good project risk management.
- The method of project control should take into account not only the knowledge about the past of the project, but also each piece of new information about its future which is available (such an approach is presented in [3], however, it is based on crisp values). The information about the past should of course be used as well, but it should not be blindly aggregated without looking at dependencies between

activities. Also, the amount of work accomplished should not be expressed in monetary units. Additionally, the method should use reactive scheduling in order to determine in each moment the optimal fuzzy planned makespan of the project.

A proposal of a method fulfilling these conditions will be presented in the next section.

3 Fuzzy Project Planning and Control Method

In the method proposed we assume the following approach:

In the planning phase we propose to estimate the cost and duration of each activity as trapezoidal fuzzy numbers, respectively as $\tilde{C}_i \ (i=1,...,n)$ and $\tilde{D}_i \ (i=1,...,n)$. We treat n, $\tilde{D}_i \ (i=1,...,n)$, $\tilde{C}_i \ (i=1,...,n)$ as synonyms of n_0, $\tilde{D}_{i,0} \ (i=1,...,n_0)$, $\tilde{C}_{i,0} \ (i=1,...,n_0)$, as these parameters will be re-evaluated in the subsequent control moment (in each control moment l=1,2..,L we will new estimations $\tilde{D}_{i,l} \ (i=1,...,n_l)$, $\tilde{C}_{i,l} \ (i=1,...,n_l)$, except for the activities for which actual values will already be known. The number of activities may change in the course of the project, let it be only for the above mentioned possibility of having to repeat an activity already executed earlier. Of course, we also define the constraints according to the knowledge in the planning phase, denoted as $\Re_0$, which may also change in the course of the project. Additionally, in the planning phase we define additional dependencies of activities, which may be of the following types:

- Activities executed by the same subcontractor or the same group of people or which will use the same type of material resource
- Activities which probably will have a similar length, but this length cannot be determined exactly yet (e.g. time necessary to get an official permission from a state office or to get the ordered parts from a company with which we have not cooperated so far)
- Activities linked by a learning process, i.e. once one of the activities is executed, the following one will be executed in a shorter time and maybe at a lower cost, because the project team has gained some experience

Also other groups of activities can be defined, between which there are some dependencies in the sense that once one of them is executed and we know its actual duration and actual cost, we can better estimate the duration and cost of another activity still to be executed.

In the planning phase we estimate $P\tilde{D}$ and $P\tilde{C}$ defined in the previous section, which will be alternatively denoted as $P\tilde{D}_0$ and $P\tilde{C}_0$. Of course, these values have to be compared with the available project budget and with requirement as to

the project completion time, to evaluate the risk of not meeting the dead line and/or the budget requirements.

Then at control moments l=1,2,..,L, where L is unknown until the project is finished (we only have chose the intervals between each control moment, like one month or one week, it depends on the degree of uncertainty and risk linked to the project) , we will re-estimate the project planned cost and duration, calculating $P\tilde{D}_l$ and $P\tilde{C}_l$. We will do it using the following algorithm (for each l=1,2,..,L):

1. For the activities which have been finished (the set of their indices will be denoted as $I_{F,l}$) we will find out their actual duration and actual cost, AD_i and AC_i respectively.
2. For the activities which have been started but not finished (the set of their indices will be denoted as $I_{S,l}$) find out how much they have taken so far ($AD_{i,l}$) and how much they have cost co far ($AC_{i,l}$), and also ask the persons responsible for each of these activities to
    a. Either estimate the percentage of work that has been accomplished at those activities $ap_{i,l}$. In this case the position of $AD_{i,l}$ and $AC_{i,l}$ with respect to $ap_{i,l}\tilde{D}_{i,l-1}$, $ap_{i,l}\tilde{C}_{i,l-1}$ will be presented to the decision maker and the position of $AD_{i,l}/ap_{i,l}$ and $AC_{i,l}/ap_{i,l}$ with respect to $\tilde{D}_{i,l-1}$, $\tilde{C}_{i,l-1}$ For example, the trapezoidal fuzzy number $ap_{i,l}\tilde{D}_{i,l-1}$ represents the planned duration of the accomplishment of $ap_{i,l}$ of the whole work to be done in the i-th activity. The actual time of executing this work ($AD_{i,l}$), as well as both values divided by $ap_{i,l}$ ($\tilde{D}_{i,l-1}$ and $AD_{i,l}/ap_{i,l}$), referring to the total planned time of the activity estimated before and the same time estimated in moment l on the basis of the time that has elapsed executing this activity so far, should give the user and indication as to a new estimate $\tilde{D}_{i,l}$ It may based on $AD_{i,l}/ap_{i,l}$ (i.e. it may a fuzzy number $\left(d_{i,l}^1, d_{i,l}^2, d_{i,l}^3, d_{i,l}^4\right)$ such that $AD_{i,l}/ap_{i,l}$ belongs to the interval $\left(d_{i,l}^1, d_{i,l}^4\right)$ or even $\left(d_{i,l}^2, d_{i,l}^3\right)$), or it may based on a "combination" of $\tilde{D}_{i,l-1}$ and $AD_{i,l}/ap_{i,l}$

($\widetilde{D}_{i,l-1}$ shifted to the direction of $AD_{i,l} \big/ ap_{i,l}$ or maybe made "slimmer" in the sense that if e.g $AD_{i,l} \big/ p_{l,i} > d^3_{i,l-1}$, than values $d^1_{i,l}, d^2_{i,l}$ will be higher than $d^1_{i,l-1}, d^2_{i,l-1}$. A similar kind of reasoning might be conducted for $\widetilde{C}_{i,l-1}$ and $AC_{i,l} \big/ p_{l,i}$. We do no think any more precise indication should be given, as the user should be open not only to the information about the past of the activity, abut also to any other information concerning its future;

   b. Or give a direct estimation of $\widetilde{D}_{i,l}$, $\widetilde{C}_{i,l}$, without using the percentage of work accomplished so far, which is meaningless or difficult to measure in some cases.

3. For all the yet unstarted activities find those which were identified as related in some way to those already finished ones or those started. Show the decision maker the actual values for the related activities (AD_i and AC_i for $i \in I_{F,l}$ and $ap_{i,l}$ (if available) $AD_{i,l}$, $AC_{i,l}$ for $i \in I_{S,l}$) and ask to give the new estimations $\widetilde{D}_{i,l}$, $\widetilde{C}_{i,l}$ for those $i \in \{1,2,...,n_l\} \setminus \left(I_{F,l} \cup I_{S,l}\right)$ for which the information about the finished or started activities was important. It may happen that some $\widetilde{D}_{i,l}$ and $\widetilde{C}_{i,l}$ will become crisp or "almost crisp" in this moment, even if they concern the not yet started activities. This will be the case of activities very similar to those already finished.
4. Identify all new information about resources, subcontractors, prices, scope concerning the future of the project, update the $\widetilde{D}_{i,l}$ and $\widetilde{C}_{i,l}$ ($i \in \{1,2,...,n_l\} \setminus \left(I_{F,l} \cup I_{S,l}\right)$) concerned.
5. Identify all the unplanned activities which will have to be additionally executed, also those which will have to be repeated for some reasons, estimate $\widetilde{D}_{i,l}$ and $\widetilde{C}_{i,l}$ for them, update n_l and $\Re_l$ (taking into account the constraints or rather relations that have actually occurred (especially, the order in which activities have been accomplished so far) and any new information about them concerning the yet unstarted activities);

With the updated information we will calculate $P\widetilde{C}_l = \sum_{i=1}^{n_l} \widetilde{C}_{i,l}$ and $P\widetilde{D}_l = \left\{\left(OPD_l\left(t_j\right), PPD_l\left(t_j\right)\right)\right\}_{j=1}^{m}$, where $OPD_l\left(t_j\right)$, will be calculated on the basis crisp durations $d^2_{i,l} - t_j\left(d^2_{i,l} - d^1_{i,l}\right)$ ($i = 1,...,n_l$) and $PPD_l\left(t_j\right)$ with crisp

durations $d_{i,l}^3 + t_j\left(d_{i,l}^4 - d_{i,l}^3\right)$, using of course rescheduling, taking into account $\Re_l$. In those cases where the estimates will not be fuzzy any more but crisp, we treat the crisp number identically as a fuzzy triangular one with all the four parameters equal. The we will compare the fuzzy estimates $P\tilde{C}_l$ and $P\tilde{D}_l$ with the available budget and the deadline, taking if necessary some steps to prevent in time problems with budget or time overrun which we are able to predict now, in the control moment *l*, but, let us emphasize it, which will come into light only once the project is finished. The estimates $P\tilde{C}_l$ and $P\tilde{D}_l$ are previsions about the future whose role is – if they are for some reasons unacceptable - to give us time to react before they (or rather some unsatisfactory crisp values from their domain) actually do occur.

4 Example

Let us consider the following example, based on [4]. Let the project consist of five activities:

- Activity 1: identifying market needs
- Activity 2: designing a new product
- Activity 3: testing the prototype
- Activity 4: manufacturing the first series of the product
- Activity 5: Marketing and selling the manufactured product

The goal of the project is to successfully introduce the product into the market, thus sell at least 90% of he first series of the product, within two years from the project start. $\Re_0$ is such that all the five activities should be executed in a sequence – the *i*th activity after the (*i*-1)th is finished (*i*=2,..,5). The initial duration estimates are as follows (in months): $D_{1,0} = (1,2,2,3), D_{2,0} = (2,4,5,6), D_{3,0} = (4,6,6,7), D_{4,0} = (3,5,6,7), D_{5,0} = (1,2,2,3)$, we will not consider cost in our example. As here we have only one path in the project network, and this will be the critical path, we can calculate $P\tilde{D}$, in this special case, directly as the sum of all the fuzzy durations. We have thus $P\tilde{D}$=(11,19,21,26). We see that the deadline two years is in danger – it is possible that, the project will last up to 26 months, but let us assume that the overrun and the degree of its possibility seems for the moment acceptable.

Let us now suppose the first control moment (*l*=1) is at the end of the 3. time unit. It stated that:

- Activity 1 has not been finished, we think it will still need about 1 month (not less that 0,75 month, but not more than 1,5 month) to be finished. It has already been executed for 3 months ($AD_{1,1} = 3$)

- Although Activity 1 still goes on, we have already started Activity 2 one month ago ($AD_{2,1} = 1$) and we think we have done 10% of the work
- The other activities are not started
- We know that marketing and selling will be overtaken by a subcontractor, who will do the marketing campaign (he says it will take 2 months and the corresponding contract has been signed) and buy from us the whole first series, taking care of selling it on the market, although offering us a slightly lower price than we expected on the market. However, he requires us, and we admit it might be useful, to do a redesigning of the product after its testing, as it seems that testing is bound to indicate some problems which will have to be improved in the design of the project.

The above information will lead to the following reestimations:

- We will now have 6 activities in the project ($n_1 = 6$), Activity 6 – redesigning the product – should be executed after Activity 3 is finished and before Activity 4 is started. This defines the new set of constraints $\Re_1$, together with the fact that Activity 2 has actually overlapped Activity 1, they are both being executed in the control moment 1.
- $\tilde{D}_{1,1} = 3 + (0.75,1,1,1.5) = (3.75,4,4,4.5)$
- We have $0,1 \cdot D_{2,0} = 0.1 \cdot (2,4,5,6) = (0.2,0.4,0.5,0.6)$ - this is how much time we thought executing 10% of Activity 2 would take. Actually it has take one month, which is outside our previsions. Calculating $AD_{1,1} / 0.1$ we get 10 months, which gives us the time the activity would be finished if we worked at the same pace as before. Confronting our prevision $\tilde{D}_{2,0} = (2,4,5,6)$ with the estimate 10 and taking into account some information from the team executing the activity, we give the following new estimate of its duration: $\tilde{D}_{2,1} = (7,8,8,9)$.
- The recently added Activity 6 will be carried out by the same team as Activity 2 and between the two activities there is also the link of learning: redesigning the product will probably be easier and shorter that designing it from the beginning. Thus, basing ourselves on the information we get from the team and on $\tilde{D}_{2,1} = (7,8,8,9)$, we estimate $\tilde{D}_{6,1} = (1,2,2,3)$
- The information from the subcontractor allows us to "defuzzyfy" the estimate of Activity 5 ($\tilde{D}_{5,1} = (2,2,2,2)$), under the assumption that the subcontractor is reliable and keeps to the contracts he has signed. But if it is so, we do not have to worry about selling our product, as it will him who will buy it and it will be his problem to sell it. The information about

the lower price would influence our cost estimate, which are not taken into account in the illustrative example.

- The other estimates ($D_{3,1} = (4,6,6,7), D_{4,1} = (3,5,6,7)$) are taken form the previous stage, as no new information concerning those two activities is available.

From the control moment on Activity 1 and Activity 2 will executed simultaneously, but as the estimated remaining time of Activity 1 is $(0.75,1,1,1.5)$ and the estimated remaining time of Activity 2 is (6,7,7,8), it is clear that Activity 1 will not be on the critical path for any value from the range (0.75,1.5). Thus, we calculate PD_1 summing up the estimates for all the other activities (including Activity 6) and the time that has elapsed to the control moment, and we get PD_1=(19,25,26,30). Thus, the deadline of 24 months is in high danger. However, we have to remember that we are only after the 3. month of the execution of the project, we still have 21 months till the deadline and we have time, knowing already now that there is a high risk of not keeping the deadline, to undertake adequate steps, negotiating, seeking for additional resources etc.

5 Conclusions

We have presented a method of planning and controlling projects characterised by a high degree of uncertainty, innovativeness and due to much changes during the project execution, which cannot be foreseen in the planning phase. The method requires the users to think while estimating project parameters in terms of trapezoidal fuzzy numbers, which in fact means only giving four parameters: an optimistic one, a pessimistic one and one or two medium ones, which may also be equal to each other. It seems that such an approach would be acceptable in practice. The approach requires in each control moment not an automatic generation of numbers which do not take into account the really important information about the project history and its future, but a deeper insight into the development of the project, the influence of its environment and the interdependencies between various project elements (activities, resources etc.). This will mean a stronger effort in planning and controlling projects, but will give as a reward a more reliable information about the project and the risk connected to it, and, what is the most important thing, this information will usually be available early enough to take adequate actions.

References

[1] Cioffi, D.F.: Designing project management: A scientific notation and an improved formalism for earned value calculations. International Journal of Project Management 24(2), 136–144 (2006)

[2] Chanas, S., Zielinski, P.: Critical path analysis in the network with fuzzy activity times. Fuzzy Sets and Systems 122(2), 195–204 (2001)

[3] Dałkowski, T.B., Kuchta, D.: Klasyczna i zmodyfikowana metoda Earned Value w zarządzaniu projektami (Classic and modified Earned Value Method in project management, in Polish). Bad. Oper. Decyz. nb 3-4, pp. 35–52 (2001)
[4] Hardie, N.: The prediction and control of project duration: a recursive model. International Journal of Project Management 19(7), 401–409 (2001)
[5] Kuchta, D.: Fuzzyfication of the earned value method. WSEAS Trans. Syst. 4(12), 2222–2229 (2005)
[6] Li, J., Moselhi, O., Alksass, S.: Forecasting project status by using fuzzy logic. J. Constr. Engrg. and Mgmt. 132(11), 1193–1202 (2006)
[7] Lipke, W., Zwikael, O., Henderson, K., Anbari, F.: Prediction of project outcome: The application of statistical methods to earned value management and earned schedule performance indexes. International Journal of Project Management (in Press) (Available online April 8, 2008)
[8] Noori, S., Bagherpour, M., Zareei, A.: Applying fuzzy control chart in Earned Value analysis. World Applied Science Journal 3(4), 684–690 (2008)
[9] Vandevoorde, S., Vanhoucke, M.: A comparison of different project duration forecasting methods using earned value metrics. International Journal of Project Management 24(4), 289–302 (2006)
[10] van de Vonder, S., Ballestín, F., Demeulemeester, E., Herroelen, W.: Heuristic procedures for reactive project scheduling. Computers & Industrial Engineering 52(1), 11–28 (2007)
[11] Zielinski, P.: On computing the latest starting times and floats of activities in a network with imprecise durations. Fuzzy Sets and Systems 150(1), 53–76 (2005)

Part V
Applications

Ant Feature Selection Using Fuzzy Decision Functions

Susana M. Vieira, João M.C. Sousa, and Uzay Kaymak

Abstract. One of the most important stages in data preprocessing for data mining is feature selection. Real-world data analysis, data mining, classification and modeling problems usually involve a large number of candidate inputs or features. Less relevant or highly correlated features decrease in general the classification accuracy, and enlarge the complexity of the classifier. Feature selection is a multi-criteria optimization problem with contradictory objectives, which are difficult to properly describe by conventional cost functions. This chapter proposes the use of fuzzy optimization to improve the performance of this type of system, since it allows for an easier and more transparent description of the criteria used in the feature selection process. In our previous work, an ant colony optimization algorithm for feature selection was proposed, which minimized two objectives: number of features and classification error. In this chapter, a fuzzy objective function is proposed to cope with the difficulty of weighting the different criteria involved in the optimization algorithm. The application of fuzzy feature selection to two benchmark problems show the usefulness of the proposed approach.

Keywords: Feature selection, fuzzy optimization, ant feature selection, ant colony optimization, fuzzy modeling.

Susana M. Vieira · João M.C. Sousa
Technical University of Lisbon, Instituto Superior Técnico
Dept. of Mechanical Engineering, Center of Intelligent Systems/IDMEC
Av. Rovisco Pais, 1049-001 Lisbon, Portugal
e-mail: susana@dem.ist.utl.pt, jmsousa@ist.utl.pt

Uzay Kaymak
Econometric Institute, Erasmus School of Economics
Erasmus University Rotterdam
P.O. Box 1738, 3000 DR Rotterdam, The Netherlands
e-mail: u.kaymak@ieee.org

W.A. Lodwick & J. Kacprzyk (Eds.): Fuzzy Optimization, STUDFUZZ 254, pp. 343–364.
springerlink.com

1 Introduction

Feature selection has been an active research area in data mining, pattern recognition and statistics communities. The main idea of feature selection is to choose a subset of available features, by eliminating features with little or no predictive information, and also redundant features that are strongly correlated. Many practical pattern classification tasks (as e.g., medical diagnosis) require learning of an appropriate classification function that assigns a given input pattern (typically represented by using a vector of feature values) to one of a set of classes. The choice of features used for classification has an impact on the accuracy of the classifier and on the time required for classification. The challenge is selecting the minimum subset of features with little or no loss of classification accuracy. The feature subset selection problem consists of identifying and selecting a useful subset of features from a larger set of often mutually redundant, possibly irrelevant, features with different associated importance [10].

Like many design problems, the feature selection problem is characterized by multiple goals, where a trade-off amongst various objectives must be made. Further, some of the objectives may be known only approximately. Fuzzy set theory provides ways of representing and dealing with flexible or soft goals and constraints. This flexibility can be exploited to obtain better solutions of the optimization problem. Various fuzzy optimization methods have been proposed in the literature to deal with different aspects of soft goals and constraints. In one formulation of fuzzy optimization due to Zimmermann [33], concepts from the Bellman and Zadeh [2] model of fuzzy decision making are used for formulating the fuzzy optimization problem. Recently, a method was proposed for satisfying the problem goals, where preference for different goals can be specified by the decision maker [13]. Fuzzy optimization admits the introduction of weight factors that represent the importance of the objectives for the optimization problem. Fuzzy optimization is used to handle feature selection problems in this chapter, where a weighted fuzzy objective function is proposed.

In this chapter, we present a feature selection algorithm based on an ant colony optimization algorithm, as proposed in [27, 28]. This algorithm uses two cooperative ant colonies, which cope with two different objectives. The two objectives we consider are minimizing the number of features and minimizing the error classification. Two pheromone matrices and two different heuristics are used for each objective.

The outline of this chapter is as follows. Section 2 gives a brief overview of the feature selection problem and its inherent difficulties. Model based feature selection is briefly described, as well as fuzzy modeling, which is used to evaluate the performance of the selected subsets. Further, feature selection is formulated as an optimization problem. A fuzzy optimization approach using fuzzy criteria is proposed for the feature selection problem in Section 3. The problem is defined, and membership functions to define the fuzzy goals are proposed. Additionally, fuzzy weighted optimization is presented, where the aggregation of fuzzy criteria for feature selection is also discussed. Section 4 presents the ant colony optimization algorithm used

to solve the feature selection problem. Section 5 presents some applications of the proposed approach, and finally, Section 6 presents some concluding remarks.

2 Feature Selection

Feature selection, or variable subset selection, is the technique commonly used in selecting a subset of relevant features for building robust learning models. By removing most irrelevant and redundant features from data, feature selection helps improve the performance of learning models, by alleviating to some extent the effect of the curse of dimensionality, enhancing generalization capability, speeding up the learning process and even improving model interpretability. Feature selection also helps to better understand the data by discovering which are the important features and how they are related with each other.

From a theoretical perspective, it can be shown that optimal feature selection for supervised learning problems requires an exhaustive search of all possible subsets of features of the chosen cardinality. If large numbers of features are available, this is impractical. For practical supervised learning algorithms, the search is made for a satisfactory set of features, instead of for an optimal set [9].

Feature selection algorithms typically fall into two categories; feature ranking and subset selection. Feature ranking typically ranks the features by a metric and eliminates all features that do not achieve an adequate score. Subset selection usually entails the search for the optimal subset of possible features [11].

Subset selection evaluates a subset of features as a group, and the group of features that produces the most accurate model is selected. Subset selection algorithms can be broken into *wrappers* and *filters*. Wrappers use a search algorithm to search through the space of possible features and evaluate each subset by running a model on the subset. Wrappers can be computationally expensive and have a risk of overfitting the model. Filters are similar to wrappers in the search approach, but instead of evaluating against a model, the features are selected by evaluating a performance measure that does not require building a model.

Many popular search approaches use greedy hill climbing, which iteratively evaluates a candidate subset of features, then modifies the subset and determines whether or not if the new subset is an improvement over the old. Evaluation of the subsets requires a scoring metric that grades a subset of features. In the case of wrapper methods, the feature subsets are evaluated by the accuracy of the produced model. Exhaustive search is generally impractical, so at some defined stopping point, the subset of features with the highest score discovered up to that point is selected as the satisfactory feature subset. The stopping criterion varies by algorithm. Possible criteria include: a subset score exceeds a threshold, a program's maximum allowed run time has been surpassed, etc.

In this work, the feature selection problem is approached in the subset selection wrapper perspective. Thus, the output of a feature selection optimization problem is the reduced feature subset chosen to model the process with sufficient or necessary accuracy. Therefore, one of the main issues in model based feature selection

(MBFS) is the optimization technique applied to derive the reduced feature subset. The decision criteria of the optimization problem are the translation of the complexity (or cardinality) and accuracy of the final model. The number of selected features (cardinality) is often used as a measure of complexity in feature selection.

One of the main issues in MBFS is the type of model used to describe the process under study. Fuzzy models are rule-based systems that can be interpreted by human experts. For these reasons, fuzzy modeling is often called a "gray-box" modeling approach. This chapter uses fuzzy modeling, which is briefly presented in the following section.

2.1 Fuzzy Modeling

Rule-based expert systems are often applied to classification problems in fault detection, biology, medicine, etc. Fuzzy logic improves classification and decision support systems by allowing the use of overlapping antecedents definitions and improves the interpretability of the results by providing more insight into the classifier structure and the decision making process [18, 24].

In general, fuzzy models can provide a more transparent model and can also give a linguistic interpretation in the form of rules, which is appealing when dealing with classification systems. Fuzzy models use rules and logical connectives to establish relations between the features defined to derive the model.

The automatic determination of fuzzy classification rules from data has been approached by several different techniques: neuro–fuzzy methods, genetic–algorithm based rule selection and fuzzy clustering in combination with GA–optimization [20]. An approach that addresses simplicity and accuracy issues is used. Interpretable fuzzy rule-based classifiers are obtained from observation data following the steps described below.

In this work, we use Takagi-Sugeno (TS) fuzzy models [26], which consist of fuzzy rules where each rule describes a local input-output relation, typically in an affine form. The rules in the affine TS model are given by:

$$R_i : \textbf{If } x_1 \text{ is } A_{i1} \textbf{ and } \ldots \textbf{ and } x_n \text{ is } A_{in} \textbf{ then } y_{C_i} = a_{i1}x_1 + \ldots + a_{in}x_n + b_i\,, \quad (1)$$

where $i = 1, \ldots, K$, K denotes the number of rules in the rule base, R_i is the i^{th} rule, $\mathbf{x} = [x_1, \ldots, x_n]^T$ is the antecedent vector, n is the number of features, $A_{i1}, \ldots, A_{in}$ are fuzzy sets defined in the antecedent space, y_{C_i} is the output variable for rule i, $\mathbf{a}_i$ is a parameter vector and b_i is a scalar offset for rule i. The consequents of the affine TS model are hyperplanes in the product space of the inputs and the output. The model output, y_C, is computed by aggregating the individual rules contribution:

$$y_C = \frac{\sum_{i=1}^{K} \beta_i y_{C_i}}{\sum_{i=1}^{K} \beta_i}, \quad (2)$$

where β_i is the degree of activation of the ith rule:

$$\beta_i = \prod_{j=1}^{n} \mu_{A_{ij}}(x_j), \tag{3}$$

and $\mu_{A_{ij}}(x_j) : \mathbb{R} \rightarrow [0, 1]$ is the membership function of the fuzzy set A_{ij} in the antecedent of R_i. Each class is considered an output of the model. A model has C classes. The output of the classifier is given by the following classification decision:

$$\max_{C} y_C \tag{4}$$

Given N available input–output data pairs $(\mathbf{x}_k, \mathbf{y}_k)$, the n–dimensional pattern matrix $\mathbf{X} = [\mathbf{x}_1, \ldots, \mathbf{x}_N]^T$, and the corresponding C–dimensional class vector $\mathbf{y} = [y_1, \ldots, y_N]^T$ are constructed.

The number of rules K, the antecedent fuzzy sets A_{ij}, and the consequent parameters b_i are determined by means of fuzzy clustering in the product space of the input and output variables [24]. Hence, the data set $\mathbf{Z}$ to be clustered is composed by $\mathbf{X}$ and $\mathbf{y}$:

$$\mathbf{Z} = [\mathbf{X}, \mathbf{y}]^T . \tag{5}$$

Given the data $\mathbf{Z}$ and the number of clusters K, several fuzzy clustering algorithms can be used. This paper uses the Gustafson-Kessel (GK) [8] clustering algorithm to compute the fuzzy partition matrix $\mathbf{U}$. The matrix $\mathbf{Z}$ provides a description of the system in terms of its local characteristic behavior in regions of the data identified by the clustering algorithm, and each cluster defines a rule. The GK algorithm applies an adaptive distance measure, finding hyper-ellipsoid regions in the data that can be efficiently approximated by the hyper-planes described by the consequents in the TS model.

The fuzzy sets in the antecedent of the rules are obtained from the partition matrix $\mathbf{U}$, whose ikth element $\mu_{ik} \in [0, 1]$ is the membership degree of the data object $\mathbf{z}_k$ in cluster i. One-dimensional fuzzy sets A_{ij} are obtained from the multidimensional fuzzy sets defined point-wise in the ith row of the partition matrix by projections onto the space of the input variables x_j:

$$\mu_{A_{ij}}(x_{jk}) = \text{proj}_j^{\mathbb{N}^{n+1}}(\mu_{ik}), \tag{6}$$

where proj is the point-wise projection operator [15]. The point-wise defined fuzzy sets A_{ij} are approximated by suitable parametric functions in order to compute $\mu_{A_{ij}}(x_j)$ for any value of x_j. This is schematically represented in Fig. 1.

The consequent parameters for each rule are obtained as a weighted ordinary least-square estimate. Let $\theta_i^T = [\mathbf{a}_i, b_i]$, let $\mathbf{X}_e$ denote the matrix $[\mathbf{X}; \mathbf{1}]$ and let $\mathbf{W}_i$ denote a diagonal matrix in having the degree of activation, $\beta_i(\mathbf{x}_k)$, as its kth diagonal element. Assuming that the columns of $\mathbf{X}_e$ are linearly independent and $\beta_i(\mathbf{x}_k) > 0$ for $1 \leq k \leq N$, the weighted least-squares solution of $\mathbf{y} = \mathbf{X}_e\theta + \varepsilon$ becomes

$$\theta_i = \left[\mathbf{X}_e^T \mathbf{W}_i \mathbf{X}_e\right]^{-1} \mathbf{X}_e^T \mathbf{W}_i \mathbf{y} . \tag{7}$$

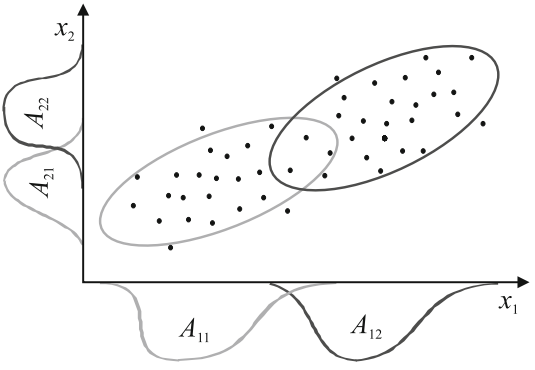


Fig. 1 Projection of multidimensional fuzzy sets onto the space of the input variables x_j

Rule bases constructed from clusters can be redundant due to the fact that the rules defined using the multidimensional antecedents are overlapping in one or more dimensions. A possible approach to solve this problem is to reduce the number of features n of the model, as addressed in this chapter.

2.2 Formulation of the Feature Selection Problem

When a classification system is designed, performance criteria must be specified. These criteria are usually defined in terms of a desired minimum error between the real classification and the model output, model complexity, number of used features, etc., representing the goals of the classification system. In MBFS, these goals must be translated as criteria into an objective function. This function normally is minimized (or maximized) over a given number of iterations.

Conventional MBFS mainly uses classification accuracy as the objective function [12]. In model based feature selection, besides classification accuracy, model complexity reduction is also desired. The model complexitiy is directly related to the number of features used for modeling.

Let the overall MBFS goals be stated as achieving good performance while reducing the features subset size, and implicitly the model complexity. The performance criterion used to evaluate the fuzzy model is the classification accuracy γ, given by:

$$\gamma = \frac{(N_n - N_e)}{N_n}. \tag{8}$$

where N_e is the number of errors in test samples and N_n is the number of used samples.

Let $\mathcal{F}$ be the n-dimensional set of features. The vector $\mathbf{z} = (z_1, z_2, \ldots, z_{N_f})$, is a subset of $\mathcal{F}$, and N_f is the features cardinality (number of used features). It is desirable that $N_f << n$. The goals can be represented by the following objective function

$$J = w_1(1-\gamma) + w_2\frac{N_f}{n}, \tag{9}$$

where γ is the percentage of correct classifications in test samples and $\{w_1, w_2\}$ represent weight factors to help setting priorities of the optimization problem. The first term of (9) represents the score of the features subset in terms of model performance. The objective function of this optimization problem aggregate both criteria: the minimization of the classification error and the minimization of the features cardinality. These are contradictory objectives, and are difficult to properly describe by this conventional cost function, once the weights are difficult to settle although the terms are normalized.

The use of fuzzy objective functions can improve the performance of this type of optimization problem, since it allows for an easier and transparent description of the different criteria used in the feature selection process. In next section, a fuzzy optimization problem is presented for MBFS to cope with the difficulty of weighting the different criteria involved in the optimization algorithm.

3 Fuzzy Optimization in Feature Selection

The objective function for feature selection can be seen as the simultaneous satisfaction of different criteria. These fuzzy criteria must be defined for different objectives inherent to the feature selection problem. When fuzzy criteria is used in the objective function, fuzzy optimization is the most obvious technique to deal with the optimization problem in MBFS.

3.1 Formulation of the Feature Selection Problem Using Fuzzy Goals

Feature selection using fuzzy goals can be defined as follows. Let G_ℓ, with $\ell = 1, \ldots, q$, be a fuzzy goal (or criterion) characterized by its membership function μ_{G_ℓ}, which is a mapping from the space of the goal G_ℓ to the interval $[0, 1]$. The goals are defined on relevant optimization criteria. Each goal is defined in the domain Φ_j, $j = 1, \ldots, q$, which can be any of the various domains used in feature selection. This optimization problem has a discrete search space with a finite and countable set Ω of subset solutions. The fuzzy criteria must be aggregated in the subset selection environment. The membership value μ_z for the subset solution z is obtained using the aggregation operator $\circledast$ to combine the fuzzy goals (criteria),

$$\mu_z = \mu_{G_1} \circledast \ldots \circledast \mu_{G_q}\,. \tag{10}$$

Various types of aggregation operations can be used as decision functions for expressing different decision strategies using the well-known properties of these operators [24]. Parametric triangular norms can generalize a large number of t-norms, and can control the degree of compensation between the different criteria. The decision criteria in (10) (e.g. small number of features and high accuracy) should be

satisfied as much as possible, which corresponds to the maximal value of the overall decision. The optimal subset z^* is found by the maximization of μ_z:

$$z^* = \arg\max_{\mathbf{z}} \mu_z. \tag{11}$$

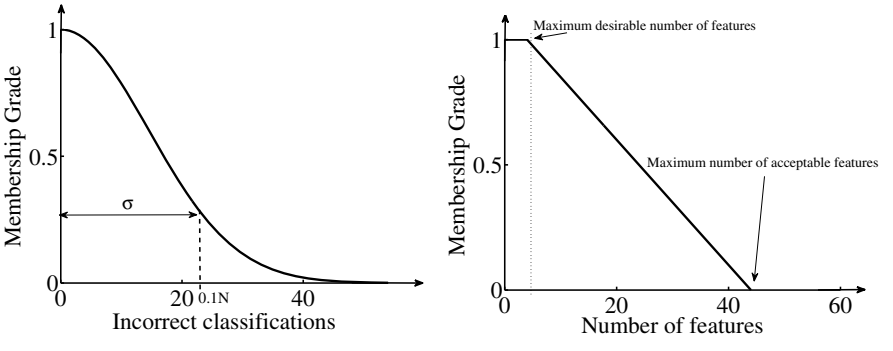
Because the membership functions for the fuzzy criteria can have an arbitrary shape, and because of the nonlinearity of the decision function, the optimization problem (11) is usually non-convex. This problem is discussed in Section 4. However, we just say at this point that given that (10) is a non-convex nonlinear optimization problem, a heuristic method such as ant colony is useful to use.

3.2 Fuzzy Criteria in Feature Selection

Fuzzy goals must be a translation of the (fuzzy) performance criteria defined for the system. The definition of these criteria in terms of model performance has shown to be quite powerful in the model based feature selection framework [9].

Additional flexibility can be introduced in MBFS, when fuzzy multicriteria is applied to determine the objective function. This flexibility provides additional control to the model builder to have control over the optimization problem. Each goal G_ℓ is described by a fuzzy set. Fuzzy criteria can be described in different ways. The most straightforward and easy way is just to adapt the classical criteria in MBFS, as defined in (9).

Figure 2 shows examples of general membership functions that can be used for the error N_e and for the features cardinality N_f. In this example, the minimization of the classification error is represented by an exponential membership function, given by (see Fig. 2a):



(a) Membership function for the number of incorrect classifications.

(b) Membership function for the number of selected features

Fig. 2 Membership functions for the feature selection goals

$$\mu_e = \exp^{-\left(\frac{N_e}{2\sigma}\right)^2} \tag{12}$$

This well-known function has the nice property of never reaching the zero value, and the membership value is still quite considerable for an error of 10%. This property is very useful, once the desirable solution should have the smaller error possible. The σ parameter can be defined as a small percentage of the number of data samples or as a percentage of error that is admissible for the problem.

The features cardinality N_f can be represented, for example, by a trapezoidal membership function, as shown in Fig. 2b. A reduced number of features is considered to be a desired outcome of the optimization algorithm. The membership function is defined so that for a low number of features the membership degree is one and linearly decreases to zero. The membership degree should be zero outside the maximum number of available features. This is because we are trying to get the number of features to be as close as possible to the minimum desired. The parameters defining the range of the trapezoidal membership function are application dependent. Sometimes it is convenient to make the upper limit of the membership function significantly lower than the maximum number of allowed features, especially if a very large number of features is being tested.

In general, the parameters of the different membership functions are application dependent. However, it is possible to derive some tuning guidelines, as described in the following. The membership functions quantify how much the system satisfies the criteria given a particular feature subset, bringing various quantities into a unified domain.

The use of the membership functions introduces additional flexibility in the goals, and it leads to increased transparency as it becomes possible to specify explicitly what kind of solution is preferred. For instance, it becomes easier to penalize more severely a subset of features that have bigger classification errors. Alternatively, if we prefer a solution with less features, a higher number of features can be penalized instead.

After the membership functions have been defined, they are combined by using a decision function, such as a parametric aggregation operator from the fuzzy sets theory, as e.g. the Yager t-norm [31]. The formulation in (10) does not use weights for the criteria. However, a general formulation that takes into account different importance of different goals must consider weights, as in the conventional objective function (9). The generalization of the fuzzy objective function (10) in order to use weighted criteria is used here, and is presented in the next section.

3.3 Weighted Fuzzy Optimization

Fuzzy objective functions using weighted criteria have been used quite extensively, especially in fuzzy decision making, where the weights are used to represent the relative importance that the decision maker attaches to different decision criteria. An averaging operator is normally used for the weighted aggregation, such as generalized means [6], fuzzy integrals [7] or the ordered weighted average (OWA) operators [32]. Consequently, weighted aggregation of fuzzy sets has been studied with averaging type of operators. The generalized means extend naturally to weighted

equivalents. The weighted generalized mean operator has been used in many fields, and it has been studied in the context of fuzzy set aggregation in [6, 14]. The OWA operators and the fuzzy integrals are inherently weighted operators, which do not need a separate extension to the weighted case. Applications of these operators have also been reported in the literature, see e.g. [7, 19].

In the fuzzy feature selection approach, the goal is to satisfy simultaneously the model accuracy and the feature subset reduction. The averaging operators are suitable for modeling compensatory aggregation. In compensatory aggregation, one criterion can compensate for another one. However, they are not suitable for modeling simultaneous satisfaction of aggregated contradictory criteria, where the aggregated value should never be larger then the least satisfied criterion. In this case, t-norms must be used to model the conjunctive aggregation [25]. Therefore, weighted aggregation using t-norms must be considered. Note that the most common axiomatic definition of t-norms does not allow for weighted aggregation. Hence, the commutativity and the associativity properties must be dropped, since weighted operators are by definition not commutative.

Weighted aggregation of fuzzy sets by using t-norms has been considered first by Yager in [29], where the membership functions are modified by associated weight factors before the fuzzy aggregation. The application of a generalized form of this idea, introduced in [30], is given by

$$\mu_z(\mathbf{w}) = t[I(\mu_{G_1}, w_1), \ldots, I(\mu_{G_i}, w_q)], \tag{13}$$

where $\mathbf{w}$ is a vector of weight factors $w_l \in [0, 1]$, t is a t-norm and I is a function that transforms the membership functions. Note that the fuzzy objective function for feature selection in this chapter has two goals ($q = 2$), where G_1 translates the classification accuracy and G_2 the features cardinality reduction. The most common fuzzy aggregation operator uses the power-raising method for the transformation and the minimum operator for the t-norm.

$$\mu_z(\mathbf{w}) = \bigwedge_{l=1}^{q} [\mu_{G_l}]^{w_l} . \tag{14}$$

This aggregation function has been used in many publications regarding the application of fuzzy weighted aggregation, especially in multicriteria decision making. Weighted aggregation of fuzzy sets has been investigated in more detail in a generalized framework [13, 14], where weighted counterparts of fuzzy t-norms have also been proposed based on a sensitivity analysis of weighted fuzzy aggregation. This analysis provides a general mechanism for introducing weight factors into Archimedean t-norms and t-conorms by considering several requirements that can be imposed on a weighted aggregation operator. In this work, weighted counterparts of Archimedean t-norms are used, namely, the weighted extension of the product t-norm

$$\mu_z(\mathbf{w}) = \prod_{l=1}^{q} [\mu_{G_l}]^{w_l}, \tag{15}$$

the extension of the Hamacher t-norm

$$\mu_z(\mathbf{w}) = \begin{cases} \dfrac{1}{1+\sum_{l=1}^{q} w_l \dfrac{1-\mu_{G_l}}{\mu_{G_l}}} & \text{if } \forall l, \mu_{G_l} > 0 \\ 0 & \text{if } \exists l, \mu_{G_l} = 0 \end{cases} \tag{16}$$

and the extension of the Yager t-norm

$$\mu_z(\mathbf{w}) = \max\left(0, 1 - \sqrt[s]{\sum_{l=1}^{q} w_l (1-\mu_{G_l})^s}\right), \qquad s > 0\,. \tag{17}$$

Note that the extension of the product t-norm in (15), according to the sensitivity based analysis, is the same as the application of (13) using the product operator as t, with power raising method as I. However, the extensions (16) and (17) cannot be obtained from (13). Applications of weighted fuzzy optimization can be found in [13, 17].

After combining the objectives, the resulting optimization is non-convex. Furthermore, gradient descent methods may not be suitable for the maximization due to possible and likely discontinuity in the first derivative of the final aggregated function. Derivative-free search and optimization algorithms such as simulated annealing, evolutionary algorithms or other bio-inspired algorithms, such as ant colony optimization, can be used to solve this type of optimization problems. In this work, an ant colony optimization approach is used to solve the MBFS optimization problem, as proposed in [27, 28]. This approach is described in the next section.

4 Ant Feature Selection

Ant algorithms were first proposed by Dorigo [3] as a multi-agent approach to difficult combinatorial optimization problems, such as traveling salesman problem, quadratic assignment problem or supply chain management [21, 22, 23]. The ant colony optimization (ACO) methodology is an optimization method suited to find minimum cost paths in optimization problems described by graphs [4].

This chapter presents an implementation of ACO applied to feature selection, where the best number of features is determined automatically [27, 28]. In this approach, two objectives are considered: minimizing the number of features and minimizing the classification error. Two cooperative ant colonies are considered, one for each objective. The first colony determines the number (cardinality) of features and the second selects the features based on the cardinality given by the first colony. Thus, two pheromone matrices and two different heuristics are used. The heuristic value is computed using the Fisher discriminant criterion for feature selection [5], which ranks the features by giving them a relative importance.

The determination of the *features cardinality* N_f is addressed in the first colony sharing the same minimization cost function with the second colony, which in this case aggregates both the maximization of the classification accuracy and the

Table 1 Variables definition

Variable	Description
General	
n	Number of features
N	Number of samples
N_n	Number of samples used for validation
I	Number of iterations
K	Number of rules/clusters of the fuzzy model
C	Number of existing classes in database
g	Number of ants
$\mathbf{x}$	Set with all the features
$\mathbf{w}$	Subset of features selected to build classifiers
J^k	Cost of the solution for each ant k
J^q	Cost of the winner ant q
Ant colony for cardinality of features	
N_f	Features cardinality (number of selected features)
$N_f(k)$	Features cardinality of ant k
I_n	Number of iterations with same feature cardinality
α_n	Pheromone weight of features cardinality
β_n	Heuristic weight of features cardinality
τ_n	Pheromone trails for features cardinality
η_n	Heuristic of features cardinality
ρ_n	Evaporation of features cardinality
Γ_n^k	Feasible neighborhood of ant k (features cardinality availability)
$\mathcal{Q}_i$	Amount of pheromone laid in the features cardinality of the best solution
Ant colony for selecting subset of features	
$L_f^k(t)$	Feature subset for ant k at tour t
α_f	Pheromone weight of features
β_f	Heuristic weight of features
τ_f	Pheromone trails for feature selection
η_f	Heuristic of features
ρ_f	Evaporation of features
Γ_f^k	Feasible neighborhood of ant k (features availability)
$\mathcal{Q}_j$	Amount of pheromone laid in the features of the best solution

minimization of the features cardinality. Hence, the first colony determines the size of the subsets of the ants in the second colony, and the second colony selects the features that will be part of the subsets.

The algorithm used in this study deals with the feature selection problem as a multi–criteria problem with a single objective function. Therefore, a pheromone matrix is computed for each criterion, and different heuristics are used. Table 1

describes the variables used in the algorithm. To evaluate the classification error, a fuzzy classifier is built for each solution using the procedure described in Section 2.

4.1 Probabilistic Rule

Consider a problem with N_f nodes and two colonies of g ants. First, g ants of the first colony randomly select the number of nodes N_f to be used by the g ants of the second colony. Following the original ACO [3], the probability that an ant k chooses the features cardinality $N_f(k)$ is given by

$$p_i^k(t) = \frac{[\tau_{n_i}]^{\alpha_n} \cdot [\eta_{n_i}]^{\beta_n}}{\sum_{l \in \Gamma_n^k} [\tau_{n_l}]^{\alpha_n} \cdot [\eta_{n_l}]^{\beta_n}} \tag{18}$$

where τ_{n_i} is the pheromone concentration matrix and η_{n_i} is the heuristic function matrix, for path (i). The values of the pheromone matrix are limited to $[\tau_{n_{\min}}, \tau_{n_{\max}}]$, with $\tau_{n_{\min}} = 0$ and $\tau_{n_{\max}} = 1$. Γ_n^k is the feasible neighborhood of ant k (available number of features to be selected), which acts as the memory of the ants, and contains all the trails that the ants have not passed and can be chosen, here the trails represent the features. The parameters α_n and β_n measure the relative importance of trail pheromone and heuristic knowledge, respectively.

After all the g ants from the first colony have chosen the features cardinality $N_f(k)$, each ant k from the second colony selects $N_f(k)$ features (nodes). The probability that an ant k chooses feature j as the next feature to visit is given by

$$p_j^k(t) = \frac{[\tau_{f_j}(t)]^{\alpha_f} \cdot [\eta_{f_j}]^{\beta_f}}{\sum_{l \in \Gamma_f^k} [\tau_{f_l}(t)]^{\alpha_f} \cdot [\eta_{f_l}]^{\beta_f}} \tag{19}$$

where τ_{f_j} is the pheromone concentration matrix and η_{f_j} is the heuristic function matrix for the path (j). Again, the pheromone matrix values are limited to $[\tau_{f_{\min}}, \tau_{f_{\max}}]$, with $\tau_{f_{\min}} = 0$ and $\tau_{f_{\max}} = 1$. Γ_f is the feasible neighborhood of ant k (available features), which contains all the features that the ants have not selected and can be chosen. Again, the parameters α_f and β_f measure the relative importance of trail pheromone and heuristic knowledge, respectively.

4.2 Updating Rule

After a complete tour, when all the g ants have visited all the $N_f(k)$ nodes, both pheromone concentration in the trails are updated by

$$\tau_{n_i}(t+1) = \tau_{n_i}(t) \times (1 - \rho_n) + \Delta\tau_{n_i}(t) \tag{20}$$

$$\tau_{f_j}(t+1) = \tau_{f_j}(t) \times (1 - \rho_f) + \Delta\tau_{f_j}(t) \tag{21}$$

where $\rho_n \in [0,1]$ is the pheromone evaporation of the features cardinality , $\rho_f \in [0,1]$ is the pheromone evaporation of the features and $\Delta\tau_{n_i}$ and $\Delta\tau_{f_j}$ are the pheromone deposited on the trails (i) and (j), respectively, by the ant q that found the best solution J^q for this tour:

$$\Delta\tau_{n_i}^q = \begin{cases} \mathcal{Q}_i & \text{if node } (i) \text{ is used by the ant } q \\ 0 & \text{otherwise} \end{cases} \tag{22}$$

$$\Delta\tau_{f_j}^q = \begin{cases} \mathcal{Q}_j & \text{if node } (j) \text{ is used by the ant } q \\ 0 & \text{otherwise} \end{cases} \tag{23}$$

The number of nodes $N_f(k)$ that each ant k has to visit on each tour t is only updated every I_n tours (iterations), in order to allow the search for the best features for each N_f. The algorithm runs I times. Both colonies share the same cost functions. Classical and fuzzy cost functions, given respectively by (9) and (10) are both tested in this chapter.

4.3 Heuristics

The heuristic value used for the second ant colony is computed as

$$\eta_{f_j} = 1/Ne_j \tag{24}$$

for $j = 1, \ldots, n$. For the features cardinality (first colony), the heuristic value is computed using the Fisher discriminant criterion for feature selection [5]. Considering a classification problem with two possible classes, class 1 and class 2, the Fisher discriminant criterion is described as

$$F(i) = \frac{|\mu_1(i) - \mu_2(i)|^2}{\sigma_1^2 + \sigma_2^2} \tag{25}$$

where $\mu_1(i)$ and $\mu_2(i)$ are the mean values of feature i for the samples in class 1 and class 2, respectively, and σ_1^2 and σ_2^2 are the variances of feature i for the samples in classes 1 and 2. The score aims to maximize the between-class difference and minimize the within-class spread. Other currently proposed rank-based criteria generally come from similar considerations and show similar performance [5]. Since our goal is to work with several classification problems, which can contain two or more possible classes, a one versus-all strategy is used to rank the features. Thus, for a C-class prediction problem, a particular class is compared with the other $C-1$ classes that are considered together. The features are weighted according to the total score summed over all C comparisons:

$$\sum_{j=1}^{C} F_j(i), \tag{26}$$

Algorithm 1. Ant Feature Selection

*/*Initialization*/*
set the parameters ρ_f, ρ_n, α_f, α_n, β_f, β_n, I, I_n, g.
for $t = 1$ *to* I ***do***
  for $k = 1$ *to* g ***do***
    Choose the subset size $N_f(k)$ *of each ant* k *using (18)*
  end for
  for $l = 1$ *to* I_n ***do***
    for $k = 1$ *to* g ***do***
      Build feature set $L_f^k(t)$ *by choosing* $N_f(k)$ *features using (19)*
      Compute the fuzzy model using the $L_f^k(t)$ *path selected by ant* k
      Compute the cost function $J^k(t)$
      Update J^q
    end for
    Update pheromone trails $\tau_{n_i}(t+1)$ *and* $\tau_{f_j}(t+1)$, *as defined in (20) and (21).*
  end for
end for

where $F_j(i)$ denotes the Fisher discriminant score for the i^{th} feature at the j^{th} comparison. Algorithm 1 presents the description of the ant feature selection algorithm.

5 Application Examples

5.1 Data Sets

The effectiveness of the proposed approach is tested using data sets taken from some well known benchmarks in the UCI repository [1]. Two real data sets, Wine and Wisconsin Breast Cancer were used to test the presented approach. The characteristics of the data are presented in Table 2.

Table 2 Description of the used data sets

Data sets used	# features	# classes	# samples
Wine	13	3	178
Breast Cancer	9	2	699

Wine. The wine data set is widely used in the literature. The classification data is available online in the repository of the University of California [1], and contains the chemical analysis of 178 wines grown in the same region in Italy, derived from three different cultivars. Thirteen continuous attributes are available for classification: alcohol, malic acid, ash, alkalinity of ash, magnesium, total phenols, flavanoids, nonflavanoids phenols, proanthocyanism, color intensity, hue, the ratio OD280/OD315 of dilluted wines and proline.

Breast Cancer. The Wisconsin breast cancer data is also widely used to test the effectiveness of classification algorithms. The aim of the classification is to distinguish between benign and malignant cancers based on the available nine measurements (attributes): clump thickness, uniformity of cell size, uniformity of cell shape, Marginal Adhesion, Single Epithelial Cell Size, Bare Nuclei, Bland Chromatin, Normal Nucleoli and Mitoses. The attributes have integer value in the range [1,10]. The original database contains 699 instances, however 16 samples are omitted because they are incomplete. The class distribution is 65.5% benign and 34.5% malignant. The breast cancer data set is also available in the repository of the University of California [1] and it was obtained from the University of Wisconsin Hospitals, Madison from Dr. William H. Wolberg.

5.2 *Performance Evaluation*

This chapter uses two types of performance evaluation. First, the data sets are divided in training and test instances (50% for training and 50% for test). As is common, the test data set contains data points different from the ones used to train the model. The models are constructed using the procedure described in Section 2.1. As these results are excellent for the considered data sets, more demanding validation tests were necessary. Thus, we used the well-known cross validation method, which is briefly described next.

Cross-validation. In cross validation, the data set with N_n samples is divided into N mutually exclusive sets of approximately equal size, with each subset consisting of approximately the same proportions of labels as the original data set, known as stratified cross validation [16]. The classifier is trained N times, with a different subset left out as the test set and the other samples used to train the classifier at each time. During the training phase, the classifier is trained on $N-1$ out of N folds in which classification accuracy is used, as defined in (8). The prediction performance of the classifier is estimated by considering the average classification accuracy of the 10 cross-validation experiments, described as

$$E_{CV} = \left(\frac{1}{N_n} \sum_{i=1}^{N} C_i \right) \times 100\% \tag{27}$$

where C_i is the number of correctly classified samples:

$$C_i = N_n - N_e.$$

The classification error rates of the final subset solutions are obtained by performing N-fold cross validation, with $N = 10$ in our case (CV10). The experimental results are presented as the best, the worst and the mean value of the classification accuracy γ, as defined in (8).

5.3 *Results*

5.3.1 Train-Test Results

First the data sets were divided into training and test data, and 10 trials are simulated. The fuzzy goals were aggregated using the Yager t-norm. The maximum, mean and minimum classification accuracies of these 10 trials are shown in Table 3. The results are compared with the ant feature selection (AFS) approach in [28]. Table 3 shows clearly that using a fuzzy objective function the algorithm always converges to the optimal solution. Further, the algorithm converges always to the same number of features, which was not the case with AFS. Clearly, more demanding tests were necessary to test the fuzzy approach. Therefore, 10-fold cross validation was applied to the data.

Table 3 Classification rates for train/test data sets

Data set	Methods	Reduced Subset	Classification accuracy (%)		
			Max.	Mean	Min.
Wine	AFS	4-8	100.0	99.8	98.9
	Fuzzy AFS	4	100.0	100.0	100.0
Breast Cancer	AFS	2-5	100.0	96.4	91.3
	Fuzzy AFS	3	100.0	100.0	100.0

5.3.2 Cross Validation Results

Ten–fold cross validation was applied to both data sets. Different aggregation operators were used to test the fuzzy optimization, namely: product, Yager and Hamacher t-norms. The results for the data sets are presented next.

Wine Results. The results obtained for the wine data set are presented in Table 4. The t-norm that constantly achieves the smaller number of features is the Yager t-norm. The best results in terms of accuracy are obtained with the Hamacher t-norm, using however a much larger number of features. The product t-norm does not converge to the same number of features, achieving results similar to the ones without the fuzzy approach, i.e. using the AFS algorithm.

Table 4 Cross validation results for the Wine data set

t-norm	Number of features	Classification accuracy (%)		
		Max.	Mean	Min.
Product	4-8	100	91.1	72.2
Yager	4	100	92.0	75.0
Hamacher	11	100	93.9	88.9

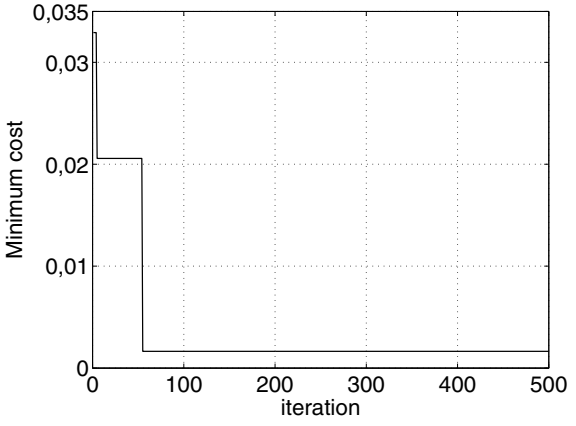


Fig. 3 Cost values for each iteration in the Wine data set

Figure 3 presents an example of the values obtained for the fuzzy objective function in (17), when the Yager t-norm is used. The convergence of the fuzzy optimization algorithm using again the Yager t-norm is depicted in Fig. 4, where the convergence to a constant number of features is clearly shown. Note that a 10-fold

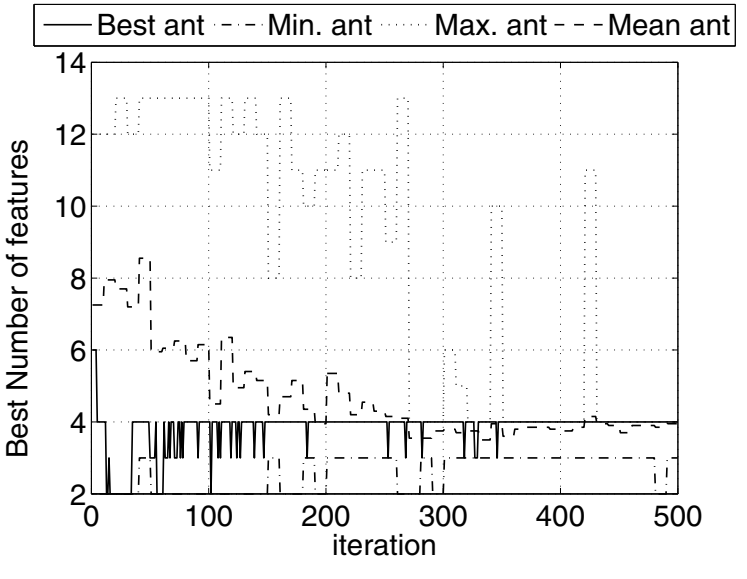


Fig. 4 Best number of features for each iteration in the Wine data set

Table 5 Cross validation results for the Wisconsin Breast Cancer data set

t-norm	Number of features	Classification accuracy (%)		
		Max.	Mean	Min.
Product	3-5	100	96.1	91.4
Yager	3-4	100	96.0	94.3
Hamacher	9	98.6	94.9	90.0

cross validation in the wine data set is clearly a very demanding test for the classification model, as the data set has very few instances.

Breast Cancer results. Table 5 presents the results obtained for this data set using several aggregation t-norms. Depending on the t-norm used, is not always possible to minimize the number of features. Again, the t-norm that constantly achieves the smaller number of features is the Yager t-norm. Further, the product t-norm does not converge to the same number of features, as happened with the Wine data set. In terms of accuracy, the best results are now obtained with the Yager t-norm (although the product has similar results), and the worst results are obtained with the Hamacher t-norm.

Figure 5 presents the convergence of the fuzzy optimization algorithm using the Yager t-norm. This t-norm converges to a small number of features (four in this case).

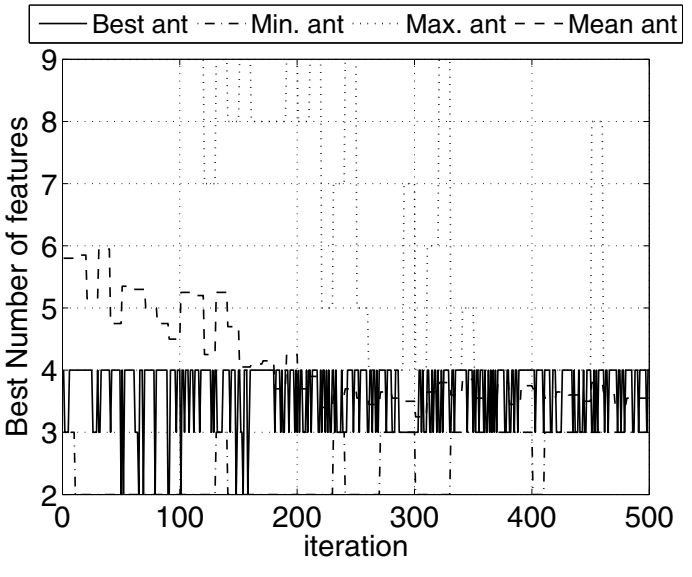


Fig. 5 Best number of features for each iteration in the breast cancer data set

5.4 Discussion

By observing Tables 4 and 5 it becomes clear that the Yager t-norm is the one obtaining the smaller number of features. The accuracy is also very good, but not always the best. In the Wine data set the best accuracy was obtained with the Hamacher t-norm.

Ant feature selection using the conventional objective function in (9) has some difficulties to satisfy both optimization criteria, even when weight factors are used. In general, the use of fuzzy optimization results in a better convergence, especially when the Yager t-norm is utilized. In summary, it is possible to conclude from the results that the performance of the optimization algorithm has improved using a fuzzy objective function.

6 Conclusions

A fuzzy objective function for ant feature selection is proposed in this chapter. The problem is divided into two objectives: minimizing the features cardinality and selecting the most relevant features. The feature selection algorithm uses fuzzy classifiers to evaluate the selected subsets of features. The proposed algorithm was applied to two well known classification databases that are considered benchmarks. Different fuzzy aggregation t-norms were tested. The results show that the proposed approach leaded to better results than ant feature selection using a classical objective function.

A systematic procedure to choose the weighting factors was not used in this study. Thus, further work is necessary to evaluate the weighting effect on the aggregation methods. This study should also test changing the weights during the evolution of the optimization algorithm. Finally, data sets with a larger number of features must be tested to confirm the obtained results.

Acknowledgements. This work is supported by: the Portuguese Government and FEDER under the programs: Programa de financiamento Plurianual das Unidades de I&D da FCT (POCTI-SFA-10-46-IDMEC) and by the FCT grant SFRH/25381/2005, Fundação para a Ciência e a Tecnologia, Ministério do Ensino Superior, da Ciência e da Tecnologia, Portugal; and by the European Science Foundation, COST Action IC0702, "Combining Soft Computing Techniques and Statistical Methods to Improve Data Analysis Solutions".

References

[1] Asuncion, A., Newman, D.J.: UCI machine learning repository (2007)
[2] Bellman, R.E., Zadeh, L.A.: Decision-making in a fuzzy environment. Management Science 17(4), 141–164 (1970)
[3] Dorigo, M.: Optimization, Learning and Natural Algorithms (in Italian). PhD thesis (1992)

[4] Dorigo, M., Birattari, M., Stützle, T.: Ant colony optimization. IEEE Computational Intelligence Magazine 1(4), 28–39 (2006)
[5] Duda, R.O., Hart, P.E., Stork, D.G.: Pattern Classification, vol. 2. Wiley–Interscience Publication, Chichester (2001)
[6] Dyckhoff, H., Pedrycz, W.: Generalized means as model of compensative connectives. Fuzzy Sets and Systems 14, 143–154 (1984)
[7] Grabisch, M., Nguyen, H.T., Walker, E.A.: Fundamentals of uncertainty calculi with applications to fuzzy inference. In: Mathematical and Statistical Methods, vol. 30. Kluwer Academic Publishers, Dordrecht (1995)
[8] Gustafson, D.E., Kessel, W.C.: Fuzzy clustering with a fuzzy covariance matrix. In: Proceedings of the 18^{th} IEEE Conference on Decision and Control, San Diego, CA, USA, pp. 761–766 (1979)
[9] Guyon, I., Elisseeff, A.: An introduction to variable and feature selection. Journal of Machine Learning Research 3, 1157–1182 (2003)
[10] Motoda, H., Liu, H.: Feature Selection for Knowledge Discovery and Data Mining. Kluwer Academic Publishers, Dordrecht (1998)
[11] Jensen, R., Shen, Q.: Finding rough set reducts with ant colony optimization. In: Proceedings of the 2003 UK Workshop on Computational Intelligence, pp. 15–22 (2003)
[12] Jensen, R., Shen, Q.: Fuzzy-rough data reduction with ant colony optimization. Fuzzy Sets and Systems 149, 5–20 (2005)
[13] Kaymak, U., Sousa, J.M.: Weighted constraint aggregation in fuzzy optimization. Constraints 8(1), 61–78 (2003)
[14] Kaymak, U., van Nauta Lemke, H.R.: A sensitivity analysis approach to introducing weight factors into decision functions in fuzzy multicriteria decision making. Fuzzy Sets and Systems 97(2), 169–182 (1998)
[15] Klir, G.J., Yuan, B.: Fuzzy Sets and Fuzzy Logic: theory and applications. Prentice-Hall, Upper Saddle River (1995)
[16] Kohavi, R.: A study of cross-validation and bootstrap for accuracy estimation and model selection. In: Proc. International Joint Conf. Artificial Intelligence (1995)
[17] Mendonça, L.F., Sousa, J.M.C., Kaymak, U., Sá da Costa, J.M.G.: Weighting goals and comstraints in fuzzy predictive control. Journal of Intelligent and Fuzzy Systems 17(5), 517–532 (2006)
[18] Roubos, J.A., Setnes, M., Abonyi, J.: Learning fuzzy classification rules from labeled data. International Journal of Information Sciences 150(1), 77–93 (2003)
[19] Salido, J.M.F., Murakami, S.: Extending Yager's orness concept for the OWA aggregators to other mean operators. Fuzzy Sets and Systems 139(3), 515–542 (2003)
[20] Setnes, M., Roubos, J.A.: GA-fuzzy modeling and classification: complexity and performance. IEEE Transactions on Fuzzy Systems 8(5), 509–522 (2000)
[21] Silva, C.A., Sousa, J.M.C., Runkler, T.A.: Rescheduling and optimization of logistic processes using GA and ACO. Engineering Applications of Artificial Intelligence 21(3), 343–352 (2007)
[22] Silva, C.A., Sousa, J.M.C., Runkler, T.A., Sá da Costa, J.M.G.: Distributed optimization of a logistic system and its suppliers using ant colony optimization. International Journal of Systems Science 37(8), 503–512 (2006)
[23] Silva, C.A., Sousa, J.M.C., Runkler, T.A., Sá da Costa, J.M.G.: Distributed supply chain management using ant colony optimization. To appear in European Journal of Operational Research (2009), doi:10.1016/j.ejor.2008.11.021
[24] Sousa, J.M.C., Kaymak, U.: Fuzzy Decision Making in Modeling and Control. World Scientific/Imperial College, Singapore/UK (2002)

[25] Sousa, J.M.: Optimization issues in predictive control with fuzzy objective functions. International Journal of Intelligent Systems 15(9), 879–899 (2000)
[26] Takagi, T., Sugeno, M.: Fuzzy identification of systems and its applications to modelling and control. IEEE Transactions on Systems, Man and Cybernetics 15(1), 116–132 (1985)
[27] Vieira, S.M., Sousa, J.M.C., Runkler, T.A.: Fuzzy classification in ant feature selection. In: Proc. of 2008 IEEE World Congress on Computational Intelligence, WCCI 2008, pp. 1763–1769, Hong Kong, China (June 2008)
[28] Vieira, S.M., Sousa, J.M.C., Runkler, T.A.: Two cooperative ant colonies for feature selection using fuzzy models. Submitted to Expert Systems with Applications (2009)
[29] Yager, R.R.: Fuzzy decision making including unequal objectives. Fuzzy Sets Systems 1, 87–95 (1978)
[30] Yager, R.R.: General multiple-objective decision functions and linguistically quantified statements. International Journal of Man-Machine Studies 21(5), 389–400 (1984)
[31] Yager, R.R.: On a general class of fuzzy connectives. Fuzzy Sets and Systems 4, 235–242 (1980)
[32] Yager, R.R.: On ordered weighted averaging aggregation operators in multicriteria decision making. IEEE Transaction Systems, Man and Cybernetics 18(1), 183–190 (1988)
[33] Zimmermann, H.J.: Description and optimization of fuzzy systems. International Journal of General Systems 2, 209–215 (1976)

Application of Fuzzy Theory to the Investment Decision Process

Hiroshi Tsuda and Seiji Saito

Abstract. In the present paper, we propose a new approach to portfolio optimization that allows portfolio managers to construct portfolios that reflect their views about risk assets by applying fuzzy theory. The proposed approach to the investment decision process is based on the mean-variance approach proposed by Markowitz (1952,1959) and uses the concept of asset market equilibrium proposed by Sharpe (1964). For portfolio managers, it is very meaningful to use the equilibrium expected excess returns associated with the capital market as a reference. The proposed approach enables a new method for incorporating the views of portfolio managers to aid in the investment decision process. Moreover, in order to estimate the distribution of an unknown true membership function of the views of portfolio managers concerning risk assets, we propose a fuzzy information criterion to evaluate the fitness of the distribution between an unknown true membership function and a hypothetical membership function. In particular, the proposed approach enables a group of portfolio managers of pension funds to obtain an important solution that realizes optimal expected excess returns of risk assets by specifying the vague views of portfolio managers as a fuzzy number.

1 Introduction

In the present paper, we propose a new approach to portfolio optimization that allows portfolio managers to construct portfolios that reflect their views concerning risk assets by applying fuzzy theory. The portfolio theory of the mean-variance approach proposed by Markowitz (1952,1959) has provided a valuable framework for the investment decision process. In this approach, the return rate of an asset is

Hiroshi Tsuda and Seiji Saito
Department of Mathematical Sciences
Doshisha University
1-3 Tsutani, Tatara, Kyotanabe, Kyoto 610-0321, Japan
e-mail: htsuda@mail.doshisha.ac.jp, ssaito@mail.doshisha.ac.jp

W.A. Lodwick & J. Kacprzyk (Eds.): Fuzzy Optimization, STUDFUZZ 254, pp. 365–387.
springerlink.com

treated as a random variable. Stochastic programming is applied to obtain the optimal solution of the portfolio. Instead of stochastic programming, fuzzy mathematical programming approaches, such as possibilistic programming, have been proposed to solve portfolio selection problems. These fuzzy mathematical programming approaches are referred to Watada (1997), Inuiguchi and Tanino (2000), and Inuiguchi and Ramik (2000). Hasuike, Katagiri, and Ishii (2009) proposed methods by which to solve portfolio selection problems, including probabilistic future returns with random fuzzy variables. However, these approaches do not consider the concept of the asset market equilibrium. Moreover, when portfolio managers directly apply the Markowitz mean-variance approach to asset management in practice, they must specify expected returns of risk assets and their standard deviations, which are defined as risk. In practical portfolio management, it is impossible to know the expected returns of random variables, fuzzy variables, or random fuzzy variables. This critical problem occurs in portfolio selection when using the fuzzy mathematical programming mentioned in Watada (1997), Inuiguchi and Tanino (2000), and Inuiguchi and Ramik (2000).

As a practical matter, portfolio managers have vague views about relative returns for different assets. There have been no reports of an approach for incorporating these vague views about the expected returns of risk assets into portfolio approaches.

The remainder of the present paper is organized as follows. In Section 2, we introduce a portfolio selection problem. In Section 3, we describe the equilibrium expected excess returns of risk assets. In Section 4, we propose a new approach of portfolio optimization, which allows portfolio managers to construct portfolios that reflect their views concerning risk assets by applying fuzzy theory. The proposed approach makes possible a new method for incorporating the views of portfolio managers that facilitates the investment decision process. In Section 5, the proposed approach is described in detail through a number of case studies. In addition, we propose a fuzzy information criterion (FIC) by which to evaluate the fitness of the distribution between an unknown true membership function generating data and a hypothetical membership function. Finally, our conclusions are presented in Section 6.

2 Portfolio Theory

Markowitz (1952,1959) proposed the portfolio theory of the mean-variance approach, in which a portfolio is considered in its entirety, rather than as a collection of individual securities. It is not sufficient to consider the characteristics of individual assets when putting together a portfolio. The co-movement between assets is a very important consideration when attempting to construct an optimal portfolio. A portfolio that generates higher expected returns at the same level of risk or the same level of expected returns at the lower level of risk can be constructed by considering the co-movements between assets.

We assume that there are M risk assets in a certain capital market. Let us denote the return rate of the j th asset by R_j and its portfolio weight by w_j. Then, the expected return rate corresponds to the return, while the variance of the return corresponds to the risk. The portfolio selection problem is formulated as the following mathematical programming problem:

$$\text{Maximize } E\left(\sum_{j=1}^{M} w_j R_j\right) = \sum_{j=1}^{M} w_j E\left(R_j\right) \tag{1}$$

$$\text{Minimize } \sum_{j=1}^{M}\sum_{k=1}^{M} w_j w_k \sigma_{jk}, \tag{2}$$

where σ_{jk} is the covariance between the j-th and k-th assets,

$$\text{subject to } \sum_{j=1}^{M} w_j = 1, \quad w_j \geq 0, \quad j = 1, 2, \cdots, M. \tag{3}$$

The curve of the expected returns and risk (standard deviations) of optimal solutions in Figure 1 is referred to as the efficient frontier. Portfolios on the efficient frontier generate higher expected returns at the same level of risk or the same level of expected returns at the lower level of risk .

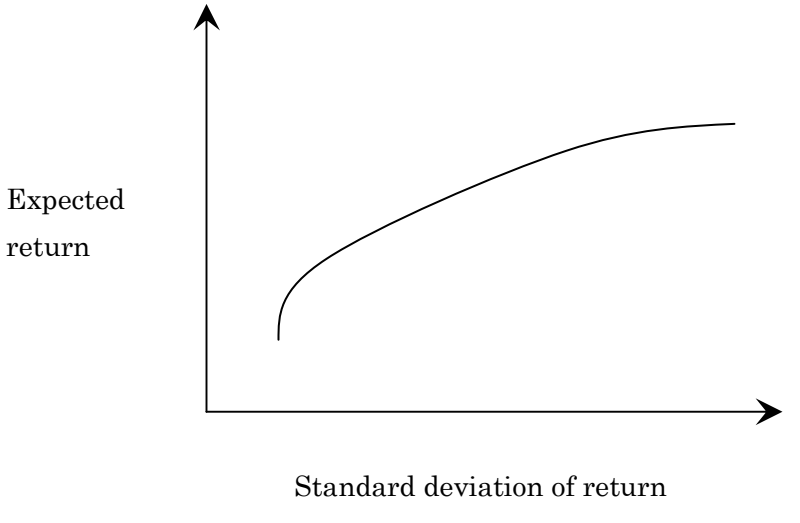


Fig. 1 Efficient frontier

Although the portfolio theory of the mean-variance approach proposed by Markowitz might seem appealing and reasonable from a theoretical standpoint, portfolio managers confront several problems when applying this theory in practice. One such problem is that the portfolio managers must specify the expected returns of risk assets and their standard deviations. Portfolio managers often use the sample mean of historical return rate data as the expected return. However, the sample mean is not a future expected return.

Basically, the Markowitz mean-variance approach overweights assets with high expected returns and underweights assets with low expected returns. The Markowitz approach almost always provides results that include large negative weights for several assets. Since pension portfolio managers are often not permitted to take short positions, a shorting constraint is added to portfolio construction.When optimizing a portfolio with short constraints, a solution is apt to take zero weights in many of assets and large weights in only a few of the assets.

Although there exist several problems when using the Markowitz approach in practice, the concept of optimizing the tradeoff relationship between the expected return and the risk is very important. Therefore, we herein propose a new approach that avoids the practical problems associated with the use of the Markowitz approach.

3 Equilibrium Expected Excess Returns of Risk Assets

As described above, when portfolio managers attempt to apply the mean-variance approach to asset management in practice, they are confronted with the problem of specifying the expected returns of risk assets. One of the solutions is to use the equilibrium expected returns of the Capital Asset Pricing Model (CAPM) proposed by Sharpe (1964) as expected returns of reference. This idea to use the expected returns associated with asset market equilibrium was proposed by Black and Litterman(1991). Our approach also uses the concept of asset market equilibrium.

In the following, we briefly explain the concept of asset market equilibrium. We assume that the number of investors is N and that the i th $\left(i \in \{1, \cdots, N\}\right)$ investor has wealth W_i. The investors invest M risk assets and one riskless asset. Here, $j = 0$ and $j = 1, \cdots, M$ denote one riskless asset and M risk assets, respectively. The weight of the portfolio held in the j th asset of the i th investor is denoted as $w_{(i)j}$. The return $R_{(i)}$ of the portfolio of the i th investor becomes

$$R_{(i)} = \sum_{j=0}^{M} w_{(i)j} R_j = R_0 + \sum_{j=1}^{M} w_{(i)j} \left(R_j - R_0\right), \tag{4}$$

where R_0 and R_j denote the returns of the riskless asset and the j th risk asset, respectively. From Eq. (4), we obtain

$$\widetilde{R}_{(i)} \triangleq R_{(i)} - R_0 = \mathbf{w}'_{(i)}\widetilde{\mathbf{R}}, \quad \widetilde{\mathbf{R}} \triangleq \begin{pmatrix} R_1 - R_0 \\ \vdots \\ R_M - R_0 \end{pmatrix}, \quad \mathbf{w}_{(i)} \triangleq \begin{pmatrix} w_{(j)1} \\ \vdots \\ w_{(j)M} \end{pmatrix}.$$

Here, $\widetilde{R}_{(i)}$ denotes the excess return of the portfolio of the i th investor. The wealth of the i th investor after one year becomes $W_{(i)}\left(1 + R_0 + \widetilde{R}_{(i)}\right)$.

We then consider the utility function $U\left(\widetilde{R}_{(i)}\right)$ that expresses the utility of the i th investor. We assume that the optimal portfolio of the i th investor is to maximize

$$U\left(\widetilde{R}_{(i)}\right) = E\left[\widetilde{R}_{(i)}\right] - \frac{1}{2}\lambda_{(i)} Var\left[\widetilde{R}_{(i)}\right], \quad \lambda_{(i)} > 0, \tag{6}$$

where $\lambda_{(i)}$ denotes the measure of the risk aversion of the i th investor. As the value of $\lambda_{(i)}$ increases, the i th investor considers the risks more carefully. We define

$$\boldsymbol{\alpha} \triangleq E\left[\widetilde{\mathbf{R}}\right], \quad \boldsymbol{\Omega} \triangleq E\left[\left(\widetilde{\mathbf{R}} - \boldsymbol{\alpha}\right)\left(\widetilde{\mathbf{R}} - \boldsymbol{\alpha}\right)'\right]. \tag{7}$$

Then, the expected return $E\left[\widetilde{R}_{(i)}\right]$ of $\widetilde{R}_{(i)}$ and the variance $\mathrm{Var}\left[\widetilde{R}_{(i)}\right]$ become

$$E\left[\widetilde{R}_{(i)}\right] = \mathbf{w}'_{(i)}\boldsymbol{\alpha}, \quad \mathrm{Var}\left[\widetilde{R}_{(i)}\right] = \mathbf{w}'_{(i)}\boldsymbol{\Omega}\mathbf{w}_{(i)}.$$

Hence, the optimal portfolio selection problem becomes a quadratic programming problem, as follows:

$$U\left(\widetilde{R}_{(i)}\right) = U\left(\mathbf{w}'_{(i)}\widetilde{\mathbf{R}}\right) = \mathbf{w}'_{(i)}\boldsymbol{\alpha} - \frac{1}{2}\lambda_{(i)}\mathbf{w}'_{(i)}\Omega\mathbf{w}_{(i)}, \quad \lambda_{(i)} > 0. \tag{8}$$

The optimal weight of the i th investor becomes

$$\mathbf{w}^*_{(i)} = \frac{1}{\lambda_{(i)}}\Omega^{-1}\boldsymbol{\alpha}. \tag{9}$$

This weight is the optimal solution of the Markowitz mean-variance approach, which is obtained by maximizing the expected utility of an investor.

Next, we describe the concept of asset market equilibrium. Let us denote the market value of the j th asset at the initial time by V_j. Then, the equilibrium conditions of the capital market are given by

$$V_0 = \sum_{i=1}^{N}\left(1-\sum_{j=1}^{M} w^*_{(i)j}\right)W_{(i)} = 0 \tag{10}$$

$$V_j = \sum_{i=1}^{N} w^*_{(i)j} W_{(i)}, \quad j \in \{1,\cdots,M\}. \tag{11}$$

Equation (10) indicates that the total sum of debit and credit in the money market is equal to zero. Equation (11) indicates that the market capitalization of stock is equal to the sum of the wealth of all investors. From the equilibrium conditions of Eqs. (10) and (11), we obtain

$$\sum_{j=1}^{M} V_j = \sum_{i=1}^{N} W_{(i)}. \tag{12}$$

Then, from Eqs. (11) and (12) we have

$$v_j = \sum_{i=1}^{N} \delta_{(i)} w^*_{(i)j}, \quad v_j = \frac{V_j}{\sum_{j=1}^{M} V_j}, \quad \delta_{(i)} = \frac{W_{(i)}}{\sum_{i=1}^{N} W_{(i)}}, \quad i \in \{1,\cdots,N\}, \quad j \in \{1,\cdots,M\}. \tag{13}$$

Next, we define $\mathbf{v}$ as follows:

$$\mathbf{v} \triangleq \begin{pmatrix} v_1 \\ \vdots \\ v_N \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^{N} \delta_{(i)} w^*_{(i)1} \\ \vdots \\ \sum_{i=1}^{N} \delta_{(i)} w^*_{(i)M} \end{pmatrix} = \sum_{i=1}^{N} \delta_{(i)} \begin{pmatrix} w^*_{(i)1} \\ \vdots \\ w^*_{(i)M} \end{pmatrix} = \sum_{i=1}^{N} \delta_{(i)} \mathbf{w}^*_{(i)} = \sum_{i=1}^{N} \frac{\delta_{(i)}}{\lambda_{(i)}} \mathbf{\Omega}^{-1}\boldsymbol{\alpha} = \frac{1}{\lambda}\mathbf{\Omega}^{-1}\boldsymbol{\alpha}, \tag{14}$$

$$\lambda = \left(\sum_{i=1}^{N} \frac{\delta_{(i)}}{\lambda_{(i)}}\right)^{-1}.$$

Then, from Eq. (14), we obtain

$$\boldsymbol{\alpha}^* = \lambda\, \mathbf{\Omega}\, \mathbf{v}, \tag{15}$$

where $\boldsymbol{\alpha}^*$ denotes the equilibrium expected excess returns of risk assets. Satchell and Scowcroft (2000) derived the value of the risk aversion factor λ from Eq. (16), as follows:

$$\mathbf{v} = (\lambda\boldsymbol{\Omega})^{-1}\boldsymbol{\alpha}^*$$

$$\alpha_p = \mathbf{v}'\boldsymbol{\alpha}^* = \boldsymbol{\alpha}^{*\prime}(\lambda\boldsymbol{\Omega})^{-1}\boldsymbol{\alpha}^* = \lambda^{-1}\boldsymbol{\alpha}^{*\prime}\boldsymbol{\Omega}^{-1}\boldsymbol{\alpha}^*$$

$$\sigma_p^2 = \mathbf{v}'\boldsymbol{\Omega}\mathbf{v} = \boldsymbol{\alpha}^{*\prime}(\lambda\boldsymbol{\Omega})^{-1}\boldsymbol{\Omega}(\lambda\boldsymbol{\Omega})^{-1}\boldsymbol{\alpha}^* = \lambda^{-1}\bullet\lambda^{-1}\boldsymbol{\alpha}^{*\prime}\boldsymbol{\Omega}^{-1}\boldsymbol{\alpha}^* = \lambda^{-1}\alpha_p \quad (16)$$

$$\therefore \quad \lambda = \frac{\alpha_p}{\sigma_p^2}$$

We can obtain $\boldsymbol{\alpha}^*$ from the market values $\mathbf{v}$ of risk assets and can obtain the covariance matrix $\boldsymbol{\Omega}$ of the risk assets and the risk aversion factor λ from α_p and σ_p^2 of the market portfolio.

If an investor believes firmly that the equilibrium expected excess returns of risk assets exist in the market, the investor may use $\boldsymbol{\alpha}^*$ as the expected excess returns of risk assets. However, practically speaking, $\boldsymbol{\alpha}^*$ is never used directly in the decision process of asset management. Instead, investors use their views about the expected excess returns of risk assets.

4 Application of Fuzzy Theory to the Investment Decision Process

4.1 Views of Risk Assets Held by Portfolio Managers

In general, portfolio managers select individual securities and decide their weights in a portfolio based on their views under a council system. Portfolio managers have vague views about the excess returns of risk assets. The proposed approach allows such vague views to be expressed in either absolute or relative terms among portfolio managers. We present two sample views as follows:

View No. 1: When the equilibrium expected excess return of the stock market index is 3.5%, a portfolio manager anticipates an absolute excess return of 3.0% .

View No. 2: A portfolio manager anticipates that the stock market index will outperform the bond market index by 5%.

View No. 2, which is a relative view, is often used practically to express the feelings of investors. In order to gauge whether View No. 2 will have a positive or negative

effect on stock market index relative to the bond market index, we need evaluate the respective equilibrium expected excess returns of the two assets in the view.

4.2 Applying Fuzzy Theory to the Views of Portfolio Managers

In this section, we explain through an example how to express the vague views of portfolio managers. We assume that the equilibrium expected excess returns $\boldsymbol{\alpha}^* = {}^t\left(\alpha_A^*, \alpha_B^*, \alpha_C^*\right)$ of three risk assets A, B, and C are as follows:

$$\boldsymbol{\alpha}^* = \begin{pmatrix} \alpha_A^* \\ \alpha_B^* \\ \alpha_C^* \end{pmatrix} = \begin{pmatrix} 5 \\ 4 \\ 3 \end{pmatrix} \tag{17}$$

Next, the views of five portfolio managers with respect to the expected excess returns of the three risk assets, A, B, and C, are as follows:

View No. 1
The first portfolio manager predicts that the excess return of asset A will be 2% larger than asset B and that the excess return of asset B will be 4%.

View No. 2
The second portfolio manager predicts that the excess return of asset A will be 3% smaller than asset B and that the excess return of asset C will be 5%.

⋮

View No. 5
The fifth portfolio manager predicts that the excess return of asset A will be 1% and that the excess return of asset C will be 4%.

Let us denote the expected excess returns of risk assets A, B, and C by $\boldsymbol{\alpha} = {}^t\left(\alpha_A\ \alpha_B\ \alpha_C\right)$ and the views of investors by $\boldsymbol{\eta}$. The relationship among the relative or absolute views of portfolio managers with respect to risk assets A, B, and C is expressed as follows by K_5 equations:

$$\mathbf{A}\boldsymbol{\alpha} = \boldsymbol{\eta}, \tag{18}$$

where $\mathbf{A}$ is a $K_5 \times 3$ matrix. The expected excess returns $\boldsymbol{\alpha}$ are fuzzy variables. Portfolio managers need not state their views about every asset in their investment universe. The equilibrium expected excess returns $\boldsymbol{\alpha}^*$ play a role as the constraint conditions of the reference points. Hence, we can combine Eqs. (17) and (18) as follows:

$$\begin{pmatrix} \boldsymbol{\alpha}^* \\ \boldsymbol{\eta} \end{pmatrix} = \begin{pmatrix} 5 \\ 4 \\ 3 \\ 2 \\ 4 \\ -3 \\ 5 \\ \vdots \\ 1 \\ 4 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & -1 & 0 \\ 0 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \\ \vdots & \vdots & \vdots \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \underbrace{\begin{pmatrix} \alpha_A \\ \alpha_B \\ \alpha_C \end{pmatrix}}_{\boldsymbol{\alpha}} \tag{19}$$

Generally speaking, in the case of N portfolio managers and M risk assets, Eq. (18) becomes

$$\mathbf{A}\boldsymbol{\alpha} = \boldsymbol{\eta}, \tag{20}$$

where $\mathbf{A}$ is a $K_N \times M$ matrix. The relationship among the views of N portfolio managers of M risk assets is expressed by K_N equations. Hence, we obtain

$$\mathbf{Y} = \begin{pmatrix} \boldsymbol{\alpha}^* \\ \boldsymbol{\eta} \end{pmatrix} = \begin{pmatrix} \mathbf{I} \\ \mathbf{A} \end{pmatrix} \boldsymbol{\alpha} = \mathbf{x}\boldsymbol{\alpha}, \tag{21}$$

where we assume that the expected excess return α_j of the jth risk asset is a fuzzy variable that reflects the views of portfolio managers and is given as follows:

$$\alpha_j = \left(b_j, c_j\right)_L \quad j = 1, \cdots, M,$$

where b_j is a center value because b_j takes the highest grade of the membership function, which expresses the fuzzy variable α_j, and c_j is the spread of the fuzzy variable α_j. We can assume various membership functions to express a fuzzy variable α_j. For example, we assume the membership function of $\mathbf{x}$, which is a symmetric triangular distribution, as follows:

$$\mu\left(\alpha_j\right) = L\left(\frac{\alpha_j - b_j}{c_j}\right) = \max\left(0, 1 - \frac{\left|\alpha_j - b_j\right|}{c_j}\right). \tag{22}$$

Here, we define h as follows:

$$h = \frac{\alpha_j - b_j}{c_j} \tag{23}$$

and $L(h)$ denotes the grade of the membership function. As $L(h)$ increases, the vagueness decreases.

Thus, we can obtain fuzzy variable $\mathbf{Y}$ in Eq. (21), as follows:

$$\mathbf{Y} = (\mathbf{bx}, \mathbf{c}|\mathbf{x}|)_L \tag{24}$$

$$\mathbf{b} = (b_1, \cdots, b_M), \quad \mathbf{c} = (c_1, \cdots, c_M), \quad |\mathbf{x}| = {}^t(|x_1|, \cdots, |x_M|).$$

We can also assume that the equilibrium expected excess returns $\boldsymbol{\alpha}^*$ are random variables. In the present paper, we assume that $\boldsymbol{\alpha}^*$ is a fuzzy number as the constraint conditions of reference points.

Then, we consider the following problem to obtain the optimal solution of $\boldsymbol{\alpha} = (\mathbf{b}, \mathbf{c})_L$

$$\text{Minimize} \quad J(\mathbf{c}) = \sum_{m=1}^{M} \{\mathbf{c}|\mathbf{x}_m|\} \tag{25}$$

subject to

$$\begin{aligned} &\mathbf{bx}_m - |L^{-1}(h)|\mathbf{c}|\mathbf{x}_m| \le y_m \quad m = 1, 2, \cdots, M \\ &\mathbf{bx}_m + |L^{-1}(h)|\mathbf{c}|\mathbf{x}_m| \ge y_m \quad m = 1, 2, \cdots, M \\ &|L^{-1}(h)| = 1 - h \\ &c_m \ge 0 \quad m = 1, \cdots, M \end{aligned} \tag{26}$$

where $J(\mathbf{c})$ denotes the vagueness level of a fuzzy linear regression model, and $\mathbf{c}|\mathbf{x}_m|$ expresses the aggregate of the width of fuzzy vector $\mathbf{bx}_m$.

Using the expected excess return $\alpha_j^h \quad (b_j - c_j \le \alpha_j^h \le b_j + c_j)$ of the j th risk asset corresponding to grade $L(h)$ of the membership function, we can obtain the optimal weight $\mathbf{w}^*$ from Eq. (9) when we assign a risk aversion. In addition to the value of the center b_j, which is the expected excess return of the j th risk asset, we can compare the cases of minimum $\alpha_j^{0.3} = b_j - 0.3c_j$ and maximum

$\alpha_j^{0.3} = b_j + 0.3c_j$ of the expected excess return in the case of $h = 0.3$. Since the covariance matrix $\mathbf{\Omega}$ between the j th and k th assets is relatively stable compared with their expected excess returns, we use the values of $\mathbf{\Omega}$ obtained by the historical data for the j th and k th assets. In practice, we must carefully consider the expected excess returns of risk assets, which greatly affect the performance of asset management.

5 Case Studies

5.1 Estimation of the Membership Function

In this section, the proposed approach is described in detail through a number of case studies. Generally speaking, it is difficult to obtain practical data on the views of portfolio managers. Moreover, the data shown in the following are virtual data. As such, we can obtain optimal values of $\boldsymbol{\alpha} = (\mathbf{b}, \mathbf{c})_L$: $\mathbf{b} = (b_1, \cdots, b_M)$, $\mathbf{c} = (c_1, \cdots, c_M)$ in Eqs. (25) and (26) by numerical nonlinear optimization (a quasi-Newton-Raphson method). In addition, we assume three risk assets, namely, A, B, and C, and assume that the values of their equilibrium expected excess returns $\boldsymbol{\alpha}^* = (\alpha_A^*, \alpha_B^*, \alpha_C^*)'$ are the same as those given by Eq. (17) for the previous sample.

Case No. 1
We assume that 10 portfolio managers (Nos. 1 though 10) have various views with respect to the expected excess returns of each risk assets, as shown in Table 1. The expected excess returns of each risk assets increase by 1 (%) from Portfolio Manager Nos. 1 through 10.

Table 1 Views and equilibrium expected excess returns of A, B, and C for 10 portfolio managers

Asset	Portfolio manager' view of expected excess returns (%)										Equilibrium expected excess returns (%)		
	No1	No2	No3	No4	No5	No6	No7	No8	No9	No10	A	B	C
A	0.0	1.0	2.0	3.0	4.0	6.0	7.0	8.0	9.0	10.0	5.0	4.0	3.0
B	−1.0	0.0	1.0	2.0	3.0	5.0	6.0	7.0	8.0	9.0			
C	−2.0	−1.0	0.0	1.0	2.0	4.0	5.0	6.0	7.0	8.0			

Here, we assume the membership function of a symmetric triangular distribution for the view of the portfolio manager with regard to expected excess returns. The results for the case in which $h = 0.0$ in Eqs. (25) and (26) are shown in Table 2. For example, the center value b_A of the expected excess return of risk asset A is 5.0 (%) and the spread c_A is 5.0 (%).

Table 2 Values of b_j, c_j

	Asset		
(%)	A	B	C
b_j	5.0	4.0	3.0
c_j	5.0	5.0	5.0

Case No. 2

Table 3 shows the results for the case in which $h = 0.0$, excluding the views of Portfolio Manager Nos. 1 and 10.

Table 3 Values of b_j, c_j

	Asset		
(%)	A	B	C
b_j	5.0	4.0	3.0
c_j	4.0	4.0	4.0

Case No. 3

Table 4 shows the results for the case in which $h = 0.0$, excluding the views of Portfolio Manager Nos. 1, 2, 9, and 10.

Table 4 Values of b_j, c_j

	Asset		
(%)	A	B	C
b_j	5.0	4.0	3.0
c_j	3.0	3.0	3.0

As an example, the membership functions of the expected excess return of risk asset A in Tables 2 through 4 are shown in Figure 2. In all cases, the center values b_A of the expected excess returns of risk asset A are identical, i.e., 5 (%). However, these spread values become smaller by 1 (%). The spread values depend on the maximum value and the minimum value among the views of the portfolio managers.

Practically speaking, portfolio managers do not have views of the expected excess returns by same interval. Therefore, let us show the results for the case in which the differences between the expected excess returns of each portfolio managers are different. We assume that 10 portfolio managers have views concerning the expected excess returns of each risk asset, as shown in Table 5.

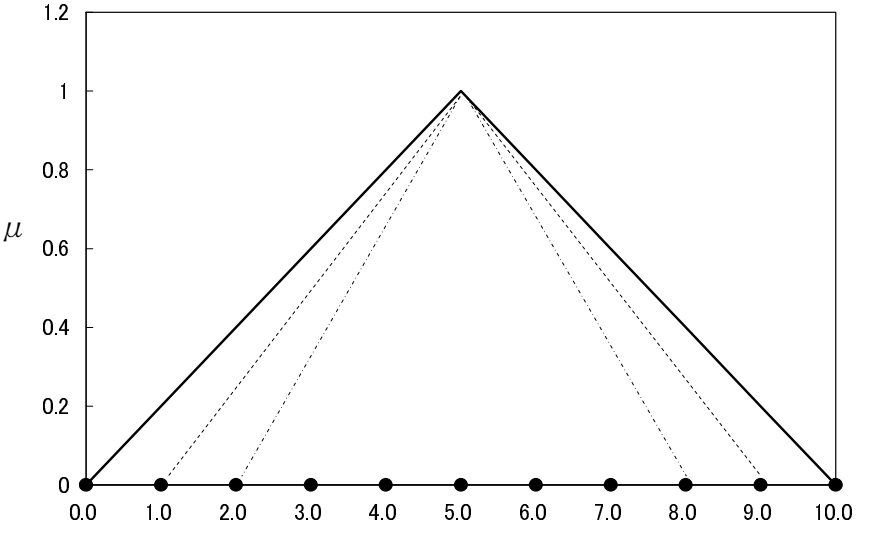


Fig. 2 Membership functions of risk asset A in Tables 2 through 4

Table 5 Views and equilibrium expected excess returns of A, B, and C for 10 portfolio managers

Asset	Portfolio manager' view of expected excess returns (%)										Equilibrium expected excess returns (%)		
	No1	No2	No3	No4	No5	No6	No7	No8	No9	No10	A	B	C
A	0.0	2.6	3.6	4.6	4.8	5.2	5.4	6.4	7.4	10.0	5.0	4.0	3.0
B	−1.0	1.6	2.6	3.6	3.8	4.2	4.4	5.4	6.4	9.0			
C	−2.0	0.6	1.6	2.6	2.8	3.2	3.4	4.4	5.4	8.0			

Case No. 4
Table 6 shows the results for the case in which $h = 0.0$ in Eqs. (25) and (26).

Table 6 Values of b_j, c_j

	Asset		
(%)	A	B	C
b_j	5.0	4.0	3.0
c_j	5.0	5.0	5.0

Case No. 5
Table 7 shows the results for the case in which $h = 0.0$, excluding the views of Portfolio Manager Nos. 1 and 10.

Table 7 Values of b_j, c_j

	Asset		
(%)	A	B	C
b_j	5.0	4.0	3.0
c_j	2.4	2.4	2.4

Case No. 6
Table 8 shows the results for the case in which $h = 0.0$, excluding the views of Portfolio Manager Nos. 1, 2, 9, and 10.

Table 8 Values of b_j, c_j

	Asset		
(%)	A	B	C
b_j	5.0	4.0	3.0
c_j	1.4	1.4	1.4

As an example, the membership functions of expected excess returns of risk asset A in Tables 6 through 8 are shown in Figure 3. In all cases, the center values b_A of the expected excess return of risk asset A are 5 (%). However, since the spread values depend on the maximum value and the minimum value among the views of the portfolio managers, the spread values decrease by different intervals.

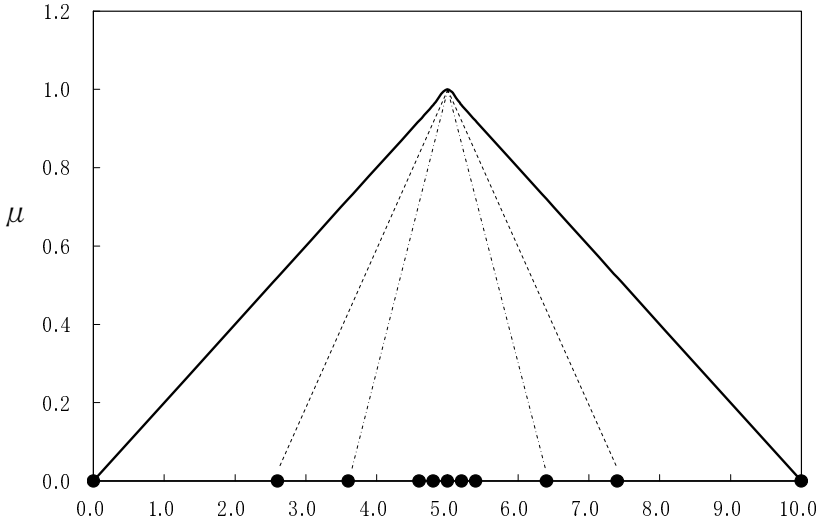


Fig. 3 Membership functions of risk asset A in Tables 6 through 8

Next, let us consider the cases of the relative views between assets of portfolio managers as shown in Table 9.

Table 9 Views and equilibrium expected excess returns of A, B, and C for 10 portfolio managers

Asset	Portfolio manager' relative view of expected excess returns (%)										Equilibrium expected excess returns (%)		
	No1	No2	No3	No4	No5	No6	No7	No8	No9	No10	A	B	C
A−B	1.0	1.0	1.0	1.0	1.0						5.0	4.0	3.0
A−C	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0			
B−C						1.0	1.0	1.0	1.0	1.0			

Case No. 7
We assume the membership function of a symmetric triangular distribution with respect to the expected excess returns of risk assets A, B, and C. The results for the case in which $h = 0.0$ in Eqs. (25) and (26) are shown in Table 10.

Table 10 Values of b_j, c_j

	Asset		
(%)	A	B	C
b_j	5.0	4.0	3.0
c_j	0.0	0.0	0.0

Next, let us consider the case shown in Table 11.

Table 11 Views and equilibrium expected excess returns of A, B, and C for 10 portfolio managers

Asset	Portfolio manager' relative view of expected excess returns (%)										Equilibrium expected excess returns (%)		
	No1	No2	No3	No4	No5	No6	No7	No8	No9	No10	A	B	C
A−B	−2.0	−1.0	0.0	1.0	2.0						5.0	4.0	3.0
A−C	−2.0	−1.0	0.0	1.0	2.0	−2.0	−1.0	0.0	1.0	2.0			
B−C						−2.0	−1.0	0.0	1.0	2.0			

Case No. 8

Table 12 shows the results for the case in which $h = 0.0$. The center values b_j of the expected excess return of all risk assets are 4.0 (%) and the spreads c_j are 1.0 (%).

Table 12 Values of b_j, c_j

	Asset		
(%)	A	B	C
b_j	4.0	4.0	4.0
c_j	1.0	1.0	1.0

Finally, we show the results for the case in which 10 portfolio managers have a mixture of absolute and relative views about the expected excess returns of each risk asset, as shown in Table 13. Table 14 shows the results for the case in which $h = 0.0$. The centers b_j and spreads c_j of the expected excess returns of risk for assets A, B, and C are not the same. Moreover, the center b_C of the expected excess returns of risk for asset C is different from the equilibrium expected excess return.

Table 13 Views and equilibrium expected excess returns of A, B, and C for 10 portfolio managers

Asset	Portfolio manager s' absolute and relative views (%)										Equilibrium Expected excess returns		
	No1	No2	No3	No4	No5	No6	No7	No8	No9	No10	A	B	C
A−B	−2.0	−1.0	0.0	1.0	2.0						5.0	4.0	3.0
A−C						−2.0	−1.0	0.0	1.0	2.0			
A	3.0	4.0	6.0	7.0									
B					3.0	4.0	5.0						
C								2.0	3.0	4.0			

Case No. 9

Table 14 Values of b_j, c_j

	Asset		
(%)	A	B	C
b_j	5.0	4.0	3.5
c_j	2.0	1.0	1.5

5.2 A Fuzzy Information Criterion

By the way, it is not clear whether the membership function of a symmetric triangular distribution expresses the true membership function of portfolio managers' views. Some portfolio managers may have views that are off the mark. Such outliers distort the distribution of the membership function. It is difficult to estimate the distribution of the membership function by the fuzzy linear regression model given by Eqs. (25) and (26).

Therefore, we propose a fuzzy information criterion (FIC) to select the membership function that properly expresses the distribution of the observed fuzzy data. Let us denote an unknown true membership function by G and a hypothetical membership function by F. The observed data $\mathbf{x}_n = \{x_1, x_2, \cdots, x_n\}$ are generated from an unknown true membership function G. The criteria of goodness with respect to the hypothetical membership function F is to evaluate the fitness for the distribution of an unknown true membership function G. Therefore, we standardize each membership function as follows:

$$N = \sum_j G(x_j), \quad M = \sum_j F(x_j)$$

$$g(x_j) = \frac{G(x_j)}{N}, \quad f(x_j) = \frac{F(x_j)}{M} \tag{27}$$

We use Kullback-Leibler information (K-L information) (1951) as a measure of the fitness between membership functions as follows:

$$I(g; f) = \int \log\left\{\frac{g(x)}{f(x)}\right\} dg(x) = E_g\left[\log\frac{g(x)}{f(x)}\right] \tag{28}$$

where E_g is the weighted average by the standardized grades of an unknown true membership function g.

The K-L information has the following characteristics:

(i) $I(g;f) \geq 0$ (29)

(ii) $I(g;f) = 0 \quad \Leftrightarrow \quad g(x) = f(x)$ (30)

Based on these characteristics, as the value of the K-L information becomes smaller, the distribution of the hypothetical membership function f becomes more similar to an unknown true membership function g. In other words, we can evaluate the fitness of the hypothetical membership function f based on the value of the K-L information. However, generally, it is limited to use K-L information values. Since the K-L information involves an unknown true distribution of the membership function g, it is impossible to calculate the value directly.

Then, the K-L information is divided as follows;

$$I(g;f) = E_g\left[\log\frac{g(X)}{f(X)}\right] = E_g\left[\log g(X)\right] - E_g\left[\log f(X)\right]. \tag{31}$$

Since the first term is a constant that depends on unknown true membership function g, it is sufficient to consider only the second term when evaluating the hypothetical membership function f. As the second term increases, the K-L information value becomes smaller.

Then, since the value of $E_g\left[\log f(X)\right]$ depends on the unknown true membership function g, we cannot calculate it directly. However, if we can obtain a meaningful estimate of unknown true membership function g from the data, we can use this estimate as an estimated criteria of the membership function. Therefore, we assume the membership function $\hat{g}(x_\tau) = \frac{1}{n}$ $(\tau = 1, 2, \cdots, n)$, which has the same grade $\frac{1}{n}$ for all of the observed data $\mathbf{x}_n = \{x_1, x_2, \cdots, x_n\}$. We refer to this as an empirical membership function. Substituting the empirical membership function $\hat{g}$ for an unknown true membership function g, we obtain

$$E_{\hat{g}}\left[\log f(X)\right] = \int \log f(x)\, d\hat{g}(x) = \sum_{\tau=1}^{n} \hat{g}(x_\tau)\log f(x_\tau) = \frac{1}{n}\sum_{\tau=1}^{n}\log f(x_\tau). \tag{32}$$

Here, we define

$$n\int \log f(x)\, d\hat{g}(x) = \sum_{\tau=1}^{n}\log f(x_\tau). \tag{33}$$

We then assume the membership function $f(x \mid \boldsymbol{\theta})$ $\left(\theta \in \Theta \subset R^p\right)$, which has an unknown p-dimension parameter $\boldsymbol{\theta} = \left(\theta_1, \theta_2, \cdots, \theta_p\right)'$. Next, we define

$$l(\boldsymbol{\theta}) = \sum_{\tau=1}^{n} \log f(x_\tau \mid \boldsymbol{\theta}) \cdot \tag{34}$$

We can obtain an estimate of $\hat{\boldsymbol{\theta}} \in \Theta$ by maximizing $l(\boldsymbol{\theta})$ as follows:

$$l(\hat{\boldsymbol{\theta}}) = \max_{\theta \in \Theta} l(\boldsymbol{\theta}). \tag{35}$$

We can evaluate the fitness of the membership function $f(x \mid \boldsymbol{\theta})$ by comparing to the value of $l(\hat{\boldsymbol{\theta}})$. In general, $l(\hat{\boldsymbol{\theta}})$ has a bias as an estimation value of $nE_g\left[\log f_j\left(z \mid \hat{\boldsymbol{\theta}}_j\right)\right]$, and the size of its bias changes depending on the dimension of the parameter $\boldsymbol{\theta}$.

Finally, we propose a fuzzy information criterion (FIC) considering the revision of the bias between $nE_g\left[\log f_j\left(z \mid \hat{\boldsymbol{\theta}}_j\right)\right]$ and $\sum_{\tau=1}^{n} \log f\left(x_\tau \mid \hat{\boldsymbol{\theta}}\right)$ as follows:

$$FIC = -2\sum_{\tau=1}^{n} \log f\left(x_\tau \mid \hat{\boldsymbol{\theta}}\right) + 2p. \tag{36}$$

Basically, this concept of FIC is similar to the Akaike information criterion used in statistics. The FIC is based on the assumption that $g(x) = f(x \mid \boldsymbol{\theta}_0)$ for $\boldsymbol{\theta}_0 \in \Theta$, and the law of large numbers functions in the same manner as in statistics. In the present paper, we omit the process of the guidance of the FIC.

We then show the values of FIC for some different membership functions based on the views of the 10 portfolio managers shown in Table 5. The results obtained by a fuzzy linear regression model assuming $L^{-1}(0.1) = 0.9$ in Eqs. (25) and (26) are shown in Table 15 and Figure 4. The grade of the maximum and the minimum value among the views of portfolio managers become 0.1.

Table 15 Values of b_j, c_j

	Asset		
(%)	A	B	C
b_j	5.0	4.0	3.0
c_j	5.6	5.6	5.6

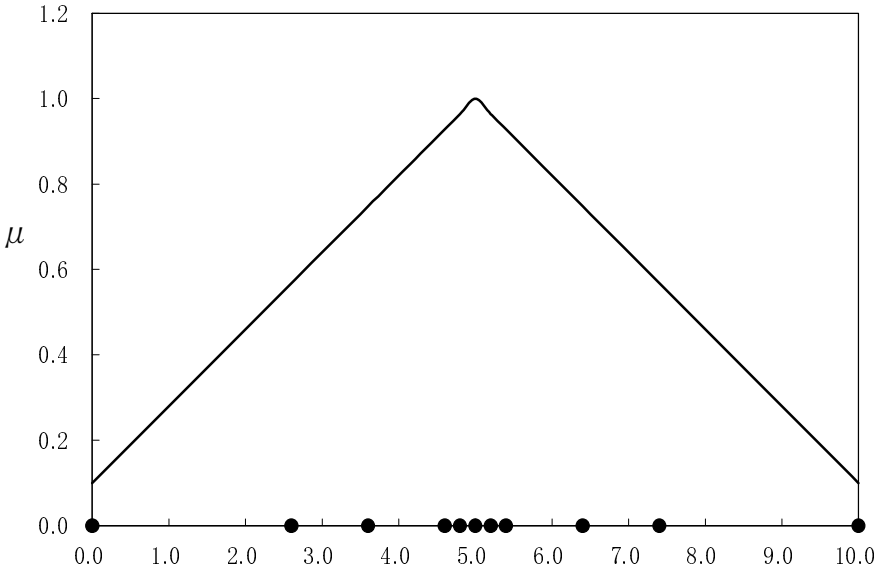


Fig. 4 Membership function of risk asset A in Table 5

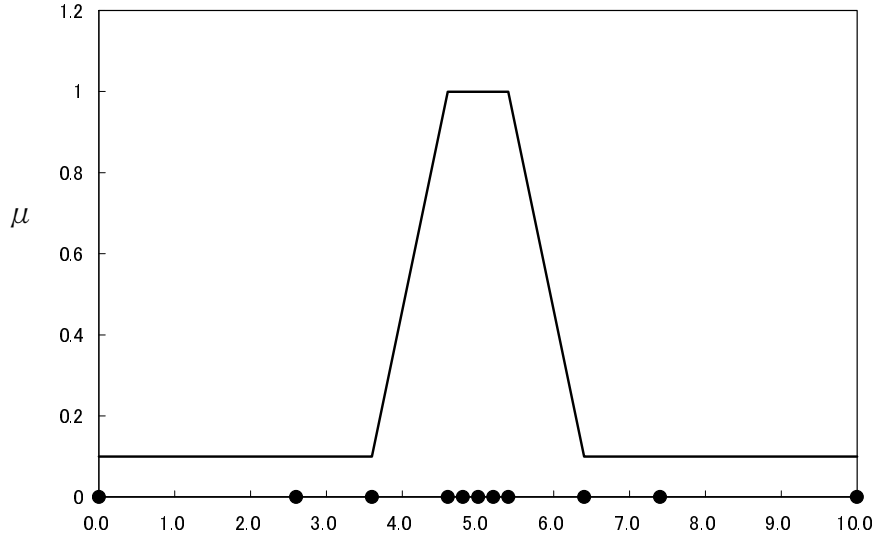


Fig. 5 Membership function of risk asset A

Concerning the membership function of Table 15, the value of $l\left(\hat{\boldsymbol{\theta}}\right)$ in Eq. (35) is −86.6 and the value of FIC is 179.24. If we assume the membership function shown in Figure 5, then the value of $l\left(\hat{\boldsymbol{\theta}}\right)$ is –98.3 and the value of FIC is 202.6. Hence we can select the membership function shown in Figure 4 in which the value of FIC is smaller .

Moreover, we can assume a normal fuzzy variable with the membership function defined by

$$\mu\left(\alpha_j\right) = \exp\left(-\frac{\left(\alpha_j - b_j\right)^2}{c_j^2}\right). \tag{37}$$

For example, in Fig. 6, we show the membership function corresponding to a normal distribution for the case in which $b_A = 5\,(\%)$, $c_A = 5.6\,(\%)$ in Table 15. The value of $l\left(\hat{\boldsymbol{\theta}}\right)$ is –80.4 and the value of FIC is 166.71. Comparing the values of $l\left(\hat{\boldsymbol{\theta}}\right)$ and FIC, the normal distribution in Figure 6 is judged to be the best among the distributions of the membership function shown in Figures 4 through 6.

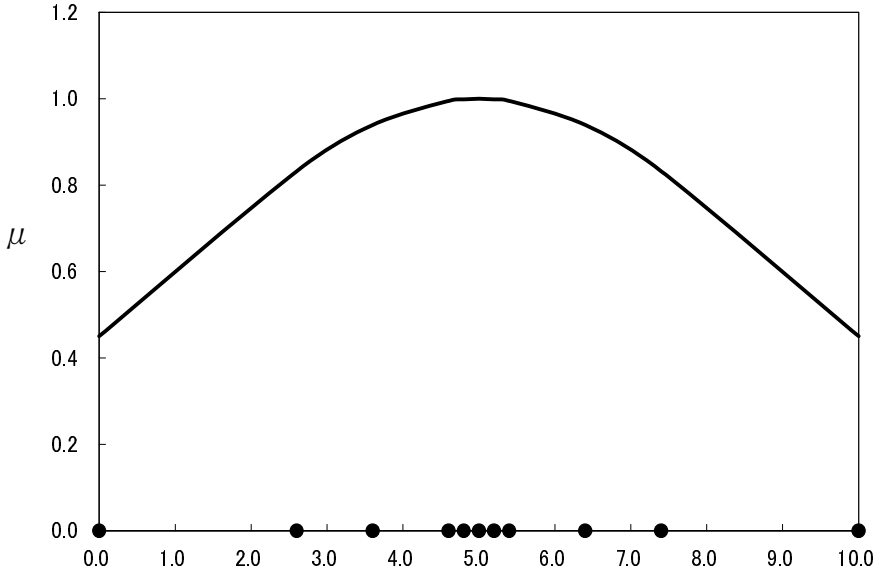


Fig. 6 Membership function corresponding to the normal distribution of risk asset A

6 Conclusion

Black and Litterman (1991) proposed an approach to combine the equilibrium expected excess returns of risk assets and the views of portfolio managers using a regression model. When each view is incorporated, their approach requires an intuitive level of confidence of the investor. Practical speaking, it is difficult to input an intuitive level of confidence of an investor.

On the other hand, the proposed fuzzy approach to obtain the expected excess returns of risk assets is an improvement over the traditional mean-variance approach in that we incorporate the vague views of an investor concerning absolute or relative excess returns on different assets. Although the proposed approach brings several new features to the traditional asset allocation problem, the main contribution of the proposed approach is that investors are able to compare and combine their vague outlooks for risk assets with equilibrium expected excess returns. The use of the equilibrium expected excess returns associated with the capital market as a reference point is very meaningful for portfolio managers. The proposed approach enables a group of portfolio managers of pension funds to obtain an important solution that realizes optimal expected excess returns of risk assets by specifying the vague views of portfolio managers as a fuzzy number.

References

1. Baks, K.P., Metrick, A., Watcher, J.: Should investors avoid all actively managed mutual funds? A study in Bayesian performance evaluation. Journal of Finance 56, 45–85 (2001)
2. Barry, C.B.: Portfolio analysis under uncertain means, variances, and covariances. Journal of Finance 29, 515–522 (1974)
3. Black, F., Litterman, R.: Asset Allocation; combining investor views with market equilibrium. Journal of Fixed Income 1, 7–18 (1991)
4. Hasuike, T., Katagiri, H., Ishii, H.: Protfolio selection problems with random fuzzy variable returns. Fuzzy Sets and Systems 160, 2579–2596 (2009)
5. Idzorek, T.: A Step-by-Step Guide to the Black-Litterman Model, Working paper (2004)
6. Inuiguchi, M., Ramki, J.: Possibilistic programming: a brief review of fuzzy mathematical programming and a comparison with stochastic programming in portfolio selection problem. Fuzzy Sets and Systems 111, 3–28 (2000)
7. Inuiguchi, M., Tanino, T.: Portfolio selection under independent possibilistic information. Fuzzy Sets and Systems 115, 83–92 (2000)
8. Kullback, S., Leibler, R.A.: On information and sufficiency. Ann. Math. Statist. 22, 79–86 (1951)
9. Konishi, S., Kitagawa, G.: Information Criteria, Asakura (2004) (in Japanese)
10. Mankert, C.: The Black-Litterman Model, Working paper, Royal Institute of Technology, Stockholm, Sweden (2006)
11. Markowitz, H.: Portfolio selection. Journal of Finance 7, 77–91 (1952)
12. Markowitz, H.: Portfolio Selection. In: Efficient Diversification of Investments. Wiley, Chichester (1959)

13. Nakatsuma, T.: Bayesian Statistics and Finance. JAFEE Journal 8, 1–62 (2009) (in Japanese)
14. Perold, A.: Large Scale Portfolio Optimization. Management Science 30(10), 1143–1160 (1984)
15. Sakawa, M.: Basement and Application of Fuzzy Theory, Morikita (1989) (in Japanese)
16. Satchell, S., Scowcroft, A.: A Demystification of the Black-Litterman Model:Managing Quantitative and Traditional Construction. Journal of Asset Management, 138–150 (September 2000)
17. Sharpe, W.F.: Capital asset prices: a theory of market equilibrium under conditions of risk. Journal of Finance 19, 425–442 (1964)
18. Tsuda, H.: Time Series Analysis of Financial Asset Price Fluctuations. The Practice of Time Series Analysis, pp. 283–297. Springer, Heidelberg (1999)
19. Tsuda, H.: Statistics of Equity, Asakura Shoten (1994) (in Japanese)
20. Watada, J.: Fuzzy Portfolio Model for Decision Making in Investment. In: Dynamical Aspects in Fuzzy Decision Making, pp. 141–162. Physica-Verlag, Heidelberg (2001)
21. Watada, J.: Fuzzy portfolio selection and its applications to decision making. Tatra Mountains Math. Publ. 13, 219–248 (1997)
22. Zadeh, L.A.: Fuzzy Sets. Information and Control 8(3), 338–353 (1965)
23. Zadeh, L.A.: Fuzzy Sets as a Basis for a Theory of Possibility. Fuzzy Sets and Systems 1 (1978)

Decision Making Techniques in Political Management

Anna M. Gil-Lafuente and José M. Merigó

Abstract. In this paper, we develop a new decision making model and apply it in political management. We use a framework based on the use of ideals in the decision process and several similarity measures such as the Hamming distance, the adequacy coefficient and the index of maximum and minimum level. For each similarity measure, we use different types of aggregation operators such as the simple average, the weighted average, the ordered weighted averaging (OWA) operator and the generalized OWA (GOWA) operator. This new approach considers several attributes and different scenarios that may occur in the uncertain environment. We see that depending on the particular type of aggregation operator used the results may lead to different decisions.

Keywords: Decision making, OWA operator, Hamming distance, Political management.

1 Introduction

Decision making problems are very common in the literature (Figueira et al., 2005; Gil-Aluja, 1999; Kaufmann and Gil-Aluja, 1986; Merigó, 2008) and can be applied in a lot of fields. For example, we can use them for the selection of policies in a government, in a company, etc. Selecting the best policy in a government is one of the key problems to be solved in order for a good development of a country. There are a lot of different types of policies such as fiscal policies, monetary policies and commercial policies. In order to select the optimal policy, the government has to develop a selection process because they have to choose the best policy in each moment. Among the great variety of studies existing in selection,

Anna M. Gil-Lafuente · José M. Merigó
Department of Business Administration, University of Barcelona,
Av. Diagonal 690, 08034 Barcelona, Spain
e-mail: amgil@ub.edu, jmerigo@ub.edu

W.A. Lodwick & J. Kacprzyk (Eds.): Fuzzy Optimization, STUDFUZZ 254, pp. 389–405.
springerlink.com

this work will follow those models that develop the decision process using ideals (Gil-Aluja, 1998; 1999; A.M. Gil-Lafuente, 2005; A.M. Gil-Lafuente and Merigó, 2006; J. Gil-Lafuente, 2001; 2002; Merigó, 2008; Merigó and A.M. Gil-Lafuente, 2006; 2007a; 2008a; 2008b; 2008c; 2008d; 2010).

Sometimes, depending on the situation that happens in the future, the government has to adapt to different necessities according to the environment of the economy and this also affects their needs regarding policy management. The aim of this paper is to develop different techniques to solve this problem. Our starting point is a selection process based on attributes where we assume that the requirements for each attribute is different depending on the environment of the economy. This methodology follows A.M. Gil-Lafuente and Merigó (2006) with the difference that now instead of considering a selection process of financial products we analyze the selection of policies. Note that we may consider a wide range of policies such as fiscal policies, monetary policies and commercial policies.

This selection process is based on the use of ideals. Thus, we use several similarity measures for analyzing the information. Particularly, we consider the Hamming distance (Hamming, 1950), the adequacy coefficient (Kaufmann and Gil-Aluja, 1986; 1987) and the index of maximum and minimum level (J. Gil-Lafuente, 2001; 2002). Moreover, we also use several aggregation operators in the analysis. We use the average, the weighted average (WA), the ordered weighted average (OWA) (Yager, 1988) and the generalized OWA (GOWA) (Karayiannis, 2000; Merigó, 2008; Yager, 2004). By using OWAs (Beliakov et al., 2007; Fodor et al., 1995; Merigó, 2008; Merigó and A.M. Gil-Lafuente, 2009a; 2009b; Yager, 1993; Yager and Kacprzyk, 1997), we are able to provide a parameterized family of aggregation operators that range from the minimum to the maximum.

This paper is organized as follows. In Section 2 we briefly revise some basic preliminaries. Section 3 presents the new decision making approach by using various types of aggregation operators. Section 4 develops a numerical example of the new decision model and Section 5 summarizes the main conclusions of the paper.

2 Preliminares

In this Section, we briefly describe some basic concepts to be used throughout the paper such as the Hamming distance, the adequacy coefficient, the index of maximum and minimum level and their extensions with the OWA operator.

2.1 The Hamming Distance

The Hamming distance (Hamming, 1950) is a useful technique for calculating the differences between two elements, two sets, etc. For two sets A and B, the weighted Hamming distance can be defined as follows.

Definition 1. A weighted Hamming distance of dimension n is a mapping WHD: $R^n \rightarrow R$ that has an associated weighting vector W of dimension n with $\sum_{j=1}^{n} w_j = 1$ and $w_j \in [0, 1]$, such that:

$$WHD\ (A, B) = \left(\sum_{i=1}^{n} w_i \mid a_i - b_i \mid \right) \tag{1}$$

where a_i and b_i are the ith arguments of the sets A and B respectively.

Note that the formulations shown above are the general expressions. For the formulation used in fuzzy set theory see for example (Kaufmann, 1975; Szmidt and Kacprzyk, 2000). Note also that if $w_i = 1/n$, for all i, then, the weighted Hamming distance becomes the normalized Hamming distance.

2.2 The Adequacy Coefficient

The adequacy coefficient (Gil-Aluja, 1998; A.M. Gil-Lafuente, 2005; J. Gil-Lafuente, 2002; Kaufmann and Gil-Aluja, 1986; 1987) is an index used for calculating the differences between two elements, two sets, etc. It is very similar to the Hamming distance with the difference that it neutralizes the result when the comparison shows that the real element is higher than the ideal one. For two sets A and B, the weighted adequacy coefficient can be defined as follows.

Definition 2. A weighted adequacy coefficient of dimension n is a mapping K: $R^n \rightarrow R$ that has an associated weighting vector W of dimension n with $\sum_{j=1}^{n} w_j = 1$ and $w_j \in [0, 1]$, such that:

$$K\ (A, B) = \sum_{i=1}^{n} w_i [1 \wedge (1 - \mu_i + \mu_i{}^{(k)}] \tag{2}$$

where a_i and b_i are the ith arguments of the sets A and B respectively.

Note that if $w_i = 1/n$, for all i, then, the weighted adequacy coefficient becomes the normalized adequacy coefficient.

2.3 The Index of Maximum and Minimum Level

The index of maximum and minimum level is an index that unifies the Hamming distance and the adequacy coefficient in the same formulation (J. Gil-Lafuente, 2001; 2002). In fuzzy set theory, it can be useful, for example, for the calculation of distances between fuzzy sets, interval-valued fuzzy sets or intuitionistic fuzzy sets. For two sets A and B, the weighted index of maximum and minimum level can be defined as follows.

Definition 3. A WIMAM of dimension n is a mapping K: $R^n \to R$ that has an associated weighting vector W of dimension n with $\sum_{j=1}^{n} w_j = 1$ and $w_j \in [0, 1]$, such that:

$$\eta(A, B) = \sum_{u} Z_i(u) \times \left| \mu_i(u) - \mu_i^{(j)}(u) \right| + \sum_{v} Z_i(v) \times \left[0 \vee (\mu_i(v) - \mu_i^{(j)}(v)) \right] \tag{3}$$

where a_i and b_i are the ith arguments of the sets A and B respectively.

Note that if $w_i = 1/n$, for all i, then, the weighted index of maximum and minimum level becomes the normalized index of maximum and minimum level.

2.4 The OWA Operator

The OWA operator (Yager, 1988) provides a parameterized family of aggregation operators which have been used in many applications (Beliakov et al., 2007; Karayiannis, 2000; Merigó, 2008; Xu, 2005; Yager, 1993; 2004; Yager and Kacprzyk, 1997). It can be defined as follows.

Definition 4. An OWA operator of dimension n is a mapping OWA: $R^n \to R$ that has an associated weighting vector W of dimension n having the properties:

1) $w_j \in [0, 1]$
2) $\sum_{j=1}^{n} w_j = 1$

and such that

$$OWA(a_1, a_2, \ldots, a_n) = \sum_{j=1}^{n} w_j b_j \tag{4}$$

where b_j is the jth largest of the a_i.

From a generalized perspective of the reordering step we can distinguish between the descending OWA (DOWA) operator and the ascending OWA (AOWA) operator. Note that the weights of these two operators are related by $w_j = w^*_{n-j+1}$, where w_j is the jth weight of the DOWA and w^*_{n-j+1} the jth weight of the AOWA operator.

Note that the OWA operator can be generalized by using generalized means, obtaining the generalized OWA (GOWA) operator. It can be defined as follows.

Definition 5. A GOWA operator of dimension n is a mapping $GOWA$: $R^n \to R$ that has an associated weighting vector W of dimension n having the properties:

1) $w_j \in [0, 1]$
2) $\sum_{j=1}^{n} w_j = 1$

and such that

$$GOWA\ (a_1, a_2, \ldots, a_n) = \left(\sum_{j=1}^{n} w_j b_j^{\lambda} \right)^{1/\lambda} \tag{5}$$

where b_j is the jth largest of the a_i and λ is a parameter such that $\lambda \in (-\infty, \infty)$.

3 Decision Making in Political Management

In this section, we describe the decision making process to follow in political management. First, we present the decision making approach and then present some aggregation operators that can be used in the process.

3.1 Decision Making Approach

We introduce a new type of decision making approach. This type of selection is very similar to the traditional methods based on different attributes with the difference that here it can occur different scenarios in the future. Thus, the selection of policies needs to consider this aspect and search for the policy that better adapts to the different situations that could occur. The mathematical process will be equal with the difference that here we will have more than one ideal fuzzy subset. With this additional information, the fuzzy subsets of each product will be compared with all the ideal fuzzy subsets that we have. Therefore, the process to follow will consist in the following steps (Gil-Aluja, 1998; A.M. Gil-Lafuente and Merigó, 2006; J. Gil-Lafuente, 2002):

Step 1: Analysis and determination of the significant characteristics of the interesting policies for the government. Theoretically, it will be represented as: $C = \{C_1, C_2, \ldots, C_i, \ldots, C_n\}$, where C_i is the ith characteristic to consider in each policy and we suppose a limited number n of required characteristics.

Step 2: Identification of the different possible scenarios that could occur in the future where we should need different ideal fuzzy subsets. These z fuzzy subsets will be considered as: $h = 1, 2, \ldots, z$; where each fuzzy subset would represent the necessities of the government in each situation h.

Step 3: Fixation of the ideal levels of each significant characteristic in order to form the different ideal policies for all the possible situations that could occur. Mathematically, it will be represented as:

	C_1	C_2	…	C_i	…	C_n
$P_h =$	μ_1^h	μ_2^h	…	μ_i^h	…	μ_n^h

where P_h is the hth ideal policy expressed by a fuzzy subset, C_i is the ith characteristic to consider and $\mu_i^h \in [0, 1]$; $i = 1, 2, \ldots, n$, is the valuation between 0 and 1 for the ith characteristic of the hth ideal policy.

Step 4: Grouping of all the different ideal policies in one fuzzy subset.

	C_1	…	C_n	…	C_1	…	C_n
$P_{1,\ldots,z} =$	μ_1^1	…	μ_n^1	…	μ_1^z	…	μ_n^z

where $P_{1,2,\ldots,z}$ refers to the grouping of the hth ideal policy expressed by a fuzzy subset, C_i is the ith characteristic to consider and $\mu_i^h \in [0, 1]$; $i = 1, 2, \ldots, n$, is the valuation between 0 and 1 for the ith characteristic of the hth ideal policy.

Step 5: Fixation of the real valuation of each characteristic for all the different policies considered. Theoretically, it is represented as:

	C_1	C_2	…	C_i	…	C_n
$P_j =$	$\mu_1^{(k)}$	$\mu_2^{(k)}$	…	$\mu_i^{(k)}$	…	$\mu_n^{(k)}$

with $k = 1, 2, \ldots, m$; where P_k is the kth policy expressed by a fuzzy subset, C_i is the ith characteristic to consider and $\mu_i \in [0, 1]$; $i = 1, 2, \ldots, n$, is the valuation between 0 and 1 for the ith characteristic of the kth policy.

Step 6: Construction of a fuzzy subset for each policy that is adapted to the ideal policy with different situations. It consists in creating a fuzzy subset with the subset of the policy k repeated z times.

	C_1	…	C_n	…	C_1	…	C_n
$P_j^{(p)} =$	$\mu_1^{(j)}$	…	$\mu_n^{(j)}$	…	$\mu_1^{(j)}$	…	$\mu_n^{(j)}$
		1		…		z	

Step 7: Comparison between the ideal policy and the different policies and determination of the level of removal.

In this step, we have to express numerically the approximation between the ideal policy and the different policies considered. To solve this problem, we have a lot of different selection indexes that can be used. In this paper, we will use the Hamming distance, the adequacy coefficient and the index of maximum and minimum level. In these three cases, we will consider the situation that the characteristics have the same level of importance, the situation with different degrees of importance, the situation found with the OWA operator and the general situation found with the GOWA operator. In the next sections, we will consider in more detail the use of these indexes in the selection of policies.

Step 8: Adoption of decisions according to the results found in the previous steps.

3.2 The Hamming Distance in Political Management

A first index that can be used in the selection process is the Hamming distance. Here we consider four cases where the Hamming distance could be used. First, we could consider the case where the characteristics have the same level of importance. Thus, we could define the Hamming distance as:

$$\delta(P_{1,2,\ldots,z}, P_k^{(p)}) = \frac{1}{zn} \sum_{h=1}^{z} \sum_{i=1}^{n} \left| \mu_i{}^h - \mu_i^{(k)} \right| \tag{6}$$

with: $i = 1,2,\ldots,n$; and $\forall\ (P_{1,2,\ldots,z}, P_k^{(p)}) \in \mu_i{}^h, \mu_i^{(k)} \in [0, 1]$.

Analysing the results, we see that they refer to the removal of each policy to the ideal one for all the different situations that could occur. The result will be between 0 and 1. Values near 0 imply that the policy is interesting because it has similar valuations than the ideal one. Values near 1 will mean that the policy is not interesting for the company.

A second case that we could consider is the case where the characteristics have different degrees of importance. In order to use different degrees of importance, we can use a convex weighting so the result is still in [0, 1]:

$$V_i^h = \frac{w_i{}^h}{\sum_{h=1}^{z} \sum_{i=1}^{n} w_i{}^h} \tag{7}$$

where V_i^h refers to the degree of importance of the characteristic C_i^h, w_i^h is the valuation done for the characteristic C_i^h and $\sum_{h=1}^{z} \sum_{i=1}^{n} w_i{}^h$ is the sum of all the valuations done for all the characteristics of each policy.

From this convex weighting, we could obtain the Hamming distance as:

$$\pi(P_{1,2,\ldots,z}, P_k^{(p)}) = \sum_{h=1}^{z} \sum_{i=1}^{n} V_i{}^h \left| \mu_i{}^h - \mu_i^{(k)} \right| \tag{8}$$

with: $i = 1,2,\ldots,n$; $k = 1,2,\ldots,m$; and $\forall\ (P_{1,2,\ldots,z}, P_k^{(p)}) \in \mu_i{}^h, \mu_i^{(k)} \in [0, 1]$.

A third case could be the introduction of the OWA operator in the selection process in order to introduce the attitudinal character of the decision maker in the decision. The OWA operator (Yager, 1988) is an aggregation operator that provides a parameterized family of aggregation operators between the maximum and the minimum that has been applied in a lot of fields (Merigó, 2008; Yager and Kacprzyk, 1997). In this case, using the same methodology as A.M. Gil-Lafuente and Merigó (2006), the formulation is:

$$\beta(P_{1,2,\ldots,z}, P_k^{(p)}) = \sum_{j=1}^{zn} w_j D_j \tag{9}$$

where D_j represents the jth smallest of the $|\mu_i^h - \mu_i^{(k)}|$, because in distances, the best alternative is the one with the smallest distance to the ideal, and $k = 1,2,\ldots,m$.

As we can see, it has been introduced an ascending OWA (AOWA) operator in the Hamming distance because the reordering step is ascendant. And w_j represents a weighting vector W with $w_j \in [0, 1]$ and $\sum_{j=1}^{zn} w_j = 1$.

With this weighting vector, we can calculate the attitudinal character (Yager, 1993) as:

$$\alpha(W) = \sum_{j=1}^{zn} \left[\frac{zn - j}{zn - 1} \right] w_j \tag{10}$$

As we can see: $\alpha \in [0, 1]$. Values near 1, show that the selection process has been developed with a high level of optimism, while values near 0, show that the selection process has been developed with a low level of optimism or with a high level of pessimism.

This instrument can be very useful when the result obtained is not clear and the decision maker wants to reconsider the decision changing his degree of optimism. Note that by choosing a different manifestation of the weighting vector, we are able to obtain different types of aggregation operators (Merigó, 2008; Yager, 1993; Yager and Kacprzyk, 1997).

A fourth case could be the introduction of the GOWA operator in the selection process in order to introduce a generalized version of the OWA operator. The GOWA operator (Karayiannis, 2000; Yager, 2004) includes a wide range of OWA operators such as the geometric OWA, the quadratic OWA and the harmonic OWA operator. Following Merigó (2008) we could formulate this situation as:

$$\beta(P_{1,2,\ldots,z}, P_k^{(p)}) = \left(\sum_{j=1}^{zn} w_j D_j^{\lambda} \right)^{1/\lambda} \tag{11}$$

where D_j represents the jth smallest of the $|\mu_i^h - \mu_i^{(k)}|$, because in distances, the best alternative is the one with the smallest distance to the ideal, and $k = 1,2,\ldots,m$, and λ is a parameter such that $\lambda \in (-\infty, \infty)$.

As it can be seen, it has been introduced an ascending GOWA (AGOWA) operator (Merigó, 2008; Merigó and A.M. Gil-Lafuente, 2008b) in the Hamming distance because the reordering step is ascendant. It is important to note that we will not include in the aggregation the $D_j = 0$; $\forall_j$. Here, w_j represents a weighting vector with $w_j \in [0, 1]$ and $\sum_{j=1}^{zn} w_j = 1$.

By choosing a different manifestation of the weighting vector, we are able to obtain different types of aggregation operators (Merigó, 2008; Yager, 1993).

3.3 The Adequacy Coefficient in Political Management

Another index that could be used in the selection process is the adequacy coefficient. Here, we also consider four cases where the adequacy coefficient could be used. First, we could consider the case where the characteristics have the same level of importance. In this case, the adequacy coefficient is as follows:

$$K\left(P_k^{(p)} \rightarrow P_{1,2,\ldots,z}\right) = \frac{1}{zn} \sum_{h=1}^{z} \sum_{i=1}^{n} \left[1 \wedge \left(1 - \mu_i^h + \mu_i^{(k)}\right)\right] \tag{12}$$

with: $k = 1,2,\ldots,m$.

Concerning the results obtained, we should note that they refer to the approximation of the policies to the ideal. Thus, our preference relation will be constructed in a decreasing order being the highest value the best result.

Analogously to this index, we could calculate its equivalent removal index. Its formulation for the case of policies is:

$$Q\left(P_k^{(p)} \rightarrow P_{1,2,\ldots,z}\right) = \frac{1}{zn} \sum_{h=1}^{z} \sum_{i=1}^{n} \left[0 \vee \left(\mu_i^h - \mu_i^{(k)}\right)\right] \tag{13}$$

In this case, the preference relation will be ascendant. That is, we will select the lowest result. Obviously, we see that the adequacy coefficient and the removal index are inversely related:

$$K\left(P_k^{(p)} \rightarrow P_{1,2,\ldots,z}\right) = 1 - Q\left(P_k^{(p)} \rightarrow P_{1,2,\ldots,z}\right) \tag{14}$$

A second case that we could consider is the case where the characteristics have different degrees of importance. In order to use different degrees of importance, we could use a convex weighting so the result is still in [0, 1].

For this case, we will use the same convex weighting as in Eq. (2). Thus, the adequacy coefficient is:

$$K\left(P_k^{(p)} \rightarrow P_{1,2,\ldots,z}\right) = \sum_{h=1}^{z} \sum_{i=1}^{n} V_i^h \left[1 \wedge \left(1 - \mu_i^h + \mu_i^{(k)}\right)\right] \tag{15}$$

And for the removal index:

$$Q\left(P_k^{(p)} \rightarrow P_{1,2,\ldots,z}\right) = \sum_{h=1}^{z} \sum_{i=1}^{n} V_i^h \left[0 \vee \left(\mu_i^h - \mu_i^{(k)}\right)\right] \tag{16}$$

A third case that could be used, would be the introduction of the OWA operators in the adequacy coefficient. Using the same methodology as in Merigó and A.M. Gil-Lafuente (2008c; 2008d; 2010), the formulation is:

$$K(P_k^{(p)} \rightarrow P_{1,2,\ldots,z}) = \sum_{j=1}^{zn} w_j K_j \tag{17}$$

where K_j represents the jth largest of the $[1 \wedge (1 - \mu_i^h + \mu_i^{(k)})]$, and $k = 1,2,\ldots,m$.

In this case, the reordering step is done in a decreasing way as the best result is the largest number. Therefore, the type of OWA operator used is the descending OWA (DOWA) operator: $K_1 \geq K_2 \geq \ldots \geq K_{zn}$. The final result will be a number between [0, 1], being the maximum possible result 1. And w_j represents a weighting vector W, with $w_j \in [0, 1]$ and $\sum_{j=1}^{zn} w_j = 1$. Note that in this case, we could also calculate the attitudinal character $\alpha(W)$ and different particular cases (Merigó, 2008; Yager, 1993). Note also that it is possible to calculate the removal index in a similar way as in Eq. (13) and (14).

Finally, a fourth case that could be used with the adequacy coefficient would be the introduction of the GOWA operator in the index in order to use the attitudinal character of the decision maker in the selection process. Here, the formulation would be:

$$K(P_k^{(p)} \rightarrow P_{1,2,\ldots,z}) = \left(\sum_{j=1}^{zn} w_j K_j^{\lambda} \right)^{1/\lambda} \tag{18}$$

where K_j represents the jth largest of the $[1 \wedge (1 - \mu_i^h + \mu_i^{(k)})]$, and $k = 1,2,\ldots,m$.

In this case, the reordering step is developed in a decreasing order as the best result is the largest number. Thus, the type of GOWA operator used is the descending GOWA (DGOWA) operator: $K_1 \geq K_2 \geq \ldots \geq K_{zn}$. The final result will be a number between [0 ,1], being the maximum possible result 1. And w_j represents a weighting vector with $w_j \in [0,1]$ and $\sum_{j=1}^{zn} w_j = 1$.

3.4 The Index of Maximum and Minimum Level in Political Management

The third index that we consider in the selection of polyvalent policies is the index of maximum and minimum level (J. Gil-Lafuente, 2001; 2002). A first case that we could consider is the case where all the characteristics have the same degree of importance. The formulation is as follows:

$$\eta(P_k^{(p)} \rightarrow P_{1,2,\ldots,z}) = \\ = \frac{1}{u+v} \left[\sum_u \left| \mu_i^h(u) - \mu_i^{(k)}(u) \right| + \sum_v \left[0 \vee (\mu_i^h(v) - \mu_i^{(k)}(v)) \right] \right] \tag{19}$$

where u refers to the characteristics to be considered with the Hamming distance and v refers to the characteristics to be considered with the adequacy coefficient.

We should note that $u + v = zn$. That is, the sum of both groups of characteristics is equal to the total number of characteristics.

As in the previous indexes, we could consider the case where the characteristics have different levels of importance. To solve this problem, we should introduce a version of Eq. (2) in Eq. (19). Thus, the formulation is:

$$\eta(P_k^{(p)} \to P_{1,2,\ldots,z}) = \\ = \sum_u Z_i^h(u)\left|\mu_i^h(u) - \mu_i^{(k)}(u)\right| + \sum_v Z_i^h(v)\left(0 \vee (\mu_i^h(v) - \mu_i^{(k)}(v))\right) \quad (20)$$

with $Z_i^h = w_i^h / \sum_{h=1}^{z}\sum_{i=1}^{n} w_i^h$; which represents the level of importance of the characteristic C_i^h.

Analogously to this removal index, we could calculate the approximation index as:

$$\upsilon(P_k^{(p)} \to P_{1,2,\ldots,z}) = 1 - \eta(P_k^{(p)} \to P_{1,2,\ldots,z}) \quad (21)$$

A third case that we could consider is the introduction of the OWA operator in the index of maximum and minimum level in order to modify the neutrality of the index. Using the same methodology as Merigó (2008) and Merigó and A.M. Gil-Lafuente (2006), the formulation is:

$$S(P_k^{(p)} \to P_{1,2,\ldots,z}) = \sum_{j=1}^{zn} w_j S_j \quad (22)$$

where S_j represents the jth smallest of all the $|\mu_i^h - \mu_i^{(k)}|$ and the $[0 \vee (\mu_i^h - \mu_i^{(k)})]$; with $k = 1,2,\ldots,m$.

In this case, an AOWA operator is used in the reordering step ($S_1 \leq S_2 \leq \ldots \leq S_{zn}$) with the particularity that it always selects the jth smallest of all the possible values, independently if it is a result coming from the Hamming distance or from the removal index of the adequacy coefficient. Here, w_j represents a weighting vector W, with $w_j \in [0, 1]$ and $\sum_{j=1}^{zn} w_j = 1$, and we could also calculate the attitudinal character $\alpha(W)$ and different particular cases (Merigó, 2008; Yager, 1988; 1993).

Finally, the last case we will consider in this paper is the introduction of the GOWA operator in the index of maximum and minimum level in order to change the neutrality of the selection process. Here, the formulation will be:

$$S(P_k^{(p)} \to P_{1,2,\ldots,z}) = \left(\sum_{j=1}^{zn} w_j S_j^{\lambda}\right)^{1/\lambda} \quad (23)$$

where S_j represents the jth smallest of all the $|\mu_i^h - \mu_i^{(k)}|$ and the $[0 \vee (\mu_i^h - \mu_i^{(k)})]$; with $k = 1,2,\ldots,m$.

In this case, an AGOWA operator is used in the reordering step ($S_1 \leq S_2 \leq \ldots \leq S_{zn}$) with the particularity that it always selects the jth smallest of all the possible values, independently if it is a result coming from the Hamming distance or from the removal index of the adequacy coefficient. It is important to note that we will not include in the aggregation the $S_j = 0$; $\forall_j$. For this case, w_j represents a weighting vector, with $w_j \in [0, 1]$ and $\sum_{j=1}^{zn} w_j = 1$.

4 Numerical Example

In this Section, we present a simple numerical example where it is possible to see the applicability of the new approach in a decision making problem regarding the selection of policies. We focus on the selection of monetary policies where a government is looking for their optimal policy the next year. We will use the similarity measures commented above. That is, the Hamming distance, the adequacy coefficient and the index of maximum and minimum level. We also use several aggregation operators in the similarity measures such as the simple average, the weighted average, the OWA operator and various particular cases.

Step 1: Assume a government is looking for its general strategy the next year concerning monetary political management. In order to analyze the information, the government uses a group of experts to assess the information. After careful evaluation of the information, the group of experts of the government considers the following monetary policies that could be developed according to the necessities of the country.

- A_1 = Develop a strong contractive monetary policy.
- A_2 = Develop a contractive monetary policy.
- A_3 = No not develop any change in the monetary policy.
- A_4 = Develop an expansive monetary policy.
- A_5 = Develop a strong expansive monetary policy.

The economic evaluation of carrying out the previous monetary policies can be described considering the following characteristics C.

- C_1 = Benefits.
- C_2 = Risk of the strategy.
- C_3 = Other variables.

Step 2: With this information, the group of experts of the government establishes the ideal results that the ideal policy should have. They consider three different ideals depending on the economic situation the next period. In summary, they consider that the economic situation for the next year can be bad, regular or good. These results are represented in Table 1. Note that we give valuations to each characteristic depending on the economic situation found the next period. Thus, we have C_{ih} where i is the ith characteristic and h is the hth scenario.

Table 1 Ideal monetary policy

	C_{11}	C_{21}	C_{31}	C_{12}	C_{22}	C_{32}	C_{13}	C_{23}	C_{33}
$P =$	0.8	0.9	0.8	0.9	0.8	0.8	1	0.7	0.9

Step 3*:* Fixation of the real level of each characteristic for all the different policies considered. For each of these characteristics depending on the scenario found the next period, we get the following information presented in Table 2:

Table 2 Available monetary policies

	C_{11}	C_{21}	C_{31}	C_{12}	C_{22}	C_{32}	C_{13}	C_{23}	C_{33}
A_1	0.6	0.8	0.4	0.7	0.7	0.7	0.6	0.7	0.6
A_2	0.7	0.7	0.4	0.6	0.8	0.6	0.7	0.7	0.5
A_3	0.8	0.6	0.5	0.5	0.8	0.7	0.7	0.8	0.4
A_4	0.6	0.7	0.6	0.6	0.8	0.7	0.8	0.6	0.5
A_5	0.8	0.4	0.6	0.9	0.7	0.6	0.9	0.6	0.4

Step 4: Comparison between the ideal policy and the different monetary policies considered, and determination of the level of removal using the different types of aggregation operators considered. By using the Hamming distance, we consider the normalized Hamming distance (NHD), the weighted Hamming distance (WHD), the OWA distance (OWAD), the ascending OWA distance (AOWAD) and the median-OWAD operator. In this example, we assume that the company decides to use the following weighting vector: $W = (0.1, 0.1, 0.1\ 0.1, 0.1, 0.1, 0.1, 0.1, 0.2)$. Note that we have to calculate the individual distances of each characteristic to the ideal value of the corresponding characteristic forming the fuzzy subset of individual distances for each policy. Once, we have the individual distances, we aggregate them with the appropriate aggregation operator. The results are shown in Table 3.

Table 3 Aggregated results with the Hamming distance

	NHD	*WHD*	*OWAD*	*AOWAD*	*Median*
A_1	0.2	0.21	0.18	0.22	0.2
A_2	0.211	0.23	0.19	0.23	0.2
A_3	0.222	0.25	0.2	0.25	0.3
A_4	0.188	0.21	0.17	0.21	0.2
A_5	0.188	0.22	0.17	0.22	0.1

If we develop the selection process with the adequacy coefficient, we get the following. First, we have to calculate how close the characteristics are to the ideal policy. Once we have calculated all the different individual values, we construct the aggregation. In this case, we consider the normalized adequacy coefficient (NAC), the weighted adequacy coefficient (WAC), the OWA adequacy coefficient (OWAAC), the ascending OWA adequacy coefficient (AOWAAC) and the median-OWAAC operator. The results are shown in Table 4.

Table 4 Aggregated results with the adequacy coefficient

	NAC	*WAC*	*OWAAC*	*AOWAAC*	*Median*
A_1	0.8	0.79	0.78	0.82	0.8
A_2	0.788	0.77	0.77	0.81	0.8
A_3	0.788	0.76	0.76	0.81	0.7
A_4	0.811	0.79	0.79	0.83	0.8
A_5	0.811	0.78	0.78	0.83	0.9

Finally, if we use the index of maximum and minimum level in the selection process as a combination of the normalized Hamming distance and the normalized adequacy coefficient, we will get the following. In this example, we assume that the characteristics C_1 and C_2 have to be treated with the adequacy coefficient and the third one with the Hamming distance. In this case, we do the following. First, we calculate the individual removal of each characteristic to the ideal, independently that the instrument used is the Hamming distance or the adequacy index. Once calculated all the values for the individual removal, we will construct the aggregation using the framework explained in section 3.4. Here, we note that in the reordering step, it will be only considered the individual value obtained for each characteristic, independently that the value has been obtained with the adequacy coefficient or with the Hamming distance. We consider the normalized index of maximum and minimum level (NIMAM), the weighted index of maximum and minimum level (WIMAM), the OWA index of maximum and minimum level (OWAIMAM), the ascending OWAIMAM (AOWAIMAM) and the median-OWAIMAM operator. The results are shown in Table 5.

Table 5 Aggregated results with the index of maximum and minimum level

	NIMAM	*WIMAM*	*OWAIMAM*	*AOWAIMAM*	*Median*
A_1	0.2	0.21	0.18	0.22	0.2
A_2	0.211	0.23	0.19	0.23	0.2
A_3	0.222	0.25	0.2	0.25	0.3
A_4	0.188	0.21	0.17	0.21	0.2
A_5	0.188	0.22	0.17	0.22	0.1

As we can see, the results obtained with the index of maximum and minimum level are the same than the ones obtained with the Hamming distance. The reason for this is that we are in a situation of unification point between both similarity measures (Merigó and A.M. Gil-Lafuente, 2007b).

In order to analyze the optimal monetary policies depending on the particular types of similarity measure used, we establish the following table with the ordering of the monetary policies. Note that this analysis is very useful when the decision maker wants to consider more than one alternative in the selection process. The results are shown in Table 6.

Table 6 Ordering of the monetary policies

	Ordering		Ordering
NHD	$A_4=A_5 \succ A_1 \succ A_2 \succ A_3$	AOWAAC	$A_4=A_5 \succ A_1 \succ A_2=A_3$
WHD	$A_1=A_4 \succ A_5 \succ A_2 \succ A_3$	Median-OWAAC	$A_5 \succ A_1=A_2=A_4 \succ A_3$
OWAD	$A_4=A_5 \succ A_1 \succ A_2 \succ A_3$	NIMAM	$A_4=A_5 \succ A_1 \succ A_2 \succ A_3$
AOWAD	$A_4 \succ A_1=A_5 \succ A_2 \succ A_3$	WIMAM	$A_1=A_4 \succ A_5 \succ A_2 \succ A_3$
Median-OWAD	$A_5 \succ A_1 \succ A_2 \succ A_4 \succ A_3$	OWAIMAM	$A_4=A_5 \succ A_1 \succ A_2 \succ A_3$
NAC	$A_4=A_5 \succ A_1 \succ A_2=A_3$	AOWAIMAM	$A_4 \succ A_1=A_5 \succ A_2 \succ A_3$
WAC	$A_1=A_4 \succ A_5 \succ A_2 \succ A_3$	Median	$A_5 \succ A_1 \succ A_2 \succ A_4 \succ A_3$
OWAAC	$A_4 \succ A_1=A_5 \succ A_2 \succ A_3$		

As we can see, we get different orderings depending on the aggregation operator used. The main advantage of this analysis is that the decision maker gets a more complete view of the different scenarios that could happen in the future depending on the method used. Although he will select the alternative that it is in accordance with his interests, he will be concerned on other potential results that could happen in the uncertain environment.

6 Conclusions

We have presented a new approach for decision making in an uncertain environment where we need to consider different attributes and scenarios. We have studied this new approach in a political management problem. Particularly, we have focussed on the selection of monetary policies. Thus, we have studied the selection of polyvalent monetary policies. We have seen that this new approach provides a more complete and flexible representation of the decision process. We have used several types of similarity measures for developing the analysis. We have considered the use of the Hamming distance, the adequacy coefficient and the index of maximum and minimum level. In each method we have used a wide range of aggregation operators such as the simple average, the weighted average, the OWA operator and the GOWA operator. We have seen that the use of a wide range of aggregation operators provides a more robust formulation of the new approach.

We have presented a numerical example where we have seen the implementation of the new approach in a real world problem. We have used a wide range of aggregation operators based on the Hamming distance, the adequacy coefficient and the index of maximum and minimum level. Each aggregation operator provides a different result leading to a different decision.

In future research we expect to develop further developments by considering other aggregation operators such as those recently presented by Merigó (2008) and considering other applications such as strategic management or production management.

Acknowledgements. This work is partly supported by the Spanish ministry Spanish Ministerio de Asuntos Exteriores y de Cooperación, Agencia Española de Cooperación Internacional para el Desarrollo (AECID) (project A/016239/08).

References

Beliakov, G., Calvo, T., Pradera, A.: Aggregation Functions: A Guide for Practitioners. Springer, Berlin (2007)

Figueira, J., Greco, S., Ehrgott, M.: Multiple criteria decision analysis: State of the art surveys. Springer, Boston (2005)

Fodor, J., Marichal, J.L., Roubens, M.: Characterization of the ordered weighted averaging operators. IEEE Transactions on Fuzzy Systems 3, 236–240 (1995)

Gil-Aluja, J.: The interactive management of human resources in uncertainty. Kluwer Academic Publishers, Dordrecht (1998)

Gil-Aluja, J.: Elements for a theory of decision under uncertainty. Kluwer Academic Publishers, Dordrecht (1999)

Gil-Lafuente, A.M.: Fuzzy logic in financial analysis. Springer, Berlin (2005)

Gil-Lafuente, A.M., Merigó, J.M.: Acquisition of financial products that adapt to different environments. Lectures on Modelling and Simulation 7, 42–48 (2006)

Gil-Lafuente, J.: The index of maximum and minimum level in the selection of players in sport management (in Spanish). In: Proceedings of the 10th International Conference of the European Academy of Management and Business Economics (AEDEM), Reggio Calabria, Italy, pp. 439–443 (2001)

Gil-Lafuente, J.: Algorithms for excelence. In: Milladoiro, V. (ed.) Keys for being successful in sport management (2002) (in Spanish)

Hamming, R.W.: Error-detecting and error-correcting codes. Bell Systems Technical Journal 29, 147–160 (1950)

Karayiannis, N.: Soft Learning Vector Quantization and Clustering Algorithms Based on Ordered Weighted Aggregation Operators. IEEE Transactions on Neural Networks 11, 1093–1105 (2000)

Kaufmann, A.: Introduction to the theory of fuzzy subsets. Academic Press, New York (1975)

Kaufmann, A., Gil-Aluja, J.: Introduction to the theory of fuzzy subsets in business management, Ed. Milladoiro, Santiago de Compostela (1986) (in Spanish)

Kaufmann, A., Gil-Aluja, J.: Técnicas operativas de gestión para el tratamiento de la incertidumbre, Ed. Hispano-europea, Barcelona (1987)

Merigó, J.M.: New extensions to the OWA operator and its application in decision making. PhD Thesis, Department of Business Administration, University of Barcelona (2008) (in Spanish)

Merigó, J.M., Gil-Lafuente, A.M.: Using the OWA operators in the selection of financial products. In: Proceedings of the 41st CLADEA Conference, Montpellier, France (2006), CD-ROM Proceedings

Merigó, J.M., Gil-Lafuente, A.M.: The ordered weighted averaging distance operator. Lectures on Modelling and Simulation 8, 1–11 (2007a)

Merigó, J.M., Gil-Lafuente, A.M.: Unification point in methods for the selection of financial products. Fuzzy Economic Review 12, 35–50 (2007b)

Merigó, J.M., Gil-Lafuente, A.M.: On the use of the OWA operator in the Euclidean distance. International Journal of Computer Science and Engineering 2, 170–176 (2008a)

Merigó, J.M., Gil-Lafuente, A.M.: Using the OWA operator in the Minkowski distance. International Journal of Computer Science 3, 149–157 (2008b)

Merigó, J.M., Gil-Lafuente, A.M.: On the use of the OWA operator in the adequacy coefficient. Modelling, Measurement and Control D 29, 1–15 (2008c)

Merigó, J.M., Gil-Lafuente, A.M.: The generalized adequacy coefficient and its application in strategic decision making. Fuzzy Economic Review 13, 17–36 (2008d)

Merigó, J.M., Gil-Lafuente, A.M.: The induced generalized OWA operator. Information Sciences 179, 729–741 (2009a)

Merigó, J.M., Gil-Lafuente, A.M.: OWA operators in generalized distances. International Journal of Applied Mathematics and Computer Science 5, 11–18 (2009b)

Merigó, J.M., Gil-Lafuente, A.M.: New decision making techniques and their application in the selection of financial products. Information Sciences (to be published, 2010)

Szmidt, E., Kacprzyk, J.: Distances between intuitionistic fuzzy sets. Fuzzy Sets and Systems 114, 505–518 (2000)

Xu, Z.S.: An overview of methods for determining OWA weights. International Journal of Intelligent Systems 20, 843–865 (2005)

Yager, R.R.: On Ordered Weighted Averaging Aggregation Operators in Multi-Criteria Decision Making. IEEE Transactions on Systems, Man and Cybernetics, B 18, 183–190 (1988)

Yager, R.R.: Families of OWA operators. Fuzzy Sets and Systems 59, 125–148 (1993)

Yager, R.R.: Generalized OWA Aggregation Operators. Fuzzy Optimization and Decision Making 3, 93–107 (2004)

Yager, R.R., Kacprzyk, J.: The Ordered Weighted Averaging Operators: Theory and Applications. Kluwer Academic Publishers, Norwell (1997)

Mathematical Approaches for Fuzzy Portfolio Selection Problems with Normal Mixture Distributions

Takashi Hasuike and Hiroaki Ishii

Abstract. This chapter considers some versatile portfolio selection models with general normal mixture distributions and fuzzy or interval numbers. Then, these mathematical approaches to obtain the optimal portfolio are developed. Furthermore, in order to compare our proposed models with standard models and represent the advantage of our proposed models, a numerical example is provided.

1 Introduction

In recent rapid expansions of investment and financial instability such as the extreme ups and downs to future markets of stocks and commodities, the role of investment theory becomes more and more important. Furthermore, in current investment environment, with information science and computers, not only big companies and institutional investors but also individual investors perform investment in stock, currency, land and property. Therefore, it is time to review investment theory, particularly portfolio theory. Practical financial markets are affected by a lot of uncertainty to which has a great influence on the future returns such as randomness derived from statistical analysis of historical data and ambiguity such as the psychological aspect of investors and lack of received efficient information. Under such uncertain conditions, the investor needs to consider how to reduce risk, and receive the greatest future profit. Such a finance assets selection problem is generally called a portfolio selection problem, and various studies have been done. Markowitz [20] first proposed Mean-Variance model in the sense of the mathematical programming. Then, it has been central to research

Takashi Hasuike · Hiroaki Ishii
Graduate School of Information Science and Technology, Osaka University, Japan
2-1 Yamadaoka, Suita, Osaka 565-0871, Japan
e-mail: thasuike@ist.osaka-u.ac.jp

W.A. Lodwick & J. Kacprzyk (Eds.): Fuzzy Optimization, STUDFUZZ 254, pp. 407–423.
springerlink.com

activity in the real financial field and numerous researchers have contributed to the development of the modern portfolio theory (for instance, Luenberger [19], Steinbach [23]). On the other hand, many researchers have proposed models of portfolio selection problems which extended Markowitz model; Mean-Absolute-Deviation model (Konno [16], Konno, et al. [17]), Semi-Variance model (Bawa and Lindenberg [1]), Safety-First model (Elton and Gruber [5]), and Value at Risk and conditional Value at Risk model (Rockafellar and Uryasev [22]).

In many previous studies in mathematical programming for investment, future returns are assumed to be continuous random variables according to normal distributions. By this assumption, they obtained useful mathematical knowledge and formulas for the portfolio theory. However, from recent experimental studies of investment markets, it is often shown that future returns do not occur according to normal distribution, but fat or heavy-tail distribution source. Then, in the case that investors predict future returns, since they need to consider much information derived from investment markets and some predictions of future returns based on subjectivity of veteran investors simultaneously, they usually assume not only one scenario but also several possibility scenarios of future returns. Therefore, in order to deal with these situations, we need to consider portfolio selection problems with more general random distribution with the heavy-tail.

Furthermore, we need to consider flexibly and ambiguously defined statistical distributions for the following cases that occur in practice: (1) financial information is incomplete, (2) expert investors are not estimating from historical data, (3) the need to mathematically deal with several marginal distributions considering higher or lower future returns simultaneously. In this paper, we propose a more extensional portfolio selection models including fuzzy factors. Until now, there is a body of research under various uncertainty conditions with respect to portfolio selection problems (Bilbao-Terol and Perez-Gladish [2], Carlsson et al. [3], Guo and Tanaka [6], Huang [10, 11], Inuiguchi et al. [12, 13], Katagiri et al. [14, 15], Tanaka et al. [24, 25], Watada [26]). We also proposed some portfolio models with both randomness and fuzziness [7, 8, 9]. However, there are few models considering both normal mixture distribution and ambiguity, simultaneously. Furthermore, there are no studies which are analytically extended and solved these types of portfolio selection problems. In this chapter, we propose more extensional portfolio selection models including the general random distribution with fuzzy factors and develop the efficient solution method.

On the other hand, in the sense of mathematical programming, these portfolio selection models with randomness and fuzziness are formulated as stochastic and fuzzy programming problems. Then, in order to solve them analytically, we need to use the stochastic and fuzzy optimization approaches. The stochastic optimization approach has been treated as a basic solution tool for portfolio selection problems. Recently, as well as the stochastic optimization approach, the fuzzy

optimization approach has been used as one of useful tools in financial and investment fields because these approaches are dealt with the investor's subjectivity and the investment case under ineffective and linguistic received information. Considering recent complex investment markets involving investor's speculation and the mixture of reliable on unreliable information, it is obvious that fuzzy optimization approaches play an important role in the investment research field. Various types of fuzzy optimization models have been proposed; fuzzy max ranking method, possibility and necessity programming, etc.. However, there are not many studies to compare these fuzzy optimization models for portfolio selection problems under each case such as favorable, poor, or erratic ups and downs economic conditions, investor's subjectivities; optimistic, pessimistic or neutral. Furthermore, there are also few studies that analyze the suitability of various approaches for the differing market conditions (for example, favorable conditions, unfavorable conditions, erratic fluctuations) by comparing fuzzy, stochastic and fuzzy-stochastic optimization models for portfolio selection problems.

Thus, in order to represent some uncertain social conditions and compare various models for portfolio selection problems, it is important to construct the analytical solution method using the fuzzy and stochastic optimization method as well as to develop a versatile model for portfolio selection problems. Therefore, in this chapter, we consider some models based on the fuzzy optimization models for our proposed portfolio models including the general random distribution with fuzzy factors.

This paper is organized as follows. In Section 2, we introduce notations of parameters in this paper and introduce the basic Mean-Variance model proposed by Markowitz. Then, in Section 3, we formulate the proposed portfolio selection problems minimizing the total variance and maximizing the total future return with normal mixture distributions, respectively. Furthermore, we introduce the uncertainty sets for mean values, weights and probabilities as fuzzy numbers. With respect to several portfolio selection problems including randomness and fuzziness, we construct the solution method. In Section 4, in order to compare our proposed models with standard models, we provide a numerical example. Finally, in Section 5, we conclude this paper and discuss future research problems.

2 Basic Mean-Variance Model

Notations of parameters in this paper are as follows:

$\boldsymbol{r}_i$: random column vector ith scenario

$\boldsymbol{m}_i$: mean value column vector of random variable $\boldsymbol{r}_i$

$\mathbf{V}_i$: variance-covariance matrix of random variable $\boldsymbol{r}_i$

r_G : target value of the total future return

σ_G : target value of the total variance

b_j : Upper limited value of purchasing rate

$\boldsymbol{x}$: Purchasing volume (Decision variable)

In this chapter, we mainly deal with the standard Markowitz model for portfolio selection problem involving normal mixture distributions with respect to future returns. First, we introduce the following Markowitz model minimizing the total variance:

$$\begin{aligned} &\text{Minimize} \quad \boldsymbol{x}^t \mathbf{V} \boldsymbol{x} \\ &\text{subject to} \quad \boldsymbol{m}^t \boldsymbol{x} \geq r_G \\ &\qquad \boldsymbol{x} \in X \triangleq \left\{ \boldsymbol{x} \left| \sum_{j=1}^{n} x_j = 1,\ 0 \leq x_j \leq b_j,\ \left(j = 1, 2, ..., n\right) \right. \right\} \end{aligned} \tag{1}$$

This model has been the centre of investment fields until now, and so there are many studies by academic and practical researchers.

In the case that we obtain the strict value of parameters $\boldsymbol{m}$ and $\mathbf{V}$, problem (1) is equivalent to a quadratic programming problem in the mathematical programming and we find an optimal portfolio using standard convex programming approaches. Furthermore, while problem (1) considers minimizing the total variance, the case maximizing the total future return is formulated as the following form:

$$\begin{aligned} &\text{Maximize} \quad \boldsymbol{m}^t \boldsymbol{x} \\ &\text{subject to} \quad \boldsymbol{x}^t \mathbf{V} \boldsymbol{x} \leq \sigma_G,\ \boldsymbol{x} \in X \end{aligned} \tag{2}$$

This problem is also a quadratic programming problem and so we obtain an optimal portfolio using the basic convex programming approach.

3 Our Proposed Model with Normal Mixture Distributions

However, since it often happens that future returns occur according to the normal distribution in practical investment fields, we need to consider that future returns occur according to heavier tailed distributions than normal. Therefore, in this paper, we consider that each random variable r_{ij} occurs according to the following normal mixture distribution with a heavier tail than general normal distributions based on the study [21]:

$$\boldsymbol{r} \sim \sum_{i=1}^{m} w_i N\left(\boldsymbol{m}_i, \mathbf{V}_i\right) \tag{3}$$

where each parameter w_i is the non-negative scalar random variable. If parameter vector $\mathbf{W}$ of scalar random variable w_i is fixed, random variable vector $\boldsymbol{r}$ occurs according to the following normal distribution:

$$\begin{aligned} &\boldsymbol{r}\left|\mathbf{W}=\overline{\boldsymbol{w}} \sim N\left(\sum_{i=1}^{m} \overline{w}_i \overline{\boldsymbol{r}}_i, \sum_{i=1}^{m} \overline{w}_i \mathbf{V}_i\right)\right. \\ &\overline{\boldsymbol{w}}=\left(\overline{w}_1, \overline{w}_2, ..., \overline{w}_m\right), \overline{w}_i : \text{fixed value} \end{aligned} \tag{4}$$

where we assume that random variable vector $\mathbf{W}$ is independent on each matrix $\mathbf{V}_i$. In case (4), since random variables r_{ij} are basic normal distributions, we analytically obtain an optimal portfolio of problems (1) and (2). However, w_i is also a random variable, and so random variables r_{ij} are not normal random distributions. In this paper, to simplify the discussion, we assume that $\mathbf{W}$ occurs according to the following discrete distribution introducing probabilities p_s:

$$\boldsymbol{w}=\begin{cases} \boldsymbol{w}^{(1)} & \Pr\left\{\boldsymbol{w}=\boldsymbol{w}^{(1)}\right\}=p_1 \\ \boldsymbol{w}^{(2)} & \Pr\left\{\boldsymbol{w}=\boldsymbol{w}^{(2)}\right\}=p_2 \\ \vdots & \vdots \\ \boldsymbol{w}^{(S)} & \Pr\left\{\boldsymbol{w}=\boldsymbol{w}^{(S)}\right\}=p_S \end{cases}, \ \sum_{s=1}^{S} p_s=1, \ \sum_{i=1}^{m} w_i^{(s)}=1 \tag{5}$$

where S is the total number of discrete random variables and each $w_i^{(s)}$ is the fixed weight. Using this discrete distribution, the expected value and covariance of random variable column vector to future return $\boldsymbol{r}$ are as follows:

$$\begin{aligned} &E\left(\boldsymbol{r}\right)=E\left(E\left(\boldsymbol{r}\left|W\right.\right)\right)=\boldsymbol{w}_E \boldsymbol{m}_i \\ &\operatorname{cov}\left(\boldsymbol{r}\right)=E\left(\operatorname{cov}\left(\boldsymbol{r}\left|W\right.\right)\right)+\operatorname{cov}\left(E\left(\boldsymbol{r}\left|W\right.\right)\right)=\boldsymbol{w}_E \mathbf{V}_i \\ &w_i^E=\sum_{s=1}^{S} p_s w_i^{(s)}, \ \boldsymbol{w}_E=\left(w_1^E, w_2^E, ..., w_m^E\right) \end{aligned} \tag{6}$$

Therefore, using this expression (6), the expected return and covariance of total profit $\boldsymbol{r}^t \boldsymbol{x}$ are as follows:

$$
\begin{aligned}
E\left(\boldsymbol{r}^t\boldsymbol{x}\right) &= E\left(E\left(\boldsymbol{r}^t\boldsymbol{x}\middle|W\right)\right) = \left(\boldsymbol{w}_E\boldsymbol{m}_i\right)^t\boldsymbol{x} \\
\mathrm{cov}\left(\boldsymbol{r}^t\boldsymbol{x}\right) &= E\left(\mathrm{cov}\left(\boldsymbol{r}^t\boldsymbol{x}\middle|W\right)\right) + \mathrm{cov}\left(E\left(\boldsymbol{r}^t\boldsymbol{x}\middle|W\right)\right) \\
&= \boldsymbol{x}^t\left(\boldsymbol{w}_E\mathbf{V}_i\right)\boldsymbol{x}
\end{aligned} \tag{7}
$$

3.1 Formulation of Our Proposed Model

Then, we equivalently transform main problems (1) and (2) into the following problems:

$$
\begin{aligned}
&\text{Minimize} \quad \boldsymbol{x}^t\left(\boldsymbol{w}_E\mathbf{V}_i\right)\boldsymbol{x} \\
&\text{subject to} \quad \left(\boldsymbol{w}_E\boldsymbol{m}_i\right)^t\boldsymbol{x} \geq r_G,\ \boldsymbol{x} \in X
\end{aligned} \tag{8}
$$

$$
\begin{aligned}
&\text{Maximize} \quad \left(\boldsymbol{w}_E\boldsymbol{m}_i\right)^t\boldsymbol{x} \\
&\text{subject to} \quad \boldsymbol{x}^t\left(\boldsymbol{w}_E\mathbf{V}_i\right)\boldsymbol{x} \leq \sigma_G,\ \boldsymbol{x} \in X
\end{aligned} \tag{9}
$$

With respect to problems (8) and (9), if random distribution of parameter $\mathbf{W}$ is obtained and each parameter is constant, we may solve these problems analytically using similar methods to problems (1) and (2). However, considering ambiguity conditions such as subjectivity of the decision maker and the lack of received reliable information, it is natural that each parameter such as the mean value and random variable W include fuzziness. Therefore, we consider the following cases where problems (8) and (9) include fuzziness.

3.2 Fuzzy Extension for Mean Values

First, we consider the case where each expected value $\boldsymbol{m}$ includes fuzziness and is assumed to be a fuzzy number. This case is considered that the decision maker is a veteran investor and performs the more aggressive or passive prediction than the statistical analysis derived from historical data.

In this subsection, since the random distribution of parameter $\mathbf{W}$ is obtained and discrete value $w_i^{(s)}$ and its probability p_s are constant, expected value vector $\boldsymbol{w}_i^E$ is also a constant. Then, the membership function of fuzzy numbers $\boldsymbol{m}$ is assumed to be a triangle fuzzy number and introduced by the following function:

$$\mu_{\tilde{r}_{ij}}(\omega)=\begin{cases}\dfrac{\omega-\left(m_{ij}-\gamma_{ij}\right)}{\gamma_{ij}} & \left(m_{ij}-\gamma_{ij}\le\omega\le m_{ij}\right)\\ \dfrac{\left(m_{ij}+\delta_{ij}\right)-\omega}{\delta_{ij}} & \left(m_{ij}<\omega\le m_{ij}+\delta_{ij}\right)\\ 0 & \left(otherwise\right)\end{cases} \tag{10}$$

where γ_{ij} and δ_{ij} are spreads of left and right side, respectively. In this paper, we assume that provided all fuzzy numbers are initially determined by the decision maker. Using these fuzzy numbers, the total future expected return $\tilde{R}_i=\tilde{\boldsymbol{m}}^t\boldsymbol{x}$ is also a fuzzy numbers characterized by the following membership function:

$$\mu_{\tilde{R}}(\omega)=\begin{cases}\dfrac{\omega-R_L}{\sum_{i=1}^{m}w_i^E\sum_{j=1}^{n}\gamma_{ij}x_j} & \left(R_L\le\omega\le\sum_{i=1}^{m}w_i^E\sum_{j=1}^{n}m_{ij}x_j\right)\\ \dfrac{R_U-\omega}{\sum_{i=1}^{m}w_i^E\sum_{j=1}^{n}\delta_{ij}x_j} & \left(\sum_{i=1}^{m}w_i^E\sum_{j=1}^{n}m_{ij}x_j<\omega\le R_U\right)\\ 0 & \left(\omega<R_L,\ R_U<\omega\right)\end{cases} \tag{11}$$

$$R_L=\sum_{i=1}^{m}w_i^E\sum_{j=1}^{n}m_{ij}x_j-\sum_{i=1}^{m}w_i^E\sum_{j=1}^{n}\gamma_{ij}x_j,\ R_U=\sum_{i=1}^{m}w_i^E\sum_{j=1}^{n}m_{ij}x_j+\sum_{i=1}^{m}w_i^E\sum_{j=1}^{n}\delta_{ij}x_j$$

Due to these fuzzy numbers, problems (8) and (9) are not well-defined problem in the sense of deterministic mathematical programming. Therefore, in order to solve the main problem analytically, we need to set some criterion for fuzzy variables. In this paper, we consider the case where the decision maker usually has a goal to earn the total profit more than the target value. Furthermore, taking account of the vagueness of human judgment and flexibility for the execution of a plan in many real decision cases, we give a fuzzy goal to the target future return as the fuzzy set characterized by a membership function. In this subsection, we consider the fuzzy goal of probability $\mu_{\tilde{G}}(\omega)$ which is represented by,

$$\mu_{\tilde{G}}(\omega)=\begin{cases}0 & \omega\le f_L\\ \dfrac{\omega-f_L}{f_U-f_L} & f_L\le\omega\le f_U\\ 1 & f_U\le\omega\end{cases} \tag{12}$$

Furthermore, using the concept of possibility measure and considering maximizing both membership functions, we introduce the degree of possibility as follows:

$$\prod_{\tilde{R}}\left(\tilde{G}\right)=\sup_{f}\min\left\{\mu_{\tilde{R}}(f),\ \mu_{\tilde{G}}(f)\right\} \tag{13}$$

Therefore, introducing this degree of possibility, we consider the following portfolio selection problems based on problems (8) and (9):

$$\begin{aligned}&\text{Minimize} && \boldsymbol{x}^t\left(\boldsymbol{w}_E\mathbf{V}_i\right)\boldsymbol{x}\\ &\text{subject to} && \prod_{\tilde{R}}\left(\tilde{G}\right)\geq h,\ \boldsymbol{x}\in X\end{aligned} \tag{14}$$

and

$$\begin{aligned}&\text{Maximize} && \prod_{\tilde{R}}\left(\tilde{G}\right)\\ &\text{subject to} && \boldsymbol{x}^t\left(\boldsymbol{w}_E\mathbf{V}_i\right)\boldsymbol{x}\leq\sigma_G,\ \boldsymbol{x}\in X\\ &\quad\text{Maximize} && h\\ \Leftrightarrow\ &\quad\text{subject to} && \prod_{\tilde{R}}\left(\tilde{G}\right)\geq h,\\ &&& \boldsymbol{x}^t\left(\boldsymbol{w}_E\mathbf{V}_i\right)\boldsymbol{x}\leq\sigma_G,\ \boldsymbol{x}\in X\end{aligned} \tag{15}$$

In this problem, each constraint $\prod_{\tilde{P}_i}\left(G_i\right)\geq h$ is transformed into the following inequality:

$$\begin{aligned}&\prod_{\tilde{R}}\left(\tilde{G}\right)\geq h\\ \Leftrightarrow\ &\sup_{f}\min\left\{\mu_{\tilde{R}}(f),\ \mu_{\tilde{G}}(f)\right\}\geq h\\ \Leftrightarrow\ &\exists f:\ \mu_{\tilde{R}}(f)\geq h,\ \mu_{\tilde{G}}(f)\geq h\\ \Leftrightarrow\ &\sum_{i=1}^{m}w_i^E\sum_{j=1}^{n}m_{ij}x_j+(1-h)\delta_{ij}x_j\geq(1-h)f_L+f_U\end{aligned} \tag{16}$$

Therefore, problems (14) and (15) are equivalently transformed into the following problems:

$$\begin{aligned}&\text{Minimize} && \boldsymbol{x}^t\left(\boldsymbol{w}_E\mathbf{V}_i\right)\boldsymbol{x}\\ &\text{subject to} && \sum_{i=1}^{m}w_i^E\sum_{j=1}^{n}m_{ij}x_j+(1-h)\delta_{ij}x_j\geq(1-h)f_L+hf_U,\\ &&& \boldsymbol{x}\in X\end{aligned} \tag{17}$$

and

$$\begin{aligned}&\text{Maximize} && \frac{\sum_{i=1}^{m}w_i^E\sum_{j=1}^{n}\left(m_{ij}x_j+\delta_{ij}x_j\right)-f_L}{\sum_{i=1}^{m}w_i^E\sum_{j=1}^{n}\delta_{ij}x_j+f_U-f_L}\\ &\text{subject to} && \boldsymbol{x}^t\left(\boldsymbol{w}_E\mathbf{V}_i\right)\boldsymbol{x}\leq\sigma_G,\ \boldsymbol{x}\in X\end{aligned} \tag{18}$$

Problem (17) is equivalent to problem (8), and so we analytically solve problem (17) using the same solution method to problem (8). Then, problem (18) is a standard fractional programming problem and the numerator and denominator of objective function are linear functions. Therefore, by performing the equivalent transformation using the fractional programming approach, problem (18) is a similar problem to problem (9), and so we also solve problem (18) analytically using the same solution method to problem (9).

3.3 Fuzzy Extension Model for Each Weight or Probability

In the previous subsection, we considered the case where each expected value is assumed to be a fuzzy number. However, in real-world decision cases, it is difficult to set not only the expected value but also the possible value of random variable w_i or the occurrence probability p_i strictly due to the ambiguity of decision maker's subjectivity. Therefore, in this subsection, we consider the case where random variable W also includes flexibility and is assumed to be a fuzzy number.

First, we assume that the possible value w_i includes the ambiguity and represents a fuzzy number. Then, the membership function of each value w_i is introduced by the following functions:

$$\boldsymbol{w} = \begin{cases} \tilde{\boldsymbol{w}}^{(1)} & \Pr\left\{\boldsymbol{w} = \tilde{\boldsymbol{w}}^{(1)}\right\} = p_1 \\ \tilde{\boldsymbol{w}}^{(2)} & \Pr\left\{\boldsymbol{w} = \tilde{\boldsymbol{w}}^{(2)}\right\} = p_2 \\ \vdots & \vdots \\ \tilde{\boldsymbol{w}}^{(S)} & \Pr\left\{\boldsymbol{w} = \tilde{\boldsymbol{w}}^{(S)}\right\} = p_S \end{cases}, \ \tilde{\boldsymbol{w}}^{(s)} = \left(\tilde{w}_1^{(s)}, \tilde{w}_2^{(s)}, ..., \tilde{w}_m^{(s)}\right)$$

$$\mu_{\tilde{w}_i^{(s)}}(\omega) = \begin{cases} \dfrac{w_i^{(s)} - \omega}{\alpha_{si}} & \left(w_i^{(s)} - \alpha_{si} \le \omega \le w_i^{(s)}\right) \\ \dfrac{\omega - w_i^{(s)}}{\beta_{si}} & \left(w_i^{(s)} < \omega \le w_i^{(s)} + \beta_{si}\right) \\ 0 & \left(\omega < w_i^{(s)} - \alpha_{si}\ , w_i^{(s)} + \beta_{si} < \omega\right) \end{cases} \tag{19}$$

Using these membership functions and the extension principle of fuzzy theory, expected value of $\tilde{w}_i^E$ is obtained as the following form:

$$\mu_{\tilde{w}_i^E}(\omega)=\begin{cases}\dfrac{\omega-E_L}{\sum_{s=1}^{S}p_s\alpha_{si}} & \left(E_L\le\omega\le\sum_{s=1}^{S}p_s w_i^{(s)}\right)\\ \dfrac{E_U-\omega}{\sum_{s=1}^{S}p_s\beta_{si}} & \left(\sum_{s=1}^{S}p_s w_i^{(s)}<\omega\le E_U\right)\\ 0 & \left(\omega<E_L\ ,E_U<\omega\right)\end{cases} \tag{20}$$

$$E_L\left(w_i^E\right)=\sum_{s=1}^{S}p_s w_i^{(s)}-\sum_{s=1}^{S}p_s\alpha_{si},$$

$$E_U\left(w_i^E\right)=\sum_{i=1}^{m}p_i\overline{w}_i+\sum_{s=1}^{S}p_s\beta_{si}$$

Therefore, we obtain the following membership function of the total variance:

$$\mu_{\tilde{V}}(\omega)=\begin{cases}\dfrac{\omega-V_L}{\boldsymbol{x}^t\left(\sum_{i=1}^{m}\left(\sum_{s=1}^{S}p_s\alpha_{si}\right)\mathbf{V}_i\right)\boldsymbol{x}}\\ \dfrac{V_U-\omega}{\boldsymbol{x}^t\left(\sum_{i=1}^{m}\left(\sum_{s=1}^{S}p_s\beta_{si}\right)\mathbf{V}_i\right)\boldsymbol{x}}\\ 0\end{cases} \tag{21}$$

$$V_L=\boldsymbol{x}^t\left(\sum_{i=1}^{m}\left(\sum_{s=1}^{S}p_s w_i^{(s)}-\sum_{s=1}^{S}p_s\alpha_{si}\right)\mathbf{V}_i\right)\boldsymbol{x},\ V_U=\boldsymbol{x}^t\left(\sum_{i=1}^{m}\left(\sum_{s=1}^{S}p_s w_i^{(s)}-\sum_{s=1}^{S}p_s\beta_{si}\right)\mathbf{V}_i\right)\boldsymbol{x}$$

In this case, we assume that each weight $w_i^{(s)}$ is represented as the h-cut set of the fuzzy number $\tilde{w}_i^{(s)}=\left[\underline{w}_i^{(s)}(h),\overline{w}_i^{(s)}(h)\right]$. Furthermore, in a way similar to subsection 3.2, using the concept of possibility measure to the total variance, we introduce the degree of possibility as follows:

$$\prod_{\tilde{V}}\left(\tilde{G}\right)=\sup_{\sigma}\min\left\{\mu_{\tilde{V}}(\sigma),\ \mu_{\tilde{G}}(\sigma)\right\}$$

$$\mu_{\tilde{G}}(\omega)=\begin{cases}1 & \omega\le\sigma_L\\ \dfrac{\sigma_U-\omega}{\sigma_U-\sigma_L} & \sigma_L\le\omega\le\sigma_U\\ 0 & \sigma_U\le\omega\end{cases} \tag{22}$$

Therefore, from this degree of possibility, problems (8) and (9) are equivalently transformed into the following problems performing the method similar to (18):

$$\begin{aligned}
&\text{Maximize} \quad \prod\nolimits_{\tilde{V}}\left(\tilde{G}\right)\\
&\text{subject to} \quad \sum_{i=1}^{m} w_i^E \sum_{j=1}^{n} m_{ij}x_j + (1-h)\delta_{ij}x_j \geq (1-h) f_L + hf_U,\\
&\qquad \boldsymbol{x} \in X\\
&\qquad \text{Minimize} \quad \frac{\boldsymbol{x}^t\left(\sum_{i=1}^{m}\left(\sum_{s=1}^{S} p_s w_i^{(s)} - \sum_{s=1}^{S} p_s \alpha_{si}\right)\mathbf{V}_i\right)\boldsymbol{x} - \sigma_U}{\boldsymbol{x}^t\left(\sum_{i=1}^{m}\left(\sum_{s=1}^{S} p_s \alpha_{si}\right)\mathbf{V}_i\right)\boldsymbol{x} + (\sigma_U - \sigma_L)}\\
&\Leftrightarrow \quad \text{subject to} \quad \sum_{i=1}^{m} w_i^E \sum_{j=1}^{n} m_{ij}x_j + (1-h)\delta_{ij}x_j \geq (1-h) f_L + hf_U,\\
&\qquad \boldsymbol{x} \in X,\ w_i^{(s)} \in \left[\underline{w}_i^{(s)}(h), \overline{w}_i^{(s)}(h)\right],\ \sum_{i=1}^{m} w_i^{(s)} = 1
\end{aligned} \tag{23}$$

and

$$\begin{aligned}
&\text{Maximize} \quad \frac{\sum_{i=1}^{m} w_i^E \sum_{j=1}^{n}\left(m_{ij}x_j + \delta_{ij}x_j\right) - f_L}{\sum_{i=1}^{m} w_i^E \sum_{j=1}^{n} \delta_{ij}x_j + f_U - f_L}\\
&\text{subject to} \quad \prod\nolimits_{\tilde{V}}\left(\tilde{G}\right) \geq \bar{h},\ \boldsymbol{x} \in X\\
&\qquad \text{Maximize} \quad \frac{\sum_{i=1}^{m} w_i^E \sum_{j=1}^{n}\left(m_{ij}x_j + \delta_{ij}x_j\right) - f_L}{\sum_{i=1}^{m} w_i^E \sum_{j=1}^{n} \delta_{ij}x_j + f_U - f_L}\\
&\Leftrightarrow \quad \text{subject to} \quad \boldsymbol{x}^t\left(\sum_{i=1}^{m}\left(\sum_{s=1}^{S} p_s w_i^{(s)} - (1-h)\sum_{s=1}^{S} p_s \alpha_{si}\right)\mathbf{V}_i\right)\boldsymbol{x} \leq \bar{h}\sigma_L + (1-\bar{h})\sigma_U,\\
&\qquad \boldsymbol{x} \in X,\ w_i^{(s)} \in \left[\underline{w}_i^{(s)}(h), \overline{w}_i^{(s)}(h)\right],\ \sum_{i=1}^{m} w_i^{(s)} = 1
\end{aligned} \tag{24}$$

Problem (23) is similar to problem (17), but it is a semi-infinite programming problem since it includes interval values $w_i^{(s)} \in \left[\underline{w}_i^{(s)}(h), \overline{w}_i^{(s)}(h)\right]$ and constraint $\sum_{i=1}^{m} w_i^{(s)} = 1$. Therefore, it is hard to solve it as the original form using the standard approaches. However, since these problems are equivalently transformed into the linear programming problems using the solution approach proposed by Lai and Wu [18] and Hasuike [7], we analytically and efficiently solve it. Then, with respect to problem (23), the numerator and denominator of objective function are convex functions. Therefore, we analytically solve it using the solution method proposed by Dinkelbach [4] and the semi-infinite solution approaches [7] and [18].

On the other hand, we also consider the case where each occurrence probability includes an ambiguity and is assumed to be the following membership function in a way similar to (19):

$$\boldsymbol{w}=\begin{cases}\boldsymbol{w}^{(1)} & \Pr\left\{\boldsymbol{w}=\boldsymbol{w}^{(1)}\right\}=\tilde{p}_1\\ \boldsymbol{w}^{(2)} & \Pr\left\{\boldsymbol{w}=\boldsymbol{w}^{(2)}\right\}=\tilde{p}_2\\ \vdots & \vdots\\ \boldsymbol{w}^{(S)} & \Pr\left\{\boldsymbol{w}=\boldsymbol{w}^{(S)}\right\}=\tilde{p}_S\end{cases}$$

$$\mu_{\tilde{p}_s}(\omega)=\begin{cases}\dfrac{\omega-\left(\overline{p}_s-\xi_s\right)}{\xi_s} & \left(\overline{p}_s-\xi_s\le\omega\le\overline{p}_s\right)\\ \dfrac{\left(\overline{p}_s+\zeta_s\right)-\omega}{\zeta_s} & \left(\overline{p}_s<\omega\le\overline{p}_s+\zeta_s\right)\\ 0 & \left(\omega<\overline{p}_s-\xi_s\ ,\overline{p}_s+\zeta_s<\omega\right)\end{cases} \tag{25}$$

In this case, we assume that each possible occurrence probability p_i includes the h-cut set of the fuzzy number $\tilde{p}_i$. Then, we also apply similar transformations and degree of possibility to this case, and we equivalently transform problems (8) and (9) into the following problems:

$$\begin{aligned}&\text{Minimize} && \boldsymbol{x}^t\left(\boldsymbol{w}_E\mathbf{V}_i\right)\boldsymbol{x}\\ &\text{subject to} && \sum_{j=1}^{n}m_jx_j+\left(1-h\right)\sum_{j=1}^{n}\delta_jx_j\ge\left(1-h\right)f_L+hf_U,\ \boldsymbol{x}\in X,\\ & && p_s\in\left\{\overline{p}_s-\left(1-\overline{h}\right)\xi_s\le p_s\le\overline{p}_s+\left(1-\overline{h}\right)\zeta_s\right\},\ \sum_{s=1}^{S}p_s=1\end{aligned} \tag{26}$$

and

$$\begin{aligned}&\text{Maximize} && \frac{\sum_{j=1}^{n}m_jx_j+\sum_{j=1}^{n}\delta_jx_j-f_L}{\sum_{j=1}^{n}\delta_jx_j+f_U-f_L}\\ &\text{subject to} && \boldsymbol{x}^t\left(\boldsymbol{w}_E\mathbf{V}_i\right)\boldsymbol{x}\le\overline{h}\sigma_L+\left(1-\overline{h}\right)\sigma_U,\ \boldsymbol{x}\in X,\\ & && p_s\in\left\{\overline{p}_s-\left(1-\overline{h}\right)\xi_s\le p_s\le\overline{p}_s+\left(1-\overline{h}\right)\zeta_s\right\},\ \sum_{s=1}^{S}p_s=1\end{aligned} \tag{27}$$

Consequently, problems (26) and (27) are similar to problems (23) and (24), respectively. Therefore, we obtain the optimal portfolio with respect to the various types of fuzzy portfolio selection problems with normal mixture random distributions. Furthermore, with respect to problems (23), (24), (26) and (27), these problems consider several weights and possible occurrence probability as their interval values, and so these problems are more versatile and robust portfolio models.

4 Numerical Example

In order to compare our proposed models and standard Markowitz models and to show the usefulness of our proposed model more clearly, we provide a brief numerical example based on historical data derived from Tokyo Stock Exchange. Let us consider four securities shown in Table 1, whose mean values and variances are based on historical data in the decade between 1995 and 2004.

Table 1 Sample data from Tokyo Exchange Market

	x_1	x_2	x_3	x_4
Mean	0.046	0.043	0.087	0.090
Variance	0.0836	0.0638	0.1507	0.1010

Subsequently, we introduce the asset allocation rate x_j to each security, and its upper value is assumed to be 0.4. Then, we consider the case where the fixed target profit and total variance for Markowitz model are 0.07 and 0.025, respectively. In this case, the optimal portfolio of Markowitz model minimizing the total variance is obtained as follows:

Table 2 Optimal portfolio of Markowitz model

	x_1	x_2	x_3	x_4
Problem (1)	0.192	0.231	0.226	0.351

Furthermore, we consider the more practical case where investors predict not only single scenario but also multi-scenario with respect to each future return. In this numerical example, based on the historical data of Tokyo Exchange Market, we consider three scenarios as the following Table 3.

Table 3 Mean values and variances for three scenarios with respect to future returns

		x_1	x_2	x_3	x_4
Scenario 1	Mean	0.046	0.043	0.087	0.090
	Variance	0.0836	0.0638	0.1507	0.1010
Scenario 2	Mean	0.068	0.114	0.130	0.236
	Variance	0.0936	0.0488	0.1007	0.1510
Scenario 3	Mean	0.023	0.015	0.012	0.032
	Variance	0.0736	0.0788	0.1807	0.0910

This example means that scenarios 1, 2 and 3 are neutral, positive and passive for the investor's subjectivity. Then, each weight of the scenario is provided as the following three forms:

Table 4 Weights and probabilities for scenarios

	w_1	w_2	w_3	*Prob.*
Case 1	0.5	0.25	0.25	1/3
Case 2	0.5	0.4	0.1	1/3
Case 3	0.5	0.1	0.4	1/3

Using Table 3 and each weight, we solve Case 1 of the weighted portfolio model and obtain the following optimal portfolio:

Table 5 Optimal portfolio for Case 1 of each weighted portfolio model

	x_1	x_2	x_3	x_4
Min. Var.	0.270	0.367	0.162	0.201

On the other hand, in order to consider our proposed models in Section 3, we set the fuzzy numbers and fuzzy goals. First, we assume mean values in all scenarios to be the following symmetric triangle fuzzy numbers:

Table 6 Triangle fuzzy numbers of mean values in all scenarios

	x_1	x_2	x_3	x_4
Scenario 1	<0.046,0.02>	<0.043,0.01>	<0.087,0.03>	<0.090,0.05>
Scenario 2	<0.068,0.03>	<0.114,0.04>	<0.130,0.08>	<0.236,0.03>
Scenario 3	<0.023,0.01>	<0.015,0.005>	<0.012,0.005>	<0.032,0.02>

Then, fuzzy numbers of weights and probabilities are assumed to be the following symmetric trapezoidal fuzzy numbers:

Table 7 Symmetric trapezoidal fuzzy numbers of weights and probabilities

	w_1	w_2	w_3	*Prob.*
Case 1	(0.42,0.58,0.1)	(0.21,0.29,0.05)	(0.21,0.29,0.05)	(0.3,0.35,0.1)
Case 2	(0.42,0.58,0.1)	(0.32,0.48,0.1)	(0.02,0.18,0.1)	(0.3,0.35.0.1)
Case 3	(0.42,0.58,0.1)	(0.02,0.18,0.1)	(0.32,0.48,0.1)	(0.3,0.35.0.1)

Furthermore, fuzzy goals for the total return and variance are provided as the following forms:

$$\mu_{\tilde{G}}(f)=\begin{cases}0\\ \dfrac{f-0.08}{0.06}\\ 1\end{cases},\ \mu_{\tilde{G}}(\omega)=\begin{cases}1\\ \dfrac{0.025-\sigma}{0.005}\\ 0\end{cases},$$

Using these parameters and fuzzy goals, we solve our proposed models (17), (23) and (26) minimizing the total variance, and obtain each optimal portfolio:

Table 8 Optimal portfolios for our proposed models

	x_1	x_2	x_3	x_4
Fuzzy mean (17)	0.270	0.367	0.162	0.201
Fuzzy weight (23): Case 1	0.064	0.394	0.188	0.354
Fuzzy weight (23): Case 2	0.041	0.400	0.227	0.332
Fuzzy weight (23): Case 3	0.039	0.400	0.187	0.374
Fuzzy probability (26)	0.051	0.400	0.216	0.333

Subsequently, we consider the case where an investor purchases securities at the end of 2004 according to each portfolio shown in Tables 2 and 8. Then, the total return of three models at term ends of 2005 and 2007 become the following values shown in Table 3.9, respectively.

Table 9 Total profit for the basic and our proposed models

	Term end of 2005	Term end of 2007
Markowitz model (1)	0.2042	0.2094
Fuzzy mean (17)	0.2329	0.3008
Fuzzy weight (23): Case 1	0.2489	0.2114
Fuzzy weight (23): Case 2	0.2465	0.2108
Fuzzy weight (23): Case 3	0.2519	0.2007
Fuzzy probability (26)	0.2473	0.2135

From the result in Table 9, we find that our proposed models earn more total profit than the basic Markowitz model. Particularly, we find that fuzzy weight and fuzzy probability models earns more profit in a semi-long term investment, and the fuzzy mean models earns more profit in a long-term investment. Surely, this numerical example is based on a little data, and so this conclusion may not hold true for the stock market in all conditions. However, this results show that it is possible to earn more profits by performing suitable investments according to subjectivity. Therefore, investors can deal with our model usefully according to their investment stiles.

5 Conclusion

In this paper, we have considered several portfolio selection models with normal mixture random distributions involving ambiguous factors extending Mean-Variance model. Since our proposed models are not well-defined problems due to randomness and fuzziness, we have set some criterion such as mean value and variance to stochastic aspect and possibility measure to fuzzy aspect. Then, by performing the equivalent transformations, we have constructed the efficient solution methods based on the standard mean-variance approaches. Therefore, we have developed more versatile portfolio models with randomness and fuzziness than previous standard portfolio models, and we may obtain more beneficial knowledge for the investment theory.

As the future studies, we are going to consider the case where random distributions are more general patterns not only based on the normal distribution but also ellipsoidal distributions including almost all statistical distributions. Then, we also need to consider the case that the optimal solutions are restricted to be integers.

References

[1] Bawa, V.S., Lindenberg, E.B.: Capital market equilibrium in a mean-lower partial moment framework. Journal of Financial Economics 5, 189–200 (1977)

[2] Bilbao-Terol, A., Perez-Gladish, B., Arenas-Parra, M., Rodriguez-Uria, M.V.: Fuzzy compromise programming for portfolio selection. Applied Mathematics and Computation 173, 251–264 (2006)

[3] Carlsson, C., Fuller, R., Majlender, P.: A possibilistic approach to selecting portfolios with highest utility score. Fuzzy Sets and Systems 131, 12–21 (2002)

[4] Dinkelbach, W.: On nonlinear fractional programming. Management Science 13, 492–498 (1967)

[5] Elton, E.J., Gruber, M.J.: Modern Portfolio Theory and Investment Analysis. Wiley, New York (1995)

[6] Guo, P., Tanaka, H.: Possibility data analysis and its application to portfolio selection problems. Fuzzy Economic Rev. 3, 3–23 (1998)

[7] Hasuike, T., Ishii, H.: Probability maximization model of portfolio selection problem under ambiguity. Central European Journal of Operational Research (to appear)

[8] Hasuike, T., Katagiri, H., Ishii, H.: Portfolio selection problems with random fuzzy variable returns. In: Proceedings of IEEE International Conference on Fuzzy Systems 2007, pp. 416–421 (2007)
[9] Hasuike, T., Katagiri, H., Ishii, H.: Probability Maximization Model of 0-1 Knapsack Problem with Random Fuzzy Variables. In: Proceedings of 2008 IEEE World Congress in Computing Intelligence (WCCI 2008), IEEE International Conference on Fuzzy Systems 2008, pp. 548–554 (2008)
[10] Huang, X.: Fuzzy chance-constrained portfolio selection. Applied Mathematics and Computation 177, 500–507 (2006)
[11] Hung, X.: Two new models for portfolio selection with stochastic returns taking fuzzy information. European Journal of Operational Research 180, 396–405 (2007)
[12] Inuiguchi, M., Ramik, J.: Possibilisitc linear programming: A brief review of fuzzy mathematical programming and a comparison with stochastic programming in portfolio selection problem. Fuzzy Sets and Systems 111, 3–28 (2000)
[13] Inuiguchi, M., Tanino, T.: Portfolio selection under independent possibilistic information. Fuzzy Sets and Systems 115, 83–92 (2000)
[14] Katagiri, H., Ishii, H., Sakawa, M.: On fuzzy random linear knapsack problems. Central European Journal of Operations Research 12(1), 59–70 (2004)
[15] Katagiri, H., Sakawa, M., Ishii, H.: A study on fuzzy random portfolio selection problems using possibility and necessity measures. Scientiae Mathematicae Japonicae 65(2), 361–369 (2005)
[16] Konno, H.: Piecewise linear risk functions and portfolio optimization. Journal of Operations Research Society of Japan 33, 139–156 (1990)
[17] Konno, H., Shirakawa, H., Yamazaki, H.: A mean-absolute deviation-skewness portfolio optimization model. Annals of Operations Research 45, 205–220 (1993)
[18] Lai, H.C., Wu, S.Y.: On linear semi-infinite programming problems: an algorithm. Numerical Functional Analysis and Optimization 13, 287–304 (1992)
[19] Luenberger, D.G.: Investment Science. Oxford Univ. Press, Oxford (1997)
[20] Markowitz, H.: Portfolio Selection. Wiley, New York (1959)
[21] McNeil, A.J., Frey, R., Embrechts, P.: Quantitative Risk Management: Concepts, Techniques & Tools. Princeton University Press, Princeton (2005)
[22] Rockafellar, R.T., Uryasev, S.: Optimization of conditional value-at-risk. Journal of Risk 2(3), 1–21 (2000)
[23] Steinbach, M.C.: Markowitz revisited: Mean-variance models in financial portfolio analysis. SIAM Review 43(1), 31–85 (2001)
[24] Tanaka, H., Guo, P.: Portfolio selection based on upper and lower exponential possibility distributions. European Journal of Operational Researches 114, 115–126 (1999)
[25] Tanaka, H., Guo, P., Turksen, I.B.: Portfolio selection based on fuzzy probabilities and possibility distributions. Fuzzy Sets and Systems 111, 387–397 (2000)
[26] Watada, J.: Fuzzy portfolio selection and its applications to decision making. Tatra Mountains Math. Pub. 13, 219–248 (1997)

Fuzzy Random Redundancy Allocation Problems

Shuming Wang and Junzo Watada

Abstract. Due to subjective judgement, imprecise human knowledge and perception in capturing statistical data, the real data of lifetimes in many systems are both random and fuzzy in nature. Based on the fuzzy random variables that are used to characterize the lifetimes, this paper studies the redundancy allocation problems to a fuzzy random parallel-series system.

Two fuzzy random redundancy allocation models (FR-RAM) are developed through reliability maximization and cost minimization, respectively. Some properties of the FR-RAM are obtained, where an analytical formula of reliability with convex lifetimes is derived and the sensitivity of the reliability is discussed. To solve the FR-RAMs, we first address the computation of reliability. A random simulation method based on the derived analytical formula is proposed to compute the reliability with convex lifetimes. As for the reliability with nonconvex lifetimes, the technique of fuzzy random simulation together with the discretization method of fuzzy random variable is employed to compute the reliability, and a convergence theorem of the fuzzy random simulation is proved. Subsequently, we integrate the computation approaches of the reliability and genetic algorithm (GA) to search for the approximately optimal redundancy allocation of the models. Finally, some numerical examples are provided to illustrate the feasibility of the solution algorithm and quantify its effectiveness.

Keywords: Reliability, Redundancy allocation, Parallel-series system, Fuzzy random variable, Sensitivity, Convergence, Genetic algorithm.

1 Introduction

Reliability engineering has attracted a lot of researchers owing to its critical importance in various kinds of systems. The primary goal of reliability

Shuming Wang · Junzo Watada
Graduate School of Information, Production and Systems, Waseda University,
2-7 Hibikino, Wakamatsu, Kitakyushu 808-0135, Fukuoka, Japan
e-mail: smwangips@gmail.com, junzow@osb.att.ne.jp

W.A. Lodwick & J. Kacprzyk (Eds.): Fuzzy Optimization, STUDFUZZ 254, pp. 425–456.
springerlink.com

engineering is to improve the reliability of a system. Redundancy allocation is a direct way of enhancing the system reliability, which involves the selection of the optimal combination of components and a system-level design configuration so as to maximize the reliability under the given cost and weight constraints, or alternatively, to meet reliability and weight constraints at a minimum cost.

In conventional reliability models, an underlying assumption is that all the lifetimes of the components are characterized by random variables. Various kinds of reliability models have been proposed for different optimization purposes, and a number of methods have been proposed to solve those classical reliability optimization models. For instance, Elegbede *et al.* [5] studied the allocation of reliability and redundancy to parallel-series systems with an objective of cost minimization, and proposed an ECAY algorithm approach to solve the model. Prasad and Raghavachari [30] developed an approximate linear programming model for the optimal allocation problem in a series-parallel system, and designed a heuristic solution method. Yu *et al.* [41] studied the design of a redundant system with the consideration of the redundant dependency, and introduced a dependency function to quantify the redundant dependency. There are more various studies that examined the stochastic reliability optimization problems [1, 13, 14, 33].

To enhance the reliability of the systems in which the lifetimes are imprecise or vague, some fuzzy reliability optimization problems were studied using the fuzzy set theory [28, 29, 42, 43]. Li *et al.* [17] proposed a fuzzy linear regression based fuzzy stress-random interference model to evaluate the fuzzy reliability of the mechanical structure. Mahapatra and Roy [25] discussed a fuzzy multi-objective optimization method for a multi-objective system reliability problem which involves several mutually conflicting objectives. Zhao and Liu [47] considered a standby redundancy system with fuzzy lifetimes, and built three kinds of standby redundancy optimization models with different optimization criteria.

Nevertheless, in real-world applications, statistic data for the lifetime distributions are never precise or completely vague. Due to the subjective judgement, and imprecise human knowledge and perception in capturing such statistic data, the randomness and fuzziness are often mixed up in the lifetime data of systems. There are only a few researches that consider reliability optimization problems with such hybrid uncertainty. Zhao and Liu [46] modeled three types of system performance based on random fuzzy lifetime parameters. In their study [46], the lifetimes of the components were treated as random fuzzy variables [18], which are some "fuzzy" variables taking on random values (or more precisely, functions from possibility space to a collection of random variables). Differing from the random fuzzy variable, fuzzy random variable was introduced by Kwakernaak [15, 16] in 1978 and was defined as a measurable function from probability space to a collection of fuzzy variables. Fuzzy random variable copes with the hybrid uncertainty where the vagueness is embedded into the random variables, or in other words, the case

when "random" variables take on the fuzzy values. For example, owing to the imprecise measurement, the real distribution of the lifetime X of a component can be given in the following form: about 6 (hours) with probability 0.05, about 8 (hours) with probability 0.1, about 10 (hours) with probability 0.15, and so on. In the above distribution, the realizations of the "random" variable X are not crisp values but fuzzy numbers, say "about 8 (hours)". In the above distribution, the realizations of the "random" variable X are not crisp values but fuzzy numbers, say "about 8 (hours)". In such a case, X becomes a fuzzy random variable, and none of random variable, fuzzy variable, and even random fuzzy variable is applicable to studying such uncertainty. Following the ideas of Kwakernaak, several variants as well as extensions of fuzzy random variable were presented subsequently by other researchers such as Kruse and Meyer [12], Liu and Liu [22], and López-Diaz and Gil [24]. Fuzzy random variable has been a basic tool in constructing the framework of decision making models under fuzzy random environment, and a number of practical optimization problems have been studied based on fuzzy random variables, such as inventory (see [2, 3, 4, 40]), risk management (see [10, 23]), portfolio selection (see [9, 32]), renewal process (see [37, 48]), and regression analysis (see [38, 39]). Nevertheless, the reliability and redundancy optimization models under fuzzy random environment have not been well established in the literature.

Making use of fuzzy random variable as a tool to characterize the lifetimes of components, this paper aims to study the redundancy allocation problems to a fuzzy random parallel-series system. In this work, two fuzzy random redundancy allocation models (FR-RAM) are developed, and some properties of the FR-RAM, the computation of fuzzy random reliability function, as well as the solution of FR-RAM are discussed. In Section 2 of this paper, some basic concepts on fuzzy variables and fuzzy random variables are recalled. Section 3 formulates the FR-RAM and discusses its properties. The computation of the reliability is discussed in Section 4, while Section 5 focuses on the solution algorithm. Two numerical examples are provided in Section 6, and Section 7 presents the conclusions.

2 Preliminaries

In this section, we recall some basic concepts on fuzzy variable and fuzzy random variable which make it easier to follow further discussions on the models. Assume that $(\Gamma, \mathcal{P}(\Gamma), \mathrm{Pos})$ is a possibility space, where $\mathcal{P}(\Gamma)$ is the power set of Γ, X is a fuzzy variable defined on $(\Gamma, \mathcal{P}(\Gamma), \mathrm{Pos})$ with membership function μ_X, and r is a real number. As a well-known fuzzy measure, possibility measure of a fuzzy event $X \leq r$ is defined as

$$\mathrm{Pos}\{X \leq r\} = \sup_{t \leq r} \mu_X(t). \tag{1}$$

Lacking the self-duality, the possibility measure is not always the optimal approach to characterizing the fuzziness or vagueness in decision making problems. As a simple example, we consider an event $X > 3$ induced by a triangular fuzzy variable $X = (1, 2, 10)$. Through possibility, we can calculate the confidence level of $X > 3$ is 0.875. However, this event with such "high" confidence level is not justifiable. Because, the possibility of the opposite event, i.e., $X \leq 3$, is 1. This fact makes decision-makers confused. To overcome the above drawback, a self-dual set function, named credibility measure, is formed by [20] as follows

$$\mathrm{Cr}\{X \leq r\} = \frac{1}{2}\left(1 + \sup_{t \leq r} \mu_X(t) - \sup_{t > r} \mu_X(t)\right). \tag{2}$$

In the above example, we can calculate by credibility the confidence of $X > 3$ is 0.4345, and the confidence level of $X \leq 3$ based on credibility is $1 - 0.4345 = 0.5655$. The readers who are interested in credibility measure may refer to [19, 20].

A fuzzy variable X is said to be positive if the credibility of $X \leq 0$ is zero, i.e., $\mathrm{Cr}\{X \leq 0\} = 0$. Furthermore, fuzzy variable X is said to be convex if all the α-cut sets of X are convex sets on $\Re$. In addition, for an n-ary fuzzy vector $\boldsymbol{X} = (X_1, X_2, \cdots, X_n)$, where each individual coordinate X_k is a fuzzy variable for $k = 1, 2, \cdots, n$, the membership function of $\boldsymbol{X}$ is given by taking the minimum of the individual coordinates as follows

$$\mu_{\boldsymbol{X}}(\boldsymbol{t}) = \bigwedge_{i=1}^{n} \mu_{X_i}(t_i), \tag{3}$$

where $\boldsymbol{t} = (t_1, \cdots, t_n) \in \Re^n$.

Next, we introduce the concepts on fuzzy random variable. Roughly speaking, a fuzzy random variable is a random variable taking on fuzzy values(see [15]). Based on Kwakernaak's pioneering work [15, 16], Kruse and Meyer [12] formalized the mathematical concept of the fuzzy random variable by defining it as a fuzzy observation of a classical real-valued random variable under different measurability conditions from [15]. López-Diaz and Gil [24] discussed constructive definitions of fuzzy random variables and integrably bounded fuzzy random variables based on Hausdorff convergence. For the purpose of fuzzy random optimization, a modified fuzzy random variable was given by Liu and Liu [22], and a mean chance was defined in [23] for measuring events in fuzzy random decision-making systems. For more detailed theoretical foundation as well as detailed discussions on fuzzy random variable, one may refer to [19, 35, 36, 44, 45].

Definition 1 ([22]). *Suppose that $(\Omega, \Sigma, \mathrm{Pr})$ is a probability space, $\mathcal{F}_v$ is a collection of fuzzy variables defined on possibility space $(\Gamma, \mathcal{P}(\Gamma), \mathrm{Pos})$. A fuzzy random variable is a map $\xi : \Omega \to \mathcal{F}_v$ such that for any Borel subset B of $\Re$, $\mathrm{Pos}\{\xi(\omega) \in B\}$ is a measurable function of ω.*

Example 1. Let X be a random variable defined on probability space $(\Omega, \Sigma, \Pr)$. We call ξ a triangular fuzzy random variable, if for every $\omega \in \Omega$, $\xi(\omega)$ is a triangular fuzzy variable defined on some possibility space $(\Gamma, \mathcal{P}(\Gamma), \mathrm{Pos})$, e.g.,

$$\xi(\omega) = \Big(X(\omega) - 1, X(\omega), X(\omega) + 1\Big).$$

We say ξ is a normal fuzzy random variable, denoted by $\mathcal{N}_{\mathcal{F}}(X, b), b > 0$, if for every $\omega \in \Omega$, the membership function of $\xi(\omega)$ is

$$\mu_{\xi(\omega)}(r) = \exp\left(\frac{-(r - X(\omega))^2}{b}\right).$$

In addition, a fuzzy random variable ξ is said to be positive if for almost every $\omega \in \Omega$, $\xi(\omega)$ is a positive fuzzy variable. For example, we can construct a positive normal fuzzy random variable ξ as

$$\mu_{\xi(\omega)}(r) = \begin{cases} \exp\left(-(r - X(\omega))^2/b\right), & r \geq 0 \\ 0, & r < 0. \end{cases} \tag{4}$$

In this paper, the above positive normal fuzzy random variable ξ is denoted by $\mathcal{N}_{\mathcal{F}}^{+}(X, b)$.

In order to measure an event $\xi \in B$ induced by fuzzy random variable ξ, where B is any Borel subset of $\Re$, the mean chance measure (see [23]) is defined as

$$\mathrm{Ch}\left\{\xi \in B\right\} = \int_{\Omega} \mathrm{Cr}\left\{\xi(\omega) \in B\right\} \mathrm{Pr}(\mathrm{d}\,\omega). \tag{5}$$

Example 2. Consider a triangular fuzzy random variable ξ with $\xi(\omega) = (X(\omega)+2, X(\omega)+3, X(\omega)+4)$, where X is a discrete random variable, which takes on values $X_1 = 2$ with probability 0.4, and $X_2 = 4$ with probability 0.6. Now we calculate the mean chance of event $\xi \leq 7$.

Note that fuzzy random variable ξ takes on fuzzy variables $\xi(X_1) = (4, 5, 6)$ with probability 0.4, and $\xi(X_2) = (6, 7, 8)$ with probability 0.6, by the definition, we can work out $\mathrm{Cr}\{\xi(X_1) \leq 7\} = 1$, and $\mathrm{Cr}\{\xi(X_2) \leq 7\} = 0.5$. From (5), we have $\mathrm{Ch}\{\xi \leq 7\} = \int_{\Omega} \mathrm{Cr}\left\{\xi(\omega) \leq 7\right\} \mathrm{Pr}(\mathrm{d}\,\omega) = 1 \times 0.4 + 0.5 \times 0.6 = 0.7$.

3 Fuzzy Random Redundancy Allocation Models

3.1 Problem Formulation

This paper considers a parallel-series system composed of s subsystems (Figure 1). Each subsystem i $(1 \leq i \leq s)$ is made up of actively redundant components in parallel. The lifetimes of the components are characterized by

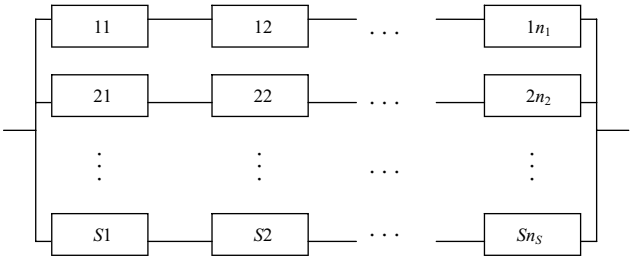


Fig. 1 An s-stage parallel-series system

fuzzy random variables. Our problem is to find the optimal redundancy allocations to this fuzzy random parallel-series system so as to maximize system reliability, or to minimize the total cost of the system. Based on those two different objectives, two fuzzy random redundancy allocation models (FRRAM) will be built in this section, respectively.

Notation

s number of subsystems
i index of subsystems, $1 \leq i \leq s$
n_i number of different component types available for subsystem i
l_i, u_i lower, and upper bounds on the number of redundant components *insubsystem* $i, 1 \leq i \leq s$
$x_{i,j}$ number of components of type j in subsystem $i, 1 \leq j \leq n_i$
$\boldsymbol{x}$ decision vector $(x_{1,1}, \cdots, x_{1,n_1}, \cdots, x_{s,1}, \cdots, x_{s,n_s})$
$\xi_{i,j,k}$ the fuzzy random lifetime of component k of type j in the *subsystem* i $1 \leq k \leq x_{ij}, 1 \leq j \leq n_i, 1 \leq i \leq s$
T^0 preselected threshold system lifetime
$R_{T^0}(\boldsymbol{x})$ system reliability for a decision $\boldsymbol{x}$ at the threshold lifetime T^0
c_{ij} the cost of each component of type j in subsystem $i, 1 \leq j \leq n_i$, $1 \leq i \leq s$
c^0 the maximum capital available
R^0 the target overall reliability of system

Assumptions

1. All the lifetimes of the components are treated as fuzzy random variables.

2. The redundancy level of subsystem i is bounded below by l_i and above by u_i.

3. The components of the same type have independent and identically distributed (i.i.d.) lifetimes.

Mathematical modeling
Based on the above assumptions and notations, in this s-stage fuzzy random parallel-series system, if we use a fuzzy random vector

$$\boldsymbol{\xi} = \Big(\xi_{1,1,1}, \cdots, \xi_{1,1,x_{1,1}}, \cdots, \xi_{1,n_1,1}, \cdots, \xi_{1,n_1,x_{1,n_1}}, \cdots, \xi_{s,1,1}, \cdots, \xi_{s,1,x_{s,1}}, \cdots, \xi_{s,n_s,1}, \cdots, \xi_{s,n_s,x_{s,n_s}}\Big)$$

to characterize the fuzzy random lifetimes of the components, the system lifetime at allocation $\boldsymbol{x}$ can be expressed as

$$T(\boldsymbol{x}, \boldsymbol{\xi}) = \bigvee_{i=1}^{s} \left[\bigwedge_{j=1}^{n_i} \left(\sum_{k=1}^{x_{i,j}} \xi_{i,j,k} \right) \right]. \tag{6}$$

Using the mean chance measure, the reliability of the fuzzy random parallel-series system can be characterized as follows

$$\begin{aligned} R_{T^0}(\boldsymbol{x}) &= \mathrm{Ch}\{T(\boldsymbol{x}, \xi) \geq T^0\}, \\ &= \int_{\Omega} \mathrm{Cr}\left\{ \bigvee_{i=1}^{s} \left[\bigwedge_{j=1}^{n_i} \left(\sum_{k=1}^{x_{i,j}} \xi_{i,j,k}(\omega) \right) \right] \geq T^0 \right\} \mathrm{Pr}(\mathrm{d}\,\omega), \end{aligned} \tag{7}$$

which is the chance that the system lifetime exceeds the threshold duration T^0.

Remark 1. If the fuzzy random vector $\boldsymbol{\xi}$ reduces to a random vector, therefore $T(\boldsymbol{x}, \boldsymbol{\xi}(\omega))$ is a crisp number for any $\omega \in \Omega$. By the definition of the mean chance, we have

$$\begin{aligned} R_{T^0}(\boldsymbol{x}) &= \mathrm{Ch}\{T(\boldsymbol{x}, \boldsymbol{\xi}) \geq T^0\}, \\ &= \int_{\Omega} \mathrm{Cr}\{T(\boldsymbol{x}, \boldsymbol{\xi}(\omega)) \geq T^0\} \,\mathrm{Pr}(\mathrm{d}\,\omega), \\ &= \int_{\Omega} I_{\{\omega | T(\boldsymbol{x}, \boldsymbol{\xi}(\omega)) \geq T^0\}}(\omega) \,\mathrm{Pr}(\mathrm{d}\,\omega), \\ &= \mathrm{Pr}\{T(\boldsymbol{x}, \boldsymbol{\xi}) \geq T^0\}, \end{aligned}$$

where I_A is the indicator function of set A. Hence, the fuzzy random reliability (7) degenerates to the reliability in the stochastic reliability theory.

Remark 2. If the fuzzy random vector $\boldsymbol{\xi}$ reduces to a fuzzy vector, then clearly the reliability (7) degenerates to $R_{T^0}(\boldsymbol{x}) = \mathrm{Cr}\{T(\boldsymbol{x}, \boldsymbol{\xi}) \geq T^0\}$ which is the reliability in the fuzzy system reliability model [47].

Maximizing the overall system reliability under the given cost c^0 and redundancy level constraints, we obtain the first model.

[FR-RAM I]

$$\begin{aligned} &\max R_{T^0}(\boldsymbol{x}) = \mathrm{Ch}\{T(\boldsymbol{x},\boldsymbol{\xi}) \geq T^0\} \\ &\text{subject to} \\ &\quad \sum_{i=1}^{s}\sum_{j=1}^{n_i} c_{ij}x_{i,j} \leq c^0, \\ &\quad l_i \leq \sum_{j=1}^{n_i} x_{i,j} \leq u_i, \text{for } i = 1,\cdots,s, \\ &\quad x_{i,j} \in \mathbb{N}, \text{for } j = 1,\cdots,n_i, i = 1,\cdots,s. \end{aligned} \tag{8}$$

Alternatively, if we minimize the total cost meeting the overall system target reliability R^0 and the redundancy level constraints, the second model can be built as follows.

[FR-RAM II]

$$\begin{aligned} &\min \sum_{i=1}^{s}\sum_{j=1}^{n_i} c_{ij}x_{i,j} \\ &\text{subject to} \\ &\quad R_{T^0}(\boldsymbol{x}) = \mathrm{Ch}\{T(\boldsymbol{x},\boldsymbol{\xi}) \geq T^0\} \geq R^0, \\ &\quad l_i \leq \sum_{j=1}^{n_i} x_{i,j} \leq u_i, \text{for } i = 1,\cdots,s, \\ &\quad x_{i,j} \in \mathbb{N}, \text{for } j = 1,\cdots,n_i, i = 1,\cdots,s. \end{aligned} \tag{9}$$

3.2 Some Properties of FR-RAM

This subsection focus on the properties of FR-RAM. In Theorem 1, we derive an analytical formula for the overall reliability $R_{T^0}(\boldsymbol{x})$ of the system, provided all the lifetimes of components have convex distributions. This formula is helpful to the computation of reliability with convex lifetimes,(which will be discussed in Section 4). Theorems 2-4 discuss the sensitivity of reliability $R_{T^0}(\boldsymbol{x})$ in FR-RAM with respect to (w.r.t.) the threshold lifetime T^0.

Theorem 1. *Assume that in the FR-RAM, the lifetimes $\xi_{i,j,k}$ of components for $i = 1,2,\cdots,s, j = 1,2,\cdots,n_i, k = 1,2,\cdots,x_{ij}$ are fuzzy random variables on probability space $(\Omega,\Sigma,\mathrm{Pr})$. Suppose for almost every $\omega \in \Omega$, $\xi_{i,j,k}(\omega)$ is a convex fuzzy variable with membership function $\mu^{\omega}_{i,j}$, and $\mu^{\omega}_{i,1}(v^{\omega}_{i,1}) = \mu^{\omega}_{i,2}(v^{\omega}_{i,2}) = \cdots = \mu^{\omega}_{i,n_i}(v^{\omega}_{i,n_i}) = 1$ with $x_{i,1}v^{\omega}_{i,1} \leq x_{i,2}v^{\omega}_{i,2} \leq \cdots \leq x_{i,n_i}v^{\omega}_{i,n_i}$. If we denote μ^{ω}_{T} the membership function of fuzzy variable $T(\boldsymbol{x},\xi(\omega))$ for any $\omega \in \Omega$, then given any allocation $\boldsymbol{x}$, the system reliability is*

$$R_{T^0}(\boldsymbol{x}) = \int_{\{\omega|T^0 \leq x_{s,1}v^{\omega}_{s,1}\}} 1 - \frac{\mu^{\omega}_{T}(T^0)}{2}\,\mathrm{Pr}(\mathrm{d}\,\omega) + \int_{\{\omega|T^0 > x_{s,1}v^{\omega}_{s,1}\}} \frac{\mu^{\omega}_{T}(T^0)}{2}\,\mathrm{Pr}(\mathrm{d}\,\omega) \tag{10}$$

where for almost every $\omega \in \Omega$, $\mu_T^\omega(t)$ *is given by*

$$\mu_T^\omega(t) = \begin{cases} \bigwedge_{i=1}^{s} \mu_i^\omega(t), & t < x_{1,1}v_{1,1}^\omega \\ \bigwedge_{i=l+1}^{s} \mu_i^\omega(t), & x_{l,1}v_{l,1}^\omega \le t < x_{l+1,1}v_{l+1,1}^\omega, 1 \le l \le s-1 \\ \bigvee_{i=1}^{s} \mu_i^\omega(t), & t \ge x_{s,1}v_{s,1}^\omega, \end{cases} \tag{11}$$

here we assume that $x_{1,1}v_{1,1}^\omega \le x_{2,1}v_{2,1}^\omega \le \cdots \le x_{s,1}v_{s,1}^\omega$ *without losing any generality, and* $\mu_i^\omega(t)$ *is calculated by*

$$\mu_i^\omega(t) = \begin{cases} \bigvee_{j=1}^{n_i} \mu_{i,j}^\omega\left(\frac{t}{x_{i,j}}\right), & t < x_{i,1}v_{i,1}^\omega \\ \bigwedge_{j=1}^{l} \mu_{i,j}^\omega\left(\frac{t}{x_{i,j}}\right), & x_{i,l}v_{i,l}^\omega \le t < x_{i,l+1}v_{i,l+1}^\omega, 1 \le l \le n_i - 1 \\ \bigwedge_{j=1}^{n_i} \mu_{i,j}^\omega\left(\frac{t}{x_{i,j}}\right), & t \ge x_{i,n_i}v_{i,n_i}^\omega. \end{cases} \tag{12}$$

Proof. First of all, recalling that all lifetimes are convex and the lifetimes of the same type of components are identically distributed, given any i, j, and $t > 0$ we have

$$\mathrm{Pos}\left\{\sum_{k=1}^{x_{i,j}} \xi_{i,j,k}(\omega) = t\right\} = \mathrm{Pos}\left\{\xi_{i,j,1}(\omega) = \frac{t}{x_{i,j}}\right\} = \mu_{i,j}^\omega\left(\frac{t}{x_{i,j}}\right)$$

for almost every $\omega \in \Omega$, which also is a convex fuzzy number. Furthermore,

$$\mathrm{Pos}\left\{\sum_{k=1}^{x_{i,j}} \xi_{i,j,k}(\omega) = x_{i,j}v_{i,j}^\omega\right\} = \mu_{i,j}^\omega\left(v_{i,j}^\omega\right) = 1$$

for $i = 1, 2, \cdots, s; j = 1, 2, \cdots, n_i$ and $x_{i,1}v_{i,1}^\omega \le x_{i,2}v_{i,2}^\omega \le \cdots \le x_{i,n_i}v_{i,n_i}^\omega$. Therefore, by the minimum t-norm operation of convex fuzzy numbers [11, Theorem 1], we obtain the membership function $\mu_i^\omega(t)$ of $\bigwedge_{j=1}^{n_i} \sum_{k=1}^{x_{i,j}} \xi_{i,j,k}(\omega)$ is

$$\mu_i^\omega(t) = \begin{cases} \bigvee_{j=1}^{n_i} \mu_{i,j}^\omega\left(\frac{t}{x_{i,j}}\right), & t < x_{i,1}v_{i,1}^\omega \\ \bigwedge_{j=1}^{l} \mu_{i,j}^\omega\left(\frac{t}{x_{i,j}}\right), & x_{i,l}v_{i,l}^\omega \le t < x_{i,l+1}v_{i,l+1}^\omega, 1 \le l \le n_i - 1 \\ \bigwedge_{j=1}^{n_i} \mu_{i,j}^\omega\left(\frac{t}{x_{i,j}}\right), & t \ge x_{i,n_i}v_{i,n_i}^\omega, \end{cases}$$

for $i = 1, 2, \cdots, s$. Note that $\mu_i^\omega(t)$ is also a convex fuzzy number, which is nondecreasing in $[-\infty, x_{i,1}v_{i,1}]$ and nonincreasing in $[x_{i,1}v_{i,1}^\omega, \infty]$ for all i, and $x_{1,1}v_{1,1}^\omega \le x_{2,1}v_{2,1}^\omega \le \cdots \le x_{s,1}v_{s,1}^\omega$, making use of the maximum t-norm operation of convex fuzzy numbers [11, Theorem 1], the membership function $\mu_T^\omega(t)$ of the system lifetime

$$T\Big(\boldsymbol{x}, \boldsymbol{\xi}(\omega)\Big) = \bigvee_{i=1}^{s} \left[\bigwedge_{j=1}^{n_i} \left(\sum_{k=1}^{x_{i,j}} \xi_{i,j,k}(\omega) \right) \right]$$

can be given in the following form

$$\mu_T^{\omega}(t) = \begin{cases} \bigwedge_{i=1}^{s} \mu_i^{\omega}(t), & t < x_{1,1} v_{1,1}^{\omega} \\ \bigwedge_{i=l+1}^{s} \mu_i^{\omega}(t), & x_{l,1} v_{l,1}^{\omega} \le t < x_{l,1} v_{l+1,1}^{\omega}, 1 \le l \le s-1 \\ \bigvee_{i=1}^{s} \mu_i^{\omega}(t), & t \ge x_{s,1} v_{s,1}^{\omega}. \end{cases}$$

Since $\mu_T^{\omega}(t)$ is nondecreasing in $[-\infty, x_{s,1} v_{s,1}^{\omega}]$ and nonincreasing in $[-\infty, x_{s,1} v_{s,1}^{\omega}]$, by the definition, we can calculate

$$\mathrm{Cr}\left\{ T\Big(\boldsymbol{x}, \boldsymbol{\xi}(\omega)\Big) \ge T^0 \right\} = \frac{1}{2} \left[1 + \sup_{t \ge T^0} \mu_T^{\omega}(t) - \sup_{t < T^0} \mu_T^{\omega}(t) \right],$$

which implies

$$\mathrm{Cr}\left\{ T\Big(\boldsymbol{x}, \boldsymbol{\xi}(\omega)\Big) \ge T^0 \right\} = \begin{cases} 1 - \frac{\mu_T^{\omega}(T^0)}{2}, & T^0 \le x_{s,1} v_{s,1}^{\omega} \\ \frac{\mu_T^{\omega}(T^0)}{2}, & \text{otherwise.} \end{cases}$$

That is

$$\mathrm{Cr}\left\{ T\Big(\boldsymbol{x}, \boldsymbol{\xi}(\omega)\Big) \ge T^0 \right\} = I_{\{\omega | T^0 \le x_{s,1} v_{s,1}^{\omega}\}}(\omega) \left(1 - \frac{\mu_T^{\omega}(T^0)}{2} \right) + I_{\{\omega | T^0 > x_{s,1} v_{s,1}^{\omega}\}}(\omega) \left(\frac{\mu_T^{\omega}(T^0)}{2} \right),$$

for almost every $\omega \in \Omega$, where I_A is the indicator function of set A. Integrating w.r.t. ω on the both sides of the above equation deduces the required result (10). The proof of the theorem is complete.

Example 3. Let us consider a 2-stage parallel-series system with 2 types component in the first subsystem and 1 type in the second, and the redundancy allocation is $\boldsymbol{x} = (x_{1,1}, x_{1,2}, x_{2,1}) = (2, 1, 1)$. Suppose that the fuzzy random lifetimes $\xi_{1,1,1}$, $\xi_{1,2,1}$ and $\xi_{2,1,1}$ are characterized by a discrete random variable which takes on values $\omega = \omega_1$ with probability 0.4 and $\omega = \omega_2$ with probability 0.6, and they have the distributions in Table 1. Selecting $T^0 = 10$, we calculate the reliability $R_{10}(\boldsymbol{x}) = \mathrm{Ch}\{T(\boldsymbol{x}, \boldsymbol{\xi}) \ge 10\}$ of the system.

Since the lifetime of the system $T(\boldsymbol{x}, \boldsymbol{\xi}) = [(\xi_{1,1,1} + \xi_{1,1,2}) \wedge \xi_{1,2,1}] \vee \xi_{2,1,1}$, we have

$$T\Big(\boldsymbol{x}, \boldsymbol{\xi}(\omega_k)\Big) = \Big[\Big(\xi_{1,1,1}(\omega_k) + \xi_{1,1,2}(\omega_k) \Big) \wedge \xi_{1,2,1}(\omega_k) \Big] \vee \xi_{2,1,1}(\omega_k)$$

for $k = 1, 2$. That is $T(\boldsymbol{x}^0, \boldsymbol{\xi})$ takes on the fuzzy values $[(4, 6, 8) \wedge (3, 5, 8)] \vee (6, 7, 8)$ with probability 0.4, and $[(6, 8, 12) \wedge (5, 8, 10)] \vee (6, 8, 10)$ with

Table 1 Lifetime of each component in Example 3

$\omega = \omega_1$ (Probability=0.4)	$\omega = \omega_2$ (Probability=0.6)
$\xi_{1,1}(\omega_1) = (2, 3, 4)$	$\xi_{1,1}(\omega_2) = (3, 4, 6)$
$\xi_{1,2}(\omega_1) = (3, 5, 8)$	$\xi_{1,2}(\omega_1) = (5, 8, 10)$
$\xi_{2,1}(\omega_1) = (6, 7, 8)$	$\xi_{2,1}(\omega_2) = (6, 8, 10)$

probability 0.6. For $\omega = \omega_1$, from (12) in Theorem 1, we can calculate that $(4, 6, 8) \wedge (3, 5, 8) = (3, 5, 8)$. Furthermore, making use of (11), we obtain $T(\boldsymbol{x}, \boldsymbol{\xi}(\omega_1)) = (3, 5, 8) \vee (6, 7, 8) = (6, 7, 8)$ and

$$\mu_T^{\omega_1}(t) = \begin{cases} t - 6, & 6 \le t < 7 \\ 8 - t, & 7 \le t < 8 \\ 0, & \text{otherwise.} \end{cases}$$

We note that $x_{2,1} = 1$ and $T^0 = 7 = x_{2,1} v_{2,1}^{\omega_1}$, therefore,

$$\mathrm{Cr}\left\{T\Big(\boldsymbol{x}, \boldsymbol{\xi}(\omega_1)\Big) \ge T^0\right\} = 1 - \frac{\mu_T^{\omega_1}(7)}{2} = 0.5.$$

Similarly, we obtain

$$\mu_T^{\omega_2}(t) = \begin{cases} (t - 6)/2, & 6 \le t < 8 \\ (10 - t)/2, & 8 \le t < 10 \\ 0, & \text{otherwise.} \end{cases}$$

Since $T^0 = 7 < x_{2,1} v_{2,1}^{\omega_2} = 8$, we have

$$\mathrm{Cr}\left\{T\Big(\boldsymbol{x}, \boldsymbol{\xi}(\omega_2)\Big) \ge T^0\right\} = 1 - \frac{\mu_T^{\omega_2}(7)}{2} = 0.75.$$

Consequently, from (10), we have

$$\begin{aligned} R_7(\boldsymbol{x}) &= \int_\Omega \mathrm{Cr}\left\{T\Big(\boldsymbol{x}, \boldsymbol{\xi}(\omega)\Big) \ge 7\right\} \mathrm{Pr}(\mathrm{d}\omega) \\ &= \int_{\{\omega | 7 \le x_{s,1} v_{s,1}^{\omega}\}} 1 - \frac{\mu_T^{\omega}(7)}{2} \mathrm{Pr}(\mathrm{d}\,\omega) \\ &= \int_\Omega 1 - \frac{\mu_T^{\omega}(7)}{2} \mathrm{Pr}(\mathrm{d}\omega) = 0.5 \times 0.4 + 0.75 \times 0.6 = 0.65. \end{aligned}$$

In the FR-RAMs I and II, we note that the reliability $R_{T^0}(\boldsymbol{x})$ is predetermined by the decision-maker, and any changes of the threshold lifetime T^0 may influence the objective value of FR-RAM I and the reliability constraint of FR-RAM II, and finally influence the allocation decision. Therefore, it is necessary to analyze the sensitivity of $R_{T^0}(\boldsymbol{x})$ w.r.t. the threshold lifetime T^0.

Theorem 2. *Let the lifetimes $\xi_{i,j,k}$ of components in the FR-RAM for $i = 1, 2, \cdots, s, j = 1, 2, \cdots, n_i, k = 1, 2, \cdots, x_{i,j}$ be fuzzy random variables on a probability space $(\Omega, \Sigma, \Pr)$ such that for almost every $\omega \in \Omega$, $\xi_{i,j,k}(\omega)$ is a convex fuzzy variable. If $\xi_{i,j,k}(\omega)$ for each i, j, k areis left continuous or upper semicontinuous, then the reliability $R_{T^0}(\boldsymbol{x})$ is left continuous w.r.t. the threshold lifetime $T^0 \in \Re$.*

Proof. First of all, we deal with the case that $\xi_{i,j,k}(\omega)$ is left continuous. Denote $\mu_{i,j}^{\omega}$ the membership function of $\xi_{i,j,k}(\omega)$ for any $\omega \in \Omega$, and $i = 1, 2, \cdots, s, j = 1, 2, \cdots, n_i, k = 1, 2, \cdots, x_{i,j}$. In the following, we prove that the membership function of fuzzy variable $\bigwedge_{j=1}^{n_i} \sum_{k=1}^{x_{i,j}} \xi_{i,j,k}(\omega)$ denoted by μ_i^{ω} is left continuous for almost every $\omega \in \Omega$.

Without losing any generality, we assume that $x_{1,1}v_{1,1}^{\omega} \leq x_{2,1}v_{2,1}^{\omega} \leq \cdots \leq x_{s,1}v_{s,1}^{\omega}$, recall that fuzzy variable $\xi_{i,j,k}(\omega)$ is convex for almost every $\omega \in \Omega$, then from Theorem 1, we have

$$\mu_i^{\omega}(t) = \begin{cases} \bigvee_{j=1}^{n_i} \mu_{i,j}^{\omega}\left(\frac{t}{x_{i,j}}\right), t < x_{i,1}v_{i,1}^{\omega} \\ \bigwedge_{j=1}^{l} \mu_{i,j}^{\omega}\left(\frac{t}{x_{i,j}}\right), x_{i,l}v_{i,l}^{\omega} \leq t < x_{i,l+1}v_{i,l+1}^{\omega}, 1 \leq l \leq n_i - 1 \\ \bigwedge_{j=1}^{n_i} \mu_{i,j}^{\omega}\left(\frac{t}{x_{i,j}}\right), t \geq x_{i,n_i}v_{i,n_i}^{\omega}. \end{cases} \tag{13}$$

For almost every $\omega \in \Omega$, from the assumption that fuzzy variable $\xi_{i,j,k}(\omega)$ is left continuous, that is, the membership function $\mu_{i,j}^{\omega}$ is a left continuous real-valued function for $i = 1, 2, \cdots, s, j = 1, 2, \cdots, n_i$, therefore, for any $t_0 \in \Re$,

$$\lim_{t \to t_0 - 0} \mu_{i,1}^{\omega}(t) \bigvee \mu_{i,2}^{\omega}(t) = \lim_{t \to t_0 - 0} \mu_{i,1}^{\omega}(t) \bigvee \lim_{t \to t_0 - 0} \mu_{i,2}^{\omega}(t) = \mu_{i,1}^{\omega}(t_0) \bigvee \mu_{i,2}^{\omega}(t_0),$$

and

$$\lim_{t \to t_0 - 0} \mu_{i,1}^{\omega}(t) \bigwedge \mu_{i,2}^{\omega}(t) = \lim_{t \to t_0 - 0} \mu_{i,1}^{\omega}(t) \bigwedge \lim_{t \to t_0 - 0} \mu_{i,2}^{\omega}(t) = \mu_{i,1}^{\omega}(t_0) \bigwedge \mu_{i,2}^{\omega}(t_0).$$

Hence, $\mu_{i,1}^{\omega} \bigvee \mu_{i,2}^{\omega}$ and $\mu_{i,1}^{\omega} \bigwedge \mu_{i,2}^{\omega}$ are left continuous. By the method of induction, we can obtain that $\bigvee_{j=1}^{n} \mu_{i,j}^{\omega}$ and $\bigwedge_{j=1}^{n} \mu_{i,j}^{\omega}$ are left continuous real-valued functions for any finite positive integer n. Therefore, from (13), we have $\mu_i^{\omega}(t)$ is left continuous in $(-\infty, x_{i,1}v_{i,1}^{\omega})$, $\left[x_{i,l}v_{i,l}^{\omega}, x_{i,l+1}v_{i,l+1}^{\omega}\right)$ for $1 \leq l \leq n_i - 1$, and $\left[x_{i,n_i}v_{i,n_i}^{\omega}, \infty\right)$, respectively. Thus, to prove the left-continuity of μ_i^{ω}, it suffices to prove μ_i^{ω} is left continuous at $x_{i,1}v_{i,1}^{\omega}, x_{i,2}v_{i,2}^{\omega}, \cdots, x_{i,n_i}v_{i,n_i}^{\omega}$ for $i = 1, 2, \cdots, s$. Given each i, for $x_{i,1}v_{i,1}^{\omega}$, we have

$$\lim_{t \to x_{i,1}v_{i,1}^{\omega} - 0} \mu_i^{\omega}(t) = \lim_{t \to x_{i,1}v_{i,1}^{\omega} - 0} \bigvee_{j=1}^{n_i} \mu_{i,j}^{\omega}\left(\frac{t}{x_{i,j}}\right) = 1 = \mu_i^{\omega}(x_{i,1}v_{i,1}^{\omega}), \tag{14}$$

which implies that μ_i^ω is left continuous at $x_{i,1}v_{i,1}^\omega$. Next, for $x_{i,l}v_{i,l}^\omega$, $2 \le l \le n_i$, we have

$$\lim_{t \to x_{i,l}v_{i,l}^\omega - 0} \mu_i^\omega(t) = \lim_{t \to x_{i,l}v_{i,l}^\omega - 0} \bigwedge_{j=1}^{l-1} \mu_{i,j}^\omega\left(\frac{t}{x_{i,j}}\right) = \bigwedge_{j=1}^{l-1} \mu_{i,j}^\omega\left(v_{i,l}^\omega\right) = \mu_i^\omega(x_{i,l}v_{i,l}^\omega). \tag{15}$$

That is, μ_i^ω is left continuous at $x_{i,l}v_{i,l}^\omega$ for $2 \le l \le n_i$. As a consequence, we have proved that μ_i^ω is left continuous for any $i = 1, 2, \cdots, s$ and almost every $\omega \in \Omega$. Furthermore, by the same reasoning, from the expression (11) of the membership function μ_T^ω of $T\Big(\boldsymbol{x}, \boldsymbol{\xi}(\omega)\Big)$, we can prove that μ_T^ω is also left continuous for almost every $\omega \in \Omega$.

In the following, we prove μ_T^ω is upper semicontinuous provided $\xi_{i,j,k}(\omega)$ is upper semicontinuous. Before this, we shall prove that given i, $\bigvee_{j=1}^n \mu_{i,j}^\omega$ and $\bigwedge_{j=1}^n \mu_{i,j}^\omega$ are upper semicontinuous real-valued functions for any finite positive integer n provided $\mu_{i,j}^\omega$ for $j = 1, 2, \cdots, n$ are upper semicontinuous. In fact, when $n = 2$, by the upper semicontinuity of $\mu_{i,1}^\omega$ and $\mu_{i,2}^\omega$, we have for any $\epsilon > 0$, there exists a $\delta > 0$ such that

$$\mu_{i,1}^\omega(t) \bigvee \mu_{i,2}^\omega(t) < \left(\mu_{i,1}^\omega(t^0) + \epsilon\right) \bigvee \left(\mu_{i,2}^\omega(t^0) + \epsilon\right) = \left(\mu_{i,1}^\omega(t^0) \bigvee \mu_{i,2}^\omega(t^0)\right) + \epsilon,$$

and

$$\mu_{i,1}^\omega(t) \bigwedge \mu_{i,2}^\omega(t) < \left(\mu_{i,1}^\omega(t^0) + \epsilon\right) \bigwedge \left(\mu_{i,2}^\omega(t^0) + \epsilon\right) = \left(\mu_{i,1}^\omega(t^0) \bigwedge \mu_{i,2}^\omega(t^0)\right) + \epsilon,$$

for any $t^0 \in \Re$. That is $\mu_{i,1}^\omega \vee \mu_{i,2}^\omega$ and $\mu_{i,1}^\omega \wedge \mu_{i,2}^\omega$ are upper semicontinuous. By the method of induction, it is not difficult to show $\bigvee_{j=1}^n \mu_{i,j}^\omega$ and $\bigwedge_{j=1}^n \mu_{i,j}^\omega$ are upper semicontinuous real-valued functions for any finite positive integer n. Furthermore, similarly as in the case that $\xi_{i,j,k}(\omega)$ is left continuous, to prove the upper semicontinuity of μ_i^ω, from (13), it suffices to prove that μ_i^ω is upper semicontinuous at $x_{i,1}v_{i,1}^\omega, x_{i,2}v_{i,2}^\omega, \cdots, x_{i,n_i}v_{i,n_i}^\omega$ for $i = 1, 2, \cdots, s$. Given $i = 1, 2, \cdots, s$, for $x_{i,1}v_{i,1}^\omega$, we have

$$\limsup_{t \to x_{i,1}v_{i,1}^\omega} \mu_i^\omega(t) \le 1 = \mu_i^\omega(x_{i,1}v_{i,1}^\omega),$$

which implies μ_i^ω is upper semicontinuous at $x_{i,1}v_{i,1}^\omega$ for each i. As to $x_{i,l}v_{i,l}^\omega$, $2 \le l \le n_i$, we note from the proof of Theorem 1 that μ_i^ω is nonincreasing in $[x_{i,1}v_{i,1}^\omega, \infty]$ for all i, hence we have

$$\limsup_{t \to x_{i,l}v_{i,l}^\omega} \mu_i^\omega(t) \le \limsup_{t \to x_{i,l}v_{i,l}^\omega - 0} \mu_i^\omega(t) = \limsup_{t \to x_{i,l}v_{i,l}^\omega - 0} \bigwedge_{j=1}^{l-1} \mu_{i,j}^\omega\left(\frac{t}{x_{i,j}}\right). \tag{16}$$

It follows from the upper semicontinuity of $\bigwedge_{j=1}^{l-1} \mu_{i,j}^{\omega}$ that

$$\limsup_{t \to x_{i,l} v_{i,l}^{\omega} - 0} \bigwedge_{j=1}^{l-1} \mu_{i,j}^{\omega}\left(\frac{t}{x_{i,j}}\right) \leq \bigwedge_{j=1}^{l-1} \mu_{i,j}^{\omega}\left(v_{i,l}^{\omega}\right) = \mu_i^{\omega}(x_{i,l} v_{i,l}^{\omega}). \tag{17}$$

Combining (16) and (17) implies that μ_i^{ω} is upper semicontinuous at $x_{i,l} v_{i,l}^{\omega}$, for $2 \leq l \leq n_i$. So far, we have proved that μ_i^{ω} is upper semicontinuous for any $i = 1, 2, \cdots, s$ and almost every $\omega \in \Omega$. Based on this fact, by the same reasoning, we can prove that μ_T^{ω} is also upper semicontinuous for almost every $\omega \in \Omega$.

Therefore, we have μ_T^{ω} is left continuous or upper semicontinuous according that $\xi_{i,j,k}(\omega)$ for each i, j, k is left continuous or upper semicontinuous. Thus, by the left-continuity condition for the distribution functions of fuzzy random variable (see [36, Corollary 3.3]), we have

$$\lim_{T \to T^0 - 0} \mathrm{Ch}\{T(\boldsymbol{x}, \boldsymbol{\xi}) \geq T\} = \mathrm{Ch}\{T(\boldsymbol{x}, \boldsymbol{\xi}) \geq T^0\}$$

for any $T^0 \in \Re$. That is the reliability $R_{T^0}(\boldsymbol{x})$ is a left continuous function of the threshold lifetime $T^0 \in \Re$.

Example 4. Consider a 2-stage parallel-series system with redundancy allocation $\boldsymbol{x} = (x_{1,1}, x_{1,2}, x_{2,1}) = (1, 1, 2)$. The components have the distributions as follows: fuzzy random lifetimes $\xi_{1,1}$, $\xi_{1,2}$ and $\xi_{2,1}$ are characterized by a discrete random variable which takes on values $\omega = \omega_1$ with probability 0.8 and $\omega = \omega_2$ with probability 0.2, and the membership functions are left continuous and upper semicontinuous which are given as below:

$$\mu_{\xi_{1,1}(\omega_1)}(t) = \begin{cases} t-2, & 2 \leq t \leq 3 \\ (4-t)/2, & 3 < t \leq 4 \\ 0, & \text{otherwise,} \end{cases}$$

$$\mu_{\xi_{1,1}(\omega_2)}(t) = \begin{cases} t-3, & 3 \leq t \leq 4 \\ (5-t)/2, & 4 < t \leq 5 \\ 0, & \text{otherwise;} \end{cases}$$

$$\mu_{\xi_{1,2}(\omega_1)}(t) = \begin{cases} (t-3)/2, & 3 \leq t \leq 5 \\ (7-t)/4, & 5 < t \leq 7 \\ 0, & \text{otherwise,} \end{cases}$$

$$\mu_{\xi_{1,2}(\omega_2)}(t) = \begin{cases} t-5, & 5 \leq t \leq 6 \\ (7-t)/2, & 6 < t \leq 7 \\ 0, & \text{otherwise;} \end{cases}$$

and

$$\mu_{\xi_{2,1}(\omega_1)}(t) = \begin{cases} t-6, & 6 \leq t \leq 7 \\ (8-t)/2, & 7 < t \leq 8 \\ 0, & \text{otherwise,} \end{cases}$$

$$\mu_{\xi_{2,1}(\omega_2)}(t) = \begin{cases} t-7, & 7 \le t \le 8 \\ (9-t)/2, & 8 < t \le 9 \\ 0, & \text{otherwise.} \end{cases}$$

Now, we verify the left continuity of $R_T(\boldsymbol{x})$ w.r.t. T.

First of all, from Theorem 1, we have the membership function $\mu_T^{\omega_1}(t)$ of

$$T\Big(\boldsymbol{x}, \boldsymbol{\xi}(\omega_1)\Big) = \Big(\xi_{1,1,1}(\omega_1) \wedge \xi_{1,2,1}(\omega_1)\Big) \bigvee \Big(\xi_{2,1,1}(\omega_1) + \xi_{2,1,2}(\omega_1)\Big)$$

is

$$\mu_T^{\omega_1}(t) = \begin{cases} (t-12)/2, & 12 < t \le 14 \\ (16-t)/4, & 14 < t \le 16 \\ 0, & \text{otherwise.} \end{cases}$$

Furthermore, we have

$$\mathrm{Cr}\left\{T\Big(\boldsymbol{x}, \boldsymbol{\xi}(\omega_1)\Big) \ge T\right\} = \begin{cases} 1, & T \le 12 \\ (16-T)/4, & 12 < T \le 14 \\ (16-T)/8, & 14 < T \le 16 \\ 0, & \text{otherwise.} \end{cases}$$

Similarly, we can obtain

$$\mu_T^{\omega_2}(t) = \begin{cases} (t-14)/2, & 14 < t \le 16 \\ (18-t)/4, & 16 < t \le 18 \\ 0, & \text{otherwise,} \end{cases}$$

and

$$\mathrm{Cr}\left\{T\Big(\boldsymbol{x}, \boldsymbol{\xi}(\omega_2)\Big) \ge T\right\} = \begin{cases} 1, & T \le 14 \\ (18-T)/4, & 14 < T \le 16 \\ (18-T)/8, & 16 < T \le 18 \\ 0, & \text{otherwise.} \end{cases}$$

As a consequence, by the definition, we have

$$\mathrm{Ch}\{T(\boldsymbol{x}, \boldsymbol{\xi}) \ge T\} = \begin{cases} 1, & T \le 12 \\ (17-T)/5, & 12 < T \le 14 \\ (50-3T)/20, & 14 < T \le 16 \\ (18-T)/20, & 16 < T \le 18 \\ 0, & \text{otherwise,} \end{cases}$$

which is a left continuous function of $T \in \Re$. This result coincides with that of Theorem 2.

Theorem 3. *Suppose that the lifetimes $\xi_{i,j,k}$ of components in the FR-RAM for $i = 1, 2, \cdots, s, j = 1, 2, \cdots, n_i, k = 1, 2, \cdots, x_{i,j}$ are fuzzy random variables on probability space $(\Omega, \Sigma, \mathrm{Pr})$ such that for almost every $\omega \in \Omega$, $\xi_{i,j,k}(\omega)$ is a convex fuzzy variable. If $\xi_{i,j,k}(\omega)$ for each i, j, k is right*

continuous and lower semicontinuous, then the reliability $R_{T^0}(\boldsymbol{x})$ *is right continuous w.r.t. the threshold lifetime* $T^0 \in \Re$.

Proof. Under the assumption that for almost every $\omega \in \Omega$, $\xi_{i,j,k}(\omega)$ for each i, j, k is right continuous and lower semicontinuous, by the same proof as in Theorem 2, we can prove μ_i^ω is also right continuous and lower semicontinuous for $i = 1, 2, \cdots, s$. Furthermore, we can similarly prove that μ_T^ω is right continuous and lower semicontinuous for $i = 1, 2, \cdots, s$ for almost every $\omega \in \Omega$. Thus from the right-continuity condition for the distribution functions of fuzzy random variable (see [36, Theorem 3.7]), we have

$$\lim_{T \to T^0 + 0} \mathrm{Ch}\{T(\boldsymbol{x}, \boldsymbol{\xi}) \geq T\} = \mathrm{Ch}\{T(\boldsymbol{x}, \boldsymbol{\xi}) \geq T^0\}$$

for any $T^0 \in \Re$. That is the reliability $R_{T^0}(\boldsymbol{x})$ is right continuous of $T^0 \in \Re$.

Theorem 4. *Assume that the lifetimes* $\xi_{i,j,k}$ *of components in the FR-RAM for* $i = 1, 2, \cdots, s, j = 1, 2, \cdots, n_i, k = 1, 2, \cdots, x_{i,j}$ *are fuzzy random variables on probability space* $(\Omega, \Sigma, \mathrm{Pr})$ *such that for almost every* $\omega \in \Omega$, $\xi_{i,j,k}(\omega)$ *is a convex fuzzy variable. If* $\xi_{i,j,k}(\omega)$ *for each* i, j, k *is continuous, then we have the reliability* $R_{T^0}(\boldsymbol{x})$ *is continuous w.r.t. the threshold lifetime* $T^0 \in \Re$.

Proof. Similarly as in Theorem 3, we can prove that for almost every $\omega \in \Omega$, μ_i^ω is continuous for all i provided that $\xi_{i,j,k}(\omega)$ for each i, j, k is continuous. Furthermore, μ_T^ω is also continuous for almost every $\omega \in \Omega$. Therefore, by the continuity condition for the distribution functions of fuzzy random variable (see [36, Theorem 3.10]), we have

$$\lim_{T \to T^0} \mathrm{Ch}\{T(\boldsymbol{x}, \boldsymbol{\xi}) \geq T\} = \mathrm{Ch}\{T(\boldsymbol{x}, \boldsymbol{\xi}) \geq T^0\}$$

for any $T^0 \in \Re$. That is the reliability $R_{T^0}(\boldsymbol{x})$ is a right continuous function of $T^0 \in \Re$.

4 Computation of Reliability

Containing fuzzy random parameters, the reliability function $R_{T^0}(\boldsymbol{x})$ in general cannot be calculated directly. Furthermore, we note from the FR-RAMs I and II that the reliability function $R_{T^0}(\boldsymbol{x})$ is the objective of FR-RAM I and a constraint of FR-RAM II, therefore, in order to solve the FR-RAMs, we at first have to deal with the computation of the system reliability $R_{T^0}(\boldsymbol{x})$.

Let $\boldsymbol{\xi} = \left(\xi_{1,1,1}, \cdots, \xi_{1,1,x_{1,1}}, \cdots, \xi_{s,n_s,1}, \cdots, \xi_{s,n_s,x_{s,n_s}}\right)$ be the fuzzy random vector involved in the fuzzy random parallel-series system, where lifetime $\xi_{i,j,k}$ can be any fuzzy random variable for $i = 1, 2, \cdots, s; j = 1, 2, \cdots, n_i; k = 1, 2, \cdots, x_{i,j}$. The discussion of computing the reliability in this section is divided into three cases: reliability with discrete lifetimes, reliability with convex lifetimes, and reliability with nonconvex (continuous) lifetimes.

4.1 Reliability with Discrete Lifetimes

$\boldsymbol{\xi}$ is a fuzzy random vector whose randomness is characterized by a discrete random vector ω assuming finite number of values ω_k, $k = 1, 2, \cdots, N$, with probability p_k, $k = 1, 2, \cdots, N$, respectively; and for each k, $\boldsymbol{\xi}(\omega_k) = \left(\xi_{1,1,1}(\omega_k), \cdots, \xi_{1,1,x_{1,1}}(\omega_k), \cdots, \xi_{s,n_s,1}(\omega_k), \cdots, \xi_{s,n_s,x_{s,n_s}}(\omega_k)\right)$ is a discrete fuzzy vector taking on M_k values

$$\widehat{\boldsymbol{\xi}}^{k,j} = \left(\hat{\xi}^{k,j}_{1,1,1}, \cdots, \hat{\xi}^{k,j}_{1,1,x_{11}}, \cdots, \hat{\xi}^{k,j}_{s,n_s,1}, \cdots, \hat{\xi}^{k,j}_{s,n_s,x_{s,n_s}}\right)$$

with possibility $\mu_k^j > 0$, $k = 1, 2, \cdots, N, j = 1, 2, \cdots, M_k$, and $\max_{1\le j\le M_k} \mu_k^j = 1$. In this case, the support of $\boldsymbol{\xi}$, $\Xi = \left\{\hat{\boldsymbol{\xi}}^{k,j} \mid k = 1, 2, \cdots, N, j = 1, 2, \cdots, M_k\right\}$, is a finite set. Hence from the definition, we have

$$R_{T^0}(\boldsymbol{x}) = \sum_{k=1}^{N} p_k Q_k(\boldsymbol{x}) \tag{18}$$

where $Q_k = \mathrm{Cr}\{T(\boldsymbol{x}, \boldsymbol{\xi}(\omega_k)) \ge T^0\}$ is calculated by

$$Q_k(\boldsymbol{x}) = \frac{1}{2}\left[\max\left\{\mu_k^j \,|\, T\left(\boldsymbol{x}, \widehat{\boldsymbol{\xi}}^{k,j}\right) \ge T^0\right\} + 1 - \max\left\{\mu_k^j \,|\, T\left(\boldsymbol{x}, \widehat{\boldsymbol{\xi}}^{k,j}\right) < T^0\right\}\right]. \tag{19}$$

For the simplicity, we abbreviate the formula (18)-(19) of the reliability with discrete lifetimes to RDL.

4.2 Reliability with Convex Lifetimes

In Subsection 3.2, we have proved Theorem 1 which supplies us with an analytical expression (10)-(12) of the reliability when all the lifetimes of components have convex distributions. By using Theorem 1, we can compute the reliability with convex lifetimes by the following random simulation method.

First of all, given a decision $\boldsymbol{x}$, for any $\omega \in \Omega$, we calculate $\mu_T^{\omega}(T^0)$ by formula (11)-(12). Furthermore, we note that formula (10) for reliability $R_{T^0}(\boldsymbol{x})$ with convex lifetimes can be rewritten as

$$R_{T^0}(\boldsymbol{x}) = E\left[I_{\{\omega | T^0 \le x_{s,1} v^{\omega}_{s,1}\}}(\omega)\left(1 - \frac{\mu_T^{\omega}(T^0)}{2}\right) + I_{\{\omega | T^0 > x_{s,1} v^{\omega}_{s,1}\}}(\omega)\left(\frac{\mu_T^{\omega}(T^0)}{2}\right)\right],$$

where $E[\cdot]$ is the expected value operator of random variable, I_A is the indicator function of set A. Therefore, making using of random simulation, we can compute the reliability by

$$R_{T^0}(\boldsymbol{x}) \leftarrow \frac{1}{M}\sum_{i=1}^{M} R_{T^0}(\boldsymbol{x}, \omega_i), \quad (M \to \infty) \tag{20}$$

where

$$R_{T^0}(\boldsymbol{x},\omega_i)=I_{\{\omega|T^0\leq x_{s,1}v^{\omega}_{s,1}\}}(\omega_i)\left(1-\frac{\mu^{\omega_i}_T(T^0)}{2}\right)+I_{\{\omega|T^0>x_{s,1}v^{\omega}_{s,1}\}}(\omega_i)\left(\frac{\mu^{\omega_i}_T(T^0)}{2}\right). \tag{21}$$

Here, ω_i for $i=1,2,\cdots,n$ are the random samples generated from the distribution of the random parameter involved in the fuzzy random vector $\boldsymbol{\xi}$. It is well known that the random simulation ([20]) is characterized by convergence with probability 1 as $M\to\infty$, which is ensured by the strong law of large numbers. The above computation procedure is summarized as the following algorithm.

Algorithm 1

Step 1. *Set* $R=0$.
Step 2. *Randomly generate a sample point* $\widehat{\omega}$ *from the distribution of the random vector involved in* $\boldsymbol{\xi}$.
Step 3. *Compute the* $R_{T^0}(\boldsymbol{x},\widehat{\omega})$ *through (21).*
Step 4. $R\leftarrow R+R_{T^0}(\boldsymbol{x},\widehat{\omega})$.
Step 5. *Repeat the Steps 2-4* M *times.*
Step 6. *Return the value of* $R_{T^0}(\boldsymbol{x})=R/M$.

4.3 Reliability with Nonconvex Lifetimes

In order to compute the system reliability with nonconvex lifetimes, in this subsection, we apply a fuzzy random simulation approach [23] to computing the reliability. Moreover, in order to attain the convergence, a discretization method [21] of continuous fuzzy random variable is embedded into the fuzzy random simulation. The convergence of the simulation approach for the $R_{T^0}(\boldsymbol{x})$ with nonconvex lifetimes is discussed to end this subsection.

Suppose the randomness of $\boldsymbol{\xi}$ is characterized by a continuous random vector, and for any random realization $\omega\in\Omega$, $\boldsymbol{\xi}(\omega)$ is a nonconvex continuous fuzzy vector with infinite support denoted by

$$\Xi=\prod_{k=1}^{K}[a_i,b_i] \tag{22}$$

where $K=\sum_{i=1}^{s}\sum_{j=1}^{n_i}x_{i,j}$, $[a_k,b_k]$ is the support of ξ_k for $k=1,2,\cdots,K$.

First of all, we employ the discretization method [21] to generate a sequence $\{\boldsymbol{\zeta}_l\}$ of discrete fuzzy random vectors which converges to the original continuous $\boldsymbol{\xi}$. For the simplicity, we denote the fuzzy random vector $\boldsymbol{\xi}=(\xi_1,\xi_2\cdots,\xi_K)$. For each integer $l=1,2,\cdots$, $\boldsymbol{\zeta}_l=(\zeta_{l,1},\zeta_{l,2},\cdots,\zeta_{l,K})$ is

constructed by the following method: define $\zeta_{l,i} = g_{l,i}(\xi_i)$ for $i = 1, 2, \cdots, K$, where the functions $g_{l,i}$'s are given by

$$g_{l,i}(v_i) = \begin{cases} a_i, & v_i \in [a_i, a_i + \frac{1}{l}) \\ \sup\left\{\frac{k_i}{l} \mid k_i \in Z, \text{ s.t. } \frac{k_i}{l} \le v_i\right\}, & v_i \in [a_i + \frac{1}{l}, b_i] \end{cases} \tag{23}$$

and Z is the set of integers. In what follows, the sequence $\{\boldsymbol{\zeta}_l\}$ of discrete fuzzy random vectors generated by (23) is referred to as the discretization of $\boldsymbol{\xi}$. It has been proved by [21] that

$$\|\boldsymbol{\zeta}_l(\omega)(\gamma) - \boldsymbol{\xi}(\omega)(\gamma)\| = \sqrt{\sum_{j=1}^{K} [\zeta_{l,i}(\omega)(\gamma) - \xi_i(\omega)(\gamma)]^2} \le \frac{\sqrt{K}}{l}, \tag{24}$$

for all $(\omega, \gamma) \in \Omega \times \Gamma$, which implies that the discretization $\{\boldsymbol{\zeta}_l\}$ converges to $\boldsymbol{\xi}$ uniformly.

Next, the fuzzy random simulation [23] is utilized to compute the $R_{T^0}(\boldsymbol{x})$. Noting that lifetime $T(\boldsymbol{x}, \boldsymbol{\xi})$ is a positive fuzzy random variable, the reliability function $R_{T^0}(\boldsymbol{x})$ can be rewritten as

$$\begin{aligned} R_{T^0}(\boldsymbol{x}) &= \mathrm{Ch}\{T(\boldsymbol{x}, \boldsymbol{\xi}) \ge T^0\} \\ &= \int_0^1 \Pr\{\omega \in \Omega \mid \mathrm{Cr}\{T(\boldsymbol{x}, \boldsymbol{\xi}(\omega)) \ge T^0\} \ge \alpha\} \mathrm{d}\alpha. \end{aligned}$$

For any $\omega \in \Omega$, we first replace the $\boldsymbol{\xi}$ with its discretization $\{\boldsymbol{\zeta}_l\}$ generated by (23), and estimate $\mathrm{Cr}\{T(\boldsymbol{x}, \boldsymbol{\xi}(\omega)) \ge T^0\}$ by $\mathrm{Cr}\{T(\boldsymbol{x}, \boldsymbol{\zeta}_l(\omega)) \ge T^0\}$ which can be calculated by (19). Furthermore, the value of

$$G(\boldsymbol{x}, U_i) = \Pr\left\{\omega \in \Omega \mid \mathrm{Cr}\left\{T\Big(\boldsymbol{x}, \boldsymbol{\zeta}_l(\omega)\Big) \ge T^0\right\} \ge U_i\right\}, \tag{25}$$

can be estimated with probability 1 by

$$G(\boldsymbol{x}, U_i) \leftarrow \frac{1}{n} \sum_{j=1}^{n} I_{\left\{\omega \mid \mathrm{Cr}\{T(\boldsymbol{x}, \boldsymbol{\zeta}_l(\omega)) \ge T^0\} \ge U_i\right\}}(\omega_j), \quad (n \to \infty), \tag{26}$$

where $U_i \sim \mathcal{U}(0, 1), i = 1, 2, \cdots$ are random variables with uniform distributions, and I_A is the indicator function of set A. Finally, the reliability $R_{T^0}(\boldsymbol{x})$ can be estimated by

$$R_{T^0}(\boldsymbol{x}) \leftarrow \frac{1}{n} \sum_{i=1}^{n} G(\boldsymbol{x}, U_i), \quad (n \to \infty) \tag{27}$$

with probability 1.

The fuzzy random simulation procedure for the $R_{T^0}(\boldsymbol{x})$ with nonconvex lifetimes is summarized as follows:

Algorithm 2 [Fuzzy Random Simulation]

Step 1. Generate $\boldsymbol{\zeta}_l(\omega)$ from the support Ξ of $\boldsymbol{\xi}(\omega)$ through (23), for any $\omega \in \Omega$.
Step 2. Calculate $\mathrm{Cr}\{T(\boldsymbol{x}, \boldsymbol{\zeta}_l(\omega)) \geq T^0\}$ through formula (19) for any $\omega \in \Omega$.
Step 3. Compute $G(\boldsymbol{x}, U_i)$ in (25) through random simulation (26) for each $U_i, i = 1, 2, \cdots, n$.
Step 4. Return the value of $R_{T^0}(\boldsymbol{x})$ by random simulation (27).

The following Theorem 5 shows that $R_{T^0,\boldsymbol{\zeta}_l}(\boldsymbol{x})$ converges to $R_{T^0}(\boldsymbol{x})$ for almost every $T^0 > 0$, as $l \to \infty$. As a consequence, the original reliability function $R_{T^0}(\boldsymbol{x})$ can be well approximated by $R_{T^0,\boldsymbol{\zeta}_l}(\boldsymbol{x})$ through Algorithm 2, provided l is sufficiently large.

Theorem 5. *Consider FR-RAMs I and II for a parallel-series system. Let $\boldsymbol{\xi}$ be the continuous fuzzy random lifetime vector of the components in the system, which has the compact interval support (22), $\{\boldsymbol{\zeta}_l\}$ be the discretization of $\boldsymbol{\xi}$, and T^0 the preselected threshold system lifetime. Then, for any feasible decision $\boldsymbol{x}$, the approximating system reliability function $R_{T^0,\boldsymbol{\zeta}_l}(\boldsymbol{x})$ converges to the original system reliability function, i.e.,*

$$\lim_{l\to\infty} R_{T^0,\boldsymbol{\zeta}_l}(\boldsymbol{x}) = R_{T^0}(\boldsymbol{x}),$$

provided $R_T(\boldsymbol{x})$ is continuous at $T = T^0$.

Proof. Recall that the lifetime of the parallel-series system is $T(\boldsymbol{x}, \boldsymbol{\xi}) = \bigvee_{i=1}^{s} \left[\bigwedge_{j=1}^{n_i} \left(\sum_{k=1}^{x_{i,j}} \xi_{i,j,k}\right)\right]$, which is a continuous function w.r.t. $\boldsymbol{\xi}$, for any given $\boldsymbol{x} = (x_{1,1}, \cdots, x_{1,n_1}, \cdots, x_{s,1}, \cdots, x_{s,n_s})$. Since the support $\Xi = \prod_{i=1}^{K}[a_i, b_i]$ of $\boldsymbol{\xi}$ is a compact set in $\Re^K$, where $K = \sum_{i=1}^{s}\sum_{j=1}^{n_i} x_{i,j}$, $T(\boldsymbol{x}, \boldsymbol{\xi})$ is uniformly continuous on Ξ. Hence, given a feasible $\boldsymbol{x}$, for any $\epsilon > 0$, there is a $\delta > 0$ such that

$$\left|T(\boldsymbol{x}, \widehat{\boldsymbol{\xi}}') - T(\boldsymbol{x}, \widehat{\boldsymbol{\xi}}'')\right| < \epsilon \tag{28}$$

whenever $\widehat{\boldsymbol{\xi}}', \widehat{\boldsymbol{\xi}}'' \in \Xi$, and

$$\left\|\widehat{\boldsymbol{\xi}}' - \widehat{\boldsymbol{\xi}}''\right\| = \sqrt{\sum_{i=1}^{s}\sum_{j=1}^{n_i}\sum_{k=1}^{xi,j}\left(\widehat{\xi}'_{i,j,k} - \widehat{\xi}''_{i,j,k}\right)^2} < \delta.$$

Noting that the discretization $\{\boldsymbol{\zeta}_l\}$ is a sequence of fuzzy random vectors which converges uniformly to $\boldsymbol{\xi}$ on $\Omega \times \Gamma$, for the above δ, there exists a positive integer L such that for all $(\omega, \gamma) \in \Omega \times \Gamma$,

$$\|\boldsymbol{\zeta}_l(\omega)(\gamma) - \boldsymbol{\xi}(\omega)(\gamma)\| = \sqrt{\sum_{i=1}^{s}\sum_{j=1}^{n_i}\sum_{k=1}^{x_{i,j}}\left(\zeta^l_{i,j,k}(\omega)(\gamma) - \xi_{i,j,k}(\omega)(\gamma)\right)^2} < \delta$$

provided $l \geq L$. Combining (28), for all $(\omega, \gamma) \in \Omega \times \Gamma$,

$$\left|T\Big(\boldsymbol{x},\boldsymbol{\zeta}_l(\omega)(\gamma)\Big)-T\Big(\boldsymbol{x},\boldsymbol{\xi}(\omega)(\gamma)\Big)\right|<\epsilon$$

whenever $l \geq L$. That is, the sequence $\{T(\boldsymbol{x},\boldsymbol{\zeta}_l)\}$ of fuzzy random variables converges uniformly to $T(\boldsymbol{x},\boldsymbol{\xi})$ on $\Omega \times \Gamma$. As a consequence, for any $\epsilon > 0$, we have

$$\lim_{l\to\infty} \mathrm{Ch}\Big\{\big|T(\boldsymbol{x},\boldsymbol{\zeta}_l)-T(\boldsymbol{x},\boldsymbol{\xi})\big|\geq\epsilon\Big\}=0.$$

Since convergence in chance implies convergence in distribution, we obtain

$$\lim_{l\to\infty} \mathrm{Ch}\big\{T(\boldsymbol{x},\boldsymbol{\zeta}_l)\geq T^0\big\}=\mathrm{Ch}\big\{T(\boldsymbol{x},\boldsymbol{\xi})\geq T^0\big\} \tag{29}$$

provided $\mathrm{Ch}\{T(\boldsymbol{x},\boldsymbol{\xi})\geq T\}$ is continuous at $T=T^0$. The proof of the theorem is complete.

5 The Algorithm

It is easy to see that the FR-RAMs I and II are tasks of fuzzy random integer programming problems. In this section, the three computation methods (formula RDL, Algorithm 1 and Algorithm 2) of the reliability $R_{T^0}(\boldsymbol{x})$ in different cases will be incorporated into the mechanism of genetic algorithm (GA) (see [7, 8, 18, 26]) to search for the approximately optimal solution the FR-RAMs I and II. In this hybrid algorithm, the GA is used to search for the best redundancy allocation and the computation methods is used to calculate the objective value of the FR-RAM I, and to check the feasibility of each chromosome in the FR-RAM II.

5.1 Solution Representation

A positive integer vector $\boldsymbol{C}=(C_1,C_2,\cdots,C_N)$ is used as a chromosome to represent a solution $\boldsymbol{x}=(x_{1,1},\cdots,x_{1,n_1},\cdots x_{s,1},\cdots,x_{s,n_s})$ of the FR-RAMs I and II, where $N=n_1+n_2+\cdots+n_s$.

5.2 Initialization Process

We first generate randomly an integer vector $\boldsymbol{C}=(C_1,C_2,\cdots,C_N)$ from a positive integer set $\{1,2,\cdots,K\}^N$, where K is a sufficiently large integer. If $\boldsymbol{C}$ is feasible, it is taken as an initial chromosome, otherwise, regenerate the vector $\boldsymbol{C}$ from $\{1,2,\cdots,K\}^N$ until the $\boldsymbol{C}$ is proved to be feasible.

Here, for the FR-RAM I, the feasibility of the chromosome $\boldsymbol{C}=(C_1,C_2,\cdots,C_N)$ is checked by

$$\sum_{i=1}^{s}\sum_{j=1}^{n_i} c_{ij}C_{ij}\leq c^0 \tag{30}$$

$$l_i \leq \sum_{j=1}^{n_i} C_{ij} \leq u_i, \text{for } i = 1, \cdots, s. \tag{31}$$

While in the FR-RAM II, to check the feasibility of the chromosome $\boldsymbol{C}$, we should compute the reliability $R_{T^0}(\boldsymbol{C})$ by the formula RDL, Algorithm 2, or Algorithm 3, according to the different distributions of lifetimes. Then, the following constraints are checked

$$R_{T^0}(\boldsymbol{C}) = \text{Ch}\{T(\boldsymbol{C}, \boldsymbol{\xi}) \geq T^0\} \geq R^0 \tag{32}$$

$$l_i \leq \sum_{j=1}^{n_i} C_{ij} \leq u_i, \text{for } i = 1, \cdots, s. \tag{33}$$

Repeating the above process *pop_size* times, we get *pop_size* initial chromosomes $\boldsymbol{C}_1$, $\boldsymbol{C}_2$, $\cdots$, $\boldsymbol{C}_{pop\_size}$.

5.3 Selection Process

The selection process is done based on *elitist strategy* and *spinning roulette wheel*. Before spinning the roulette wheel, we first calculate the objective function for each chromosome, i.e.,

$$R_{T^0}(\boldsymbol{C}) \text{ for FR-RAM I,}$$

and

$$\sum_{i=1}^{s} \sum_{j=1}^{n_i} c_{ij} C_{ij} \text{ for FR-RAM II,}$$

respectively, and the *pop_size* chromosomes are rearranged from good to bad based on the values of their objective functions. Here, the objective of the FR-RAM I is computed by using the formula RDL, Algorithm 1, or Algorithm 2 for the different cases of lifetimes, respectively.

In order to ensure that the best chromosome $\boldsymbol{C}_1$ of current population can always be selected successfully as an offspring, that is to move $\boldsymbol{C}_1$ directly into the next generation (elitist strategy). Then, we operate the selection process to the rest $pop\_size - 1$ chromosomes as follows: Employing the *evaluation function*, we assign a probability of reproduction to each chromosome $\boldsymbol{C}_k, k = 2, 3, \cdots, pop\_size$, so that the chromosome with the higher fitness will have more chance to be reproduced. There are several kinds of evaluation functions, here we adopt a popular one, *rank-based evaluation function*, which is defined as below:

$$eval(\boldsymbol{C}_k) = a(1-a)^{k-2}, k = 2, 3, \cdots, pop\_size,$$

where $a \in (0, 1)$ is a system parameter, and $k = 1$ means the best individual, while $k = pop\_size$ the worst one. Next, we calculate the cumulative

probability p_k for each chromosome $\boldsymbol{C}_k, k = 2, 3, \cdots, pop\_size$ as follows $p_1 = 0, p_k = eval(\boldsymbol{C}_2) + eval(\boldsymbol{C}_3) + \cdots + eval(\boldsymbol{C}_k)$, then normalize all $p'_k s$ dividing each $p_k, k = 2, 3, \cdots, pop\_size$ by p_{pop_size} such that $p_{pop\_size} = 1$. After that, generate a random number $r \in (0, 1]$, the probability of $p_{k-1} < r \leq p_k$ is the probability that the kth chromosome will be selected for the new population for $k = 2, 3, \cdots, pop\_size$. Repeating the following process $pop\_size - 1$ times, we can select $pop\_size - 1$ copies of chromosomes: generate a random number $r \in (0, 1]$, and select the kth chromosome $\boldsymbol{C}_k$ for $2 \leq k \leq pop\_size$ if $p_{k-1} < r \leq p_k$. Combining the previous $\boldsymbol{C}_1$, we obtain pop_size offspring.

5.4 Crossover Operation

In this process, a system parameter $p_c \in (0, 1)$ is predetermined as the probability of crossover. We repeat the following process pop_size times to determine the parents for the crossover operation: generate a random number r from interval $(0, 1]$, the chromosome $\boldsymbol{C}_k$ is selected as a parent for crossover provided $r < p_c$, where $k = 1, 2, \cdots, pop\_size$. Denote $\boldsymbol{C}'_1, \boldsymbol{C}'_2, \boldsymbol{C}'_3, \cdots$ the selected parents. They are divided into pairs: $(\boldsymbol{C}'_1, \boldsymbol{C}'_2), (\boldsymbol{C}'_3, \boldsymbol{C}'_4), (\boldsymbol{C}'_5, \boldsymbol{C}'_6), \cdots$. The crossover operation on each pair $(\boldsymbol{C}'_1, \boldsymbol{C}'_2)$ is done in the following way: Let

$$\boldsymbol{C}'_1 = \left(C_1^{(1)}, C_2^{(1)}, \cdots, C_N^{(1)}\right), \quad \boldsymbol{C}'_2 = \left(C_1^{(2)}, C_2^{(2)}, \cdots, C_N^{(2)}\right).$$

We randomly choose an integer N_c between 1 and N as the crossover point. Then, exchange the genes of the chromosomes $\boldsymbol{C}'_1$ and $\boldsymbol{C}'_2$ and produce two children as follows:

$$\boldsymbol{C}''_1 = \left(C_1^{(2)}, C_2^{(2)}, \cdots, C_{N_c-1}^{(2)}, C_{N_c}^{(2)}, C_{N_c+1}^{(1)}, \cdots, C_N^{(1)}\right)$$

$$\boldsymbol{C}''_2 = \left(C_1^{(1)}, C_2^{(1)}, \cdots, C_{N_c-1}^{(1)}, C_{N_c}^{(1)}, C_{N_c+1}^{(2)}, \cdots, C_N^{(2)}\right).$$

If both children are feasible, then the parents are replaced by them. Otherwise, keep the feasible one if exists, and then repeat the crossover process by generating a new crossover points until two feasible children are obtained.

5.5 Mutation Operation

Similar to the crossover operation, a parameter $p_m \in (0, 1)$ is predetermined as the probability of mutation. We repeat the following process pop_size times: randomly generate a real number r from $(0, 1]$, the chromosome $\boldsymbol{C}_k$ is selected as parents for mutation provided $r < p_m$, where $k = 1, 2, \cdots, pop\_size$. On each selected parent, denoted $\boldsymbol{C} = (C_1, C_2, \cdots, C_N)$, the mutation is done in the following way. We first randomly choose a mutation position N_m between 1

and N. Then, initialize $C'_1, C'_2, \cdots, C'_{N_m-1}, C'_{N_m}$ from integer set $\{1, 2, \cdots, K\}$, and produce a new chromosome

$$\boldsymbol{C}' = \left(C'_1, C'_2, \cdots, C'_{N_m-1}, C'_{N_m}, C_{N_m+1}, \cdots, C_N\right).$$

If $\boldsymbol{C}'$ is feasible for the constraints, then replace $\boldsymbol{C}$ with it. Otherwise, repeat this process until a feasible child is obtained.

5.6 Algorithm Procedure

A new population is produced after selection, crossover and mutation operation. The new cycles of evolution will continue until a given number of cyclic repetitions is met. The algorithm for solving the FR-RAMs I and II is summarized as follows.

Algorithm 3

Step 1. *Input the parameters: pop_size, p_c, p_m, and a.*
Step 2. *Initialize pop_size chromosomes from the positive integer set $\{1, 2, \cdots, K\}^N$. Here, the feasibility of the chromosomes in FR-RAM II is checked by the formula RDL, Algorithm 1, or Algorithm 2.*
Step 3. *Compute the objective values of all chromosomes. Here, for FR-RAM I, the objective values are computed by formula RDL, Algorithm 1, or Algorithm 2.*
Step 4. *Calculate the rank-based evaluation function for all the chromosomes according to their objective values.*
Step 5. *Select the chromosomes by spinning the roulette wheel with elitist strategy.*
Step 6. *Update the chromosomes by crossover and mutation operations. Again, for FR-RAM II, the feasibility of the chromosomes is checked by the formula RDL, Algorithm 1, or Algorithm 2.*
Step 7. *Repeat Step 3 to Step 6 for a given number of cycles;*
Step 8. *Return the best chromosome as the optimal solution.*

6 Numerical Examples

Two numerical examples are covered in this section to illustrate the modelling ideas of the FR-RAMs I and II, and the effectiveness of the designed solution algorithm. The numerical experiments are all performed on a personal computer, Intel(R) Core(TM) 2 Duo CPU 2.00 GHz, 1.0 GB memory. The parameters of GA include the population size *pop_size*, the probability of crossover p_c, the probability of mutation p_m, and the parameter a in the rank-based evaluation function.

Table 2 Lifetime and cost of each component in Example 4

Component ij	Cost c_{ij}	Lifetime ξ_{ij}	Random parameter Y_{ij}
11	10	$(2+Y_{11}, 3+Y_{11}, 5+Y_{11})$	$Y_{11} \sim \mathcal{U}(2,3)$
12	12	$(3+Y_{12}, 4+Y_{12}, 6+Y_{12})$	$Y_{12} \sim \mathcal{U}(3,5)$
13	14	$(4+Y_{13}, 5+Y_{13}, 6+Y_{13})$	$Y_{13} \sim \mathcal{U}(1,3)$
21	10	$(1+Y_{21}, 3+Y_{21}, 4+Y_{21})$	$Y_{21} \sim \mathcal{U}(2,4)$
22	12	$(2+Y_{22}, 4+Y_{22}, 5+Y_{22})$	$Y_{22} \sim \mathcal{U}(1,3)$
31	16	$(4+Y_{31}, 6+Y_{31}, 8+Y_{31})$	$Y_{31} \sim \mathcal{U}(0,2)$
32	11	$(3+Y_{32}, 4+Y_{32}, 5+Y_{32})$	$Y_{32} \sim \mathcal{U}(1,3)$
33	14	$(4+Y_{33}, 5+Y_{33}, 6+Y_{33})$	$Y_{33} \sim \mathcal{U}(2,3)$

Example 5. Consider a 3-stage parallel-series system, where there are 3 types of components in the first subsystem, 2 types in the second and 3 types in the third. The redundancy allocation decision vector is

$$\boldsymbol{x} = (x_{1,1}, x_{1,2}, x_{1,3}, x_{2,1}, x_{2,2}, x_{3,1}, x_{3,2}, x_{3,3}),$$

the vector of fuzzy random lifetimes is

$$\boldsymbol{\xi} = (\xi_{1,1,1}, \cdots, \xi_{1,1,x_{1,1}}, \cdots, \xi_{2,1,1}, \cdots, \xi_{2,1,x_{2,1}}, \cdots, \xi_{3,3,1}, \cdots, \xi_{3,3,x_{3,3}})$$

in which fuzzy random lifetimes together with the cost of each component is given in Table 2 (we use $\xi_{i,j}$ and c_{ij} to represent the distribution and the cost of all lifetimes $\xi_{i,j,k}$ for $k = 1, 2, \cdots, x_{i,j}$, since they are the same type of components). The threshold system lifetime is $T^0 = 6$. The total available capital is $c^0 = 300$. For each subsystem $i, i = 1, 2, 3$, the lower and upper bounds of the number of the redundant components are given as $l_1 = 3, u_1 = 9; l_2 = 2, u_2 = 6; l_3 = 4, u_3 = 11$, respectively. Maximizing the reliability and making use of FR-RAM I, we can build a redundancy allocation model for this system as follows:

$$\max R_6(\boldsymbol{x}) = \mathrm{Ch}\{T(\boldsymbol{x}, \boldsymbol{\xi}) \geq 6\}$$

subject to

$$\begin{aligned}
&11x_{1,1} + 12x_{1,2} + 14x_{1,3} + 10x_{2,1} + 12x_{2,2} + 16x_{3,1} + 11x_{3,2} + 14x_{3,3} \leq 300,\\
&3 \leq x_{1,1} + x_{1,2} + x_{1,3} \leq 9,\\
&2 \leq x_{2,1} + x_{2,2} \leq 6,\\
&4 \leq x_{3,1} + x_{3,2} + x_{3,3} \leq 11,\\
&x_{i,j} \in \mathbb{N}, \text{for } j = 1, \cdots, n_i, i = 1, \cdots, 3,
\end{aligned} \tag{34}$$

where the system lifetime

$$T(\boldsymbol{x}, \boldsymbol{\xi}) = \bigvee_{i=1}^{3} \left[\bigwedge_{j=1}^{n_i} \left(\sum_{k=1}^{x_{i,j}} \xi_{i,j,k} \right) \right]. \tag{35}$$

Noting that all the lifetimes of components have the convex distributions, therefore, at each given allocation decision $\boldsymbol{x} = (x_{1,1}, x_{1,2}, x_{1,3}, x_{2,1}, x_{2,2}, x_{3,1}, x_{3,2}, x_{3,3})$, the system reliability

$$R_6(\boldsymbol{x}) = \mathrm{Ch}\left\{T(\boldsymbol{x}, \boldsymbol{\xi}) \geq 6\right\}, \tag{36}$$

can be calculated by the Algorithm 1. Incorporating the Algorithm 1 into the GA, we use Algorithm 3 to search for the optimal solution of problem (34).

Table 3 The parameters and comparison solutions of Example 6

pop_size	p_c	p_m	a	Gen	Optimal solution	Objective value	Error(%)
20	0.2	0.2	0.05	400	(2,2,4,2,2,2,1,3)	0.8598	1.54
20	0.3	0.2	0.10	400	(2,2,2,2,2,2,2,3)	0.8588	1.66
20	0.3	0.2	0.05	400	(2,2,4,2,1,3,3,2)	0.8575	1.80
20	0.3	0.1	0.10	400	(2,2,2,1,1,2,3,3)	0.8627	1.17
20	0.3	0.1	0.05	400	(2,2,5,2,1,3,2,3)	0.8571	1.85
30	0.2	0.2	0.05	400	(2,2,5,1,1,2,2,2)	0.8675	0.66
30	0.3	0.2	0.10	400	(2,2,3,2,2,2,2,2)	0.8618	1.31
30	0.3	0.2	0.05	400	(1,1,1,3,3,1,1,2)	0.8733	0.00
30	0.3	0.1	0.10	400	(2,2,2,2,1,2,2,2)	0.8665	0.78
30	0.3	0.1	0.05	400	(2,2,5,1,1,3,3,2)	0.8607	1.44

The Algorithm 3 has been run with 400 generations in GA and 6,000 times of random simulation (20) in Algorithm 1, and in Table 3 we compare solutions by careful variations of parameters of GA with the same stopping rule. The parameters are given in Table 3 from the first to the fourth column, and the computational results are provided in fifth and sixth columns. In addition, the *relative error* is given in the last column, which is defined by *(objective value-optimal value)/optimal value* $\times$ *100%*. It follows from Table 3 that the relative error does not exceed 1.85% when different parameters of GA are selected. In addition, the convergence of the objective value (system reliability) is shown in Figure 2. The performance implies the solution algorithm is robust to the parameter settings and effective to solve the FR-RAM I.

Example 6. For the fuzzy random parallel-series system in Example 6, if the decision-maker intends to minimize the total cost to meet some reliability constraint, i.e., $R_{T^0}(\boldsymbol{x}) \geq \alpha^0$, then we can model this problem by FR-RAM II. Here, we suppose the lifetimes of components have the same costs with

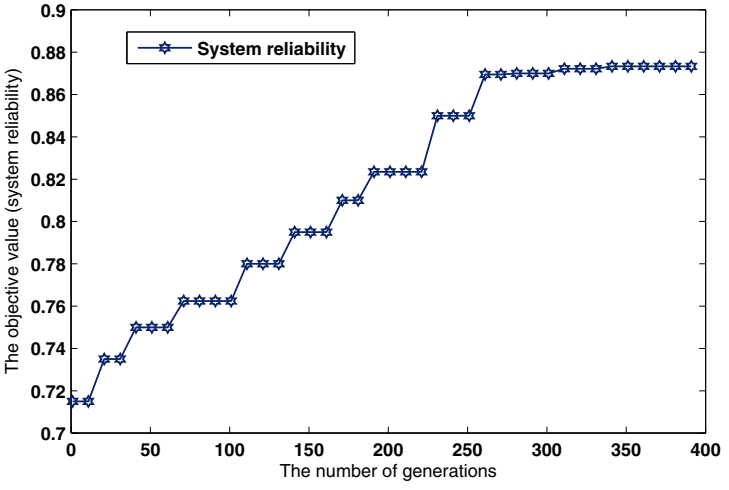


Fig. 2 The convergence of the objective value of Example 4

Table 4 Lifetime and cost of each component in Example 5

Component ij	Cost c_{ij}	Lifetime ξ_{ij}	Random parameter Y_{ij}
11	10	$(2+Y_{11}, 3+Y_{11}, 5+Y_{11})$	$Y_{11} \sim \mathcal{U}(2,3)$
12	12	$\mathcal{N}_{\mathcal{F}}^{+}(Y_{12}, 4)$	$Y_{12} \sim \mathcal{U}(2,4)$
13	14	$\perp(4+Y_{13}, 5+Y_{13}, 6+Y_{13})$	$Y_{13} \sim \mathcal{U}(1,3)$
21	10	$\perp(1+Y_{21}, 3+Y_{21}, 4+Y_{21})$	$Y_{21} \sim \mathcal{U}(1,2)$
22	12	$(2+Y_{22}, 4+Y_{22}, 5+Y_{22})$	$Y_{22} \sim \mathcal{U}(1,3)$
31	16	$\mathcal{N}_{\mathcal{F}}^{+}(Y_{12}, 6)$	$Y_{31} \sim \mathcal{U}(1,3)$
32	11	$\perp(3+Y_{32}, 4+Y_{32}, 5+Y_{32})$	$Y_{32} \sim \mathcal{U}(2,4)$
33	14	$(4+Y_{33}, 5+Y_{33}, 6+Y_{33})$	$Y_{33} \sim \mathcal{U}(2,3)$

that in Example 6 but different distributions partly as listed in Table 4, where $\mathcal{N}_{\mathcal{F}}^{+}(Y_{12}, 4)$ is a positive normal fuzzy random variable defined in (4), and $\perp(4+Y_{13}, 5+Y_{13}, 6+Y_{13})$ is an inverse triangular fuzzy random variable, in which the membership function of $\perp(a, b, c)$ is given by

$$\mu_{\perp(a,b,c)}(r) = \begin{cases} (b-r)/(b-a), & a \le r < b \\ (x-b)/(c-b), & b \le r \le c \\ 0, & \text{otherwise.} \end{cases}$$

Taking the target reliability $R^0 = 0.8$ and the threshold lifetime $T^0 = 6$, and the same lower and upper bounds of numbers of the redundant components as in Example 6, a cost minimization based FR-RAM can be formed by

$$\min\ 11x_{1,1} + 12x_{1,2} + 14x_{1,3} + 10x_{2,1} + 12x_{2,2} + 16x_{3,1} + 11x_{3,2} + 14x_{3,3}$$

subject to

$$\begin{aligned} & R_6(\boldsymbol{x}) = \mathrm{Ch}\left\{T(\boldsymbol{x}, \boldsymbol{\xi}) \geq 6\right\} \geq 0.8, \\ & 3 \leq x_{1,1} + x_{1,2} + x_{1,3} \leq 9, \\ & 2 \leq x_{2,1} + x_{2,2} \leq 6, \\ & 4 \leq x_{3,1} + x_{3,2} + x_{3,3} \leq 11, \\ & x_{i,j} \in \mathbb{N}, \text{for } j = 1, \cdots, n_i, i = 1, \cdots, 3. \end{aligned} \tag{37}$$

When using the Algorithm 3 (for FR-RAM II) to solve the problem (37), in the processes of initialization, crossover and mutation, we need to check the feasibility of each chromosome, which means to compute $R_6(\boldsymbol{x}) = \mathrm{Ch}\left\{T(\boldsymbol{x}, \boldsymbol{\xi}) \geq 6\right\}$ in each checking.

We note from Table 4 that the lifetimes $\perp(4+Y_{13}, 5+Y_{13}, 6+Y_{13})$ in subsystem 1, $\perp(1+Y_{21}, 3+Y_{21}, 4+Y_{21})$ in subsystem 2 and $\perp(3+Y_{32}, 4+Y_{32}, 5+Y_{32})$ in subsystem 3 are nonconvex fuzzy random variables, the Algorithm 1 therefore is not suitable for computing the system reliability $R_6(\boldsymbol{x})$. As a consequence, we use the fuzzy random simulation (Algorithm 2) to compute $R_6(\boldsymbol{x})$. That is, for any realization $\widehat{\boldsymbol{Y}} = \left(\hat{Y}_{11}, \hat{Y}_{12}, \hat{Y}_{13}, \hat{Y}_{21}, \hat{Y}_{22}, \hat{Y}_{31}, \hat{Y}_{32}, \hat{Y}_{33}\right)$ of random vector $\boldsymbol{Y} = (Y_{11}, Y_{12}, Y_{13}, Y_{21}, Y_{22}, Y_{31}, Y_{32}, Y_{33})$, we generate the discretization $\boldsymbol{\zeta}_l = \left(\zeta^l_{1,1,1}, \cdots, \zeta^l_{1,3,x_{1,3}}, \cdots, \zeta^3_{3,3,x_{3,3}}\right)$ of $\boldsymbol{\xi}$ by (23), where l is taken as 5,000. Noting from (37) that

$$x_{1,1} + x_{1,2} + x_{1,3} \leq 9, x_{2,1} + x_{2,2} \leq 6, x_{3,1} + x_{3,2} + x_{3,3} \leq 11,$$

hence, by (24) the discretization error $\mathcal{E}$ therefore can be controlled and

$$\mathcal{E} \leq \sqrt{\sum_{j=1}^{3} x_{1,j} + \sum_{j=1}^{2} x_{2,j} + \sum_{j=1}^{3} x_{3,j}} \Big/ 5,000 \leq \frac{\sqrt{26}}{5,000} \approx 0.001.$$

After that, we calculate

$$\mathrm{Cr}\left\{\bigvee_{i=1}^{3}\left[\bigwedge_{j=1}^{n_i}\left(\sum_{k=1}^{x_{ij}} \zeta^l_{i,j,k}\left(\hat{Y}_{ij}\right)\right)\right] \geq 6\right\}$$

through formula (19) for all $\hat{Y}_{ij}$, and can obtain $R_6(\boldsymbol{x})$ by the random simulation (26)-(27).

The above computation (Algorithm 2) is embedded into GA to search for the best solution of the problem (37). We run the Algorithm 3 with 6,000 times of random simulation (26)-(27) in Algorithm 2, and 400 generations in GA, the comparison of solutions with different parameters are collected in Table 5, and the convergence of the objective value (total cost) is shown in Figure 3. We see from Table 5 that the relative error dose not exceed 4.43%

Table 5 The parameters and comparison solutions of Example 5

pop_size	p_c	p_m	a	Gen	Optimal solution	Objective value	Error(%)
20	0.2	0.2	0.05	400	(1,2,2,1,1,2,3,3)	163	3.16
20	0.3	0.2	0.10	400	(1,3,1,2,1,2,2,2)	160	1.27
20	0.3	0.2	0.05	400	(1,3,1,1,1,2,2,2)	164	3.80
20	0.3	0.1	0.10	400	(1,1,2,1,1,2,3,2)	165	4.43
20	0.3	0.1	0.05	400	(1,1,1,1,1,2,4,2)	162	2.53
30	0.2	0.2	0.05	400	(1,1,1,1,2,2,3,2)	163	3.16
30	0.3	0.2	0.10	400	(1,1,2,1,1,1,5,1)	158	0.00
30	0.3	0.2	0.05	400	(1,1,1,1,2,2,4,1)	160	1.27
30	0.3	0.1	0.10	400	(2,1,1,1,1,1,3,3)	161	1.90
30	0.3	0.1	0.05	400	(1,2,2,1,2,2,2,1)	164	3.80

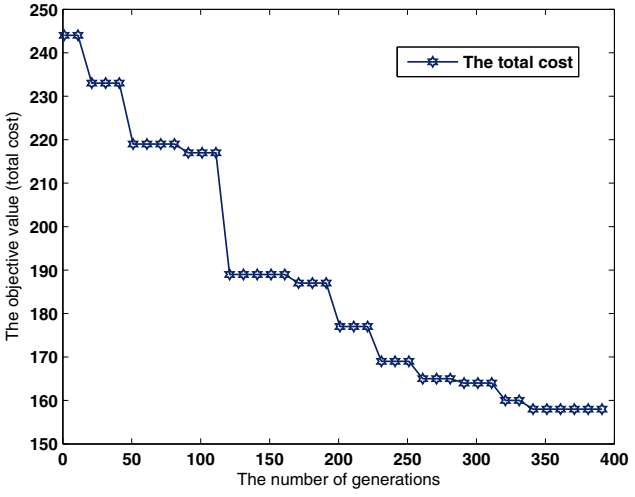


Fig. 3 The convergence of the objective value of Example 5

which shows that Algorithm 3 is also robust to the parameter settings and effective to solve the FR-RAM II.

7 Concluding Remarks

In this work, by considering a parallel-series system with fuzzy random lifetimes, we developed two redundancy allocation models (FR-RAMs I and II) through reliability maximization and cost minimization, respectively. Some properties on FR-RAM were discussed, where an analytical formula of reliability with convex lifetimes was derived (Theorem 1), and the sensitivity

of the reliability with respect to the threshold lifetime was studied (Theorems 2-4).

The FR-RAMs are fuzzy random integer programming tasks, where the objective function of the FR-RAM I and the constraints of FR-RAM II contain a fuzzy random reliability function, respectively. Since the reliability with fuzzy random parameters in general cannot be calculated directly, the classical mathematical programming methods are not applicable to the FR-RAMs. Thus, to solve the FR-RAMs, we dealt with the following two issues.

- The computation of reliability: Based on the analytical formulation (Theorem 1), we proposed a random simulation method (Algorithm 1) to compute the reliability with convex lifetimes. Furthermore, we computed the reliability with nonconvex lifetimes by combining the fuzzy random simulation with the discretization method (Algorithm 2), and the convergence of the fuzzy random simulation was proved (Theorem 5).
- The solution algorithm: By incorporating three different computation methods for reliability into GA, a hybrid solution algorithm (Algorithm 3) was produced to solve the FR-RAMs. Two numerical experiments were provided to illustrate the performance of the solution algorithm.

Nevertheless, there is much room for further development of our research. For instance, although the convergence of the fuzzy random simulation for the reliability with nonconvex lifetimes was proved in this paper, such a bifold simulation is a time consuming process, since it requires the generation of a sufficiently large number of discrete fuzzy random variables to ensure the precision of the simulation. Therefore, a method to speed up the computation for this case is a significant topic for future studies. In addition, this paper only modeled single-objective redundancy allocation problems under fuzzy random environment, while multi-objective FR-RAMs have not been considered, which should be another interesting topic that needs further investigation.

Acknowledgements. This work was supported by the Research Fellowships of the Japan Society for the Promotion of Science (JSPS) for Young Scientists, and by "Ambient SoC Global COE Program of Waseda University" of Ministry of Education, Culture, Sports, Science and Technology, Japan.

References

1. Barlow, R.E., Proschan, F.: Statistical Theory of Reliability and Life Testing. Holt, Rinehart (1975)
2. Bag, S., Chakraborty, D., Roy, A.R.: A production inventory model with fuzzy random demand and with flexibility and reliability considerations. Computers & Industrial Engineering 56, 411–416 (2009)
3. Das, K., Roy, T.K., Maiti, M.: Multi-item stochastic and fuzzy-stochastic inventory models under two restrictions. Computers & Operations Research 31, 1793–1806 (2004)

4. Dutta, P., Chakraborty, D., Roy, A.R.: A single-period inventory model with fuzzy random variable demand. Mathematical and Computer Modelling 41, 915–922 (2005)
5. Elegbede, A.O.C., Chu, C., Adjallah, K.H., Yalaoui, F.: Reliability allocation through cost minimization. IEEE Transactions on Reliability 52, 106–111 (2003)
6. Fishman, G.S.: Monte Carlo: Concepts, Algorithms, and Applications. Springer, New York (1996)
7. Goldberg, D.E.: Genetic Algorithms in Search, Optimization and Machine Learning. Addison-Wesley, Reading (1989)
8. Holland, J.: Adaptation in Natural and Artifical Systems. University of Michigan Press, Ann Arbor (1975)
9. Huang, X.: A new perspective for optimal portfolio selection with random fuzzy returns. Information Sciences 177, 5404–5414 (2007)
10. Huang, T., Zhao, R., Tang, W.: Risk model with fuzzy random individual claim amount. European Journal of Operational Research 192, 879–890 (2009)
11. Karnik, N.N., Mendel, J.M.: Operations on type-2 fuzzy sets. Fuzzy Sets and Systems 122, 327–348 (2001)
12. Kruse, R., Meyer, K.D.: Statistics with Vague Data. D. Reidel Publishing Company, Dordrecht (1987)
13. Kuo, W., Prasad, V.R., Tillman, F.A., Hwang, C.L.: Optimal Reliability Design. Cambridge University Press, Cambridge (2001)
14. Kuo, W., Wan, R.: Recent Advances in Optimal Reliability Allocation. IEEE Transactions on Systems, Man, and Cybernetics – Part A: Systems and Humans 38, 143–156 (2008)
15. Kwakernaak, H.: Fuzzy random variables–I. Definitions and theorems. Information Sciences 15, 1–29 (1978)
16. Kwakernaak, H.: Fuzzy random variables–II. Algorithm and examples. Information Sciences 17, 253–278 (1979)
17. Li, B., Zhu, M., Xu, K.: A practical engineering method for fuzzy reliability analysis of mechanical structures. Reliability Engineering & System Safety 67, 311–315 (2000)
18. Liu, B.: Theory and Practice of Uncertain Programming. Physica-Verlag, Heidelberg (2002)
19. Liu, B.: Uncertainty Theory, 2nd edn. Springer, Berlin (2007)
20. Liu, B., Liu, Y.K.: Expected value of fuzzy variable and fuzzy expected value models. IEEE Transaction on Fuzzy Systems 10, 445–450 (2002)
21. Liu, Y.K.: The approximation method for two-stage fuzzy random programming with recourse. IEEE Transactions on Fuzzy Systems 15, 1197–1208 (2007)
22. Liu, Y.K., Liu, B.: Fuzzy random variable: A scalar expected value operator. Fuzzy Optimization and Decision Making 2, 143–160 (2003)
23. Liu, Y.K., Liu, B.: On minimum-risk problems in fuzzy random decision systems. Computers & Operations Research 32, 257–283 (2005)
24. López-Diaz, M., Gil, M.A.: Constructive definitions of fuzzy random variables. Statistics and Probability Letters 36, 135–143 (1997)
25. Mahapatra, G.S., Roy, T.K.: Fuzzy multi-objective mathematical programming on reliability optimization model. Applied Mathematics and Computation 174, 643–659 (2006)
26. Michalewicz, Z.: Genetic Algorithms + Data Stractures = Evolution Programs. Springer, New York (1994)

27. Nahmias, S.: Fuzzy variable. Fuzzy Sets and Systems 1, 97–101 (1978)
28. Pedrycz, W., Gomide, F.: An Introduction to Fuzzy Sets, Analysis and Design. MIT Press, Cambridge (1998)
29. Pedrycz, W., Gomide, F.: Fuzzy Systems Engineering: Toward Human-Centric Computing. J. Wiley, Hoboken (2007)
30. Prasad, V.R., Raghavachari, M.: Optimal allocation of interchangeable components in a series-parallel system. IEEE Transactions on Reliability 47, 255–260 (1998)
31. Rubinstein, R.Y.: Simulation and the Monte Carlo Method. Wiley, New York (1981)
32. Smimou, K., Bector, C.R., Jacoby, G.: Portfolio selection subject to experts' judgments. International Review of Financial Analysis 117, 1036–1054 (2008)
33. Tavakkoli-Moghaddam, R., Safari, J., Sassani, F.: Reliability optimization of series-parallel system with a choice of redundancy strategies using a genetic algorithm. Reliability Engineering & System Safety 93, 550–556 (2008)
34. Wang, S., Liu, Y., Dai, X.: On the continuity and absolute continuity of credibility functions. Journal of Uncertain Systems 1, 185–200 (2007)
35. Wang, S., Watada, J.: T-independence condition for fuzzy random vector based on continuous triangular norms. Journal of Uncertain Systems 2, 155–160 (2008)
36. Wang, S., Watada, J.: Studying distribution functions of fuzzy random variables and its applications to critical value functions. International Journal of Innovative Computing, Information & Control 5, 279–292 (2009)
37. Wang, S., Liu, Y.K., Watada, J.: Fuzzy random renewal process with queueing applications. Computers & Mathematics with Applications 57, 1232–1248 (2009)
38. Watada, J., Wang, S.: Regression model based on fuzzy random variables. In: Rodulf, S. (ed.) Views on Fuzzy Sets and Systems from Different Perspectives, ch. 26. Spring, Berlin (2009)
39. Watada, J., Wang, S., Pedrycz, W.: Building confidence-interval-based fuzzy random regression models. IEEE Transactions on Fuzzy Systems (to be published)
40. Xu, J., Liu, Y.G.: Multi-objective decision making model under fuzzy random environment and its application to inventory problems. Information Sciences 178, 2899–2914 (2008)
41. Yu, H., Chu, C., Chatelet, E., Yalaoui, F.: Reliability optimization of a redunant system with failure dependencies. Reliability Engineering & System Safety 92, 1627–1634 (2007)
42. Zadeh, L.A.: Fuzzy sets. Information and Control 8, 338–353 (1965)
43. Zadeh, L.A.: Fuzzy sets as a basis for a theory of possibility. Fuzzy Sets and Systems 1, 3–28 (1978)
44. Zadeh, L.A.: Toward a generalized theory of uncertainty (GTU) – an outline. Information Sciences 172, 1–40 (2005)
45. Zadeh, L.A.: Generalized theory of uncertainty (GTU) – principal concepts and ideas. Computational Statistics & Data Analysis 51, 15–46 (2006)
46. Zhao, R., Liu, B.: Redundancy optimization problems with uncertainty of combining randomness and fuzziness. European Journal of Operational Research 157, 716–735 (2004)
47. Zhao, R., Liu, B.: Standby redundancy optimization problems with fuzzy lifetimes. Computers & Industrial Engineering 49, 318–338 (2005)
48. Zhao, R., Tang, W.S.: Some properties of fuzzy random renewal processes. IEEE Transactions on Fuzzy Systems 14, 173–179 (2006)

Reliable Biological Circuit Design Including Uncertain Kinetic Parameters

Eva Sciacca and Salvatore Spinella

Abstract. In the context of possibilistic decision making, this work deals with biological design problems particularly important in the near future when it will be possible to produce biological entities and synthetic organisms for pharmacological and medical usage. The biological systems is investigated in terms of performances or main key features of the system. The analysis of the biological system is based on the idea that the set of parameters involved in the model can be classified into two different typologies: the uncertain kinetic parameters and the control design parameters. In order to design a robust and reliable biological system with respect to a target performance, the design parameter values are set up to balance the uncertainty of the kinetic parameters. To take into account these uncertainties arising from the estimations of the kinetic parameters, the function representing the feedback of the system is fuzzified and a measure of failure of the designed biological circuit is minimized to reach the required performance. An application of this methodology is illustrated on a case study of an autonomously oscillatory system: the Drosophila Period Protein which is a central component of the Drosophila circadian clocks. Finally, the results of the fuzzy methodology are compared with a deterministic method.

1 Introduction

In the context of possibilistic decision making, this work deals with biological design problems particularly important in the near future when it will be possible to produce biological entities and synthetic organisms for pharmacological and medical usage. Generally, the design frameworks that support Research & Development

Eva Sciacca · Salvatore Spinella
Department of Mathematics and Computer Science, University of Catania, Viale A. Doria 6, I 95125 Catania, Italy
e-mail: sciacca@dmi.unict.it, spins@unical.it

W.A. Lodwick & J. Kacprzyk (Eds.): Fuzzy Optimization, STUDFUZZ 254, pp. 457–477.
springerlink.com

technologies must face with uncertainty modelling. In this field the uncertainty arise from different sources related to:

- behavioral models, that connect uncertain model parameters to observed evolution of system state;
- equivalent models, that connect uncertain model parameters to system feedbacks;
- approximated models, that approximate various aspects of a system in a computational tractable manner.

One possible way to model biological networks is to employ Ordinary Differential Equations (ODE) systems. These models involve several parameters such as reaction rate constants or protein concentrations. A precise measurement of these parameters is difficult to experimentally estimate and their values are quite uncertain because they vary from one cell to another or throughout the lifetime of any individual.

The aim of this work is to set up a methodology to perform a reliable target satisfaction in the framework of biological systems design that include uncertain parameters. Biological pathways modeled by mass action reactions include uncertain parameters that are usually related to the equilibrium of the system; feedback loops between tumor suppressors and oncogene agents are some examples of these kinds of systems.

This chapter includes the following sections. A general survey of optimization in biological design is outlined in Sect. 2. In Sect. 3 the design optimization problem under uncertainty is formulated. The biological system was investigated in terms of performances or main key features of the model. The analysis of the biological system is based on the idea that the set of parameters involved in the model can be classified into different typologies, two different classes of parameters were identified: the uncertain kinetic parameters and the control design parameters. In order to design a robust and reliable biological system with respect to a target performance, the design parameter values were set up to balance the uncertainty of the kinetic parameters. To take into account these uncertainties arising from the estimations of the kinetic parameters, the function representing the feedback of the system was fuzzified as described in Sect. 4. One way to deal with design specifications is to compare the fuzzified performance with a crisp number representing a reasonable threshold and give a measure of satisfaction of these constraints. For this purpose the possibility measures of failure with respect to the specification constraints can give a useful information to improve the design. The measures of failure are then minimized using an appropriate optimization algorithm to reach the required biological design. Sect. 5 gives a description of the main concepts of the possibility theory. A summary of the whole methodology and a clear diagram of the numerical algorithms and components involved in the design simulation are illustrated in Sect. 6. Among all the possible biological circuit designs, autonomously oscillating systems are the most investigated because they provide invaluable resources for controlling and orchestrating the main biological functions. These kinds of systems are introduced in Sect. 7 where also a summary of the typical biochemical rate functions involved in biological mathematical modelling is given. The particular case

study of the Drosophila Period Protein (PER) and the involved model parameters are described in Sect. 8. Some results of the application of the methodology and the comparison with a deterministic method on the case study are illustrated in Sect. 9. Finally, the conclusions on this study and some perspectives for future research are outlined in Sect. 10.

2 Biological Design Optimization

The concept of optimization is certainly not new in biology [Banga, 2008]. Model-based optimization is a key methodology in engineering, helping in the design, analysis, construction and operation of all kind of devices. Optimization methods have been applied in both metabolic control analysis [Heinrich and Schuster, 1998] and biochemical systems theory [Torres and Voit, 2002]. Further, optimization (in particular, linear programming) has been the engine behind metabolic flux balance analysis, where the optimal flux distributions are calculated using linear optimization, and are used to represent the metabolic phenotype for certain conditions. This flux balance methodology provides a guide to metabolic engineering and a method for bio-process optimization [Banga, 2008].

Coupling constraint-based analysis with optimization has been used to generate a consistent framework for the generation of hypotheses and the testing of functions of microbial cells using genome-scale models. Extensions and modifications of flux balance analysis continue to use optimization methods extensively [Segre *et al*., 2002]. Constrained evolutionary optimization has been used to understand optimal circuit design. Moreover, optimization principles have also been used to explain the complexity and robustness found in biochemical networks [Stelling *et al*., 2004]. Reverse engineering in systems biology aims to reconstruct the biochemical interactions from data sets of a particular biological system. Optimization has been used for inferring important biomolecular networks, such as transcriptional regulatory networks [Wang *et al*., 2007], gene regulatory networks, signaling pathways and protein interaction networks.

System identification is a methodology widely used in engineering for building mathematical models of dynamical systems based on measured data. This methodology involves the selection of the model structure and the parameter estimation for the model from the available experimental data. The parameter estimation problem in biochemical pathways is formulated as a nonlinear programming problem subject to the pathway model acting as constraints. Since these problems are frequently multi-modal, global optimization methods are needed in order to avoid local solutions. A local solution can be very misleading when calibrating models: it would indicate a bad fit even for a model which could potentially match perfectly a set of experimental data [Banga, 2008].

In general, for biological design optimization, it is assumed that the mathematical model and the involved species are already chosen by the biologist and are fixed. Let r be the number of specifications of the biological system to be optimized, the

desired response $R^* \in \mathbb{R}^r$ is expressed in terms of design specifications or design goals. The problem of biological model design then can be formulated as:

$$x^* = \arg\min_{x \in X} U(R(x)) \tag{1}$$

where x^* is the optimal design, X is the feasible region, U is a suitable objective function, and hopefully $R(x^*) = R^*$. In general, the above problem corresponds to a constrained nonlinear programming problem. The objective function U is typically a combination of multiple objectives with conflicting criteria.

Generally, "classical methods to solve equation 1 include Line Search and Trust Region strategies, based on methods such as Conjugate Gradient, Newton and Quasi-Newton methods. Usually methods that use only function evaluations are more suitable for problems that are very nonlinear or have many discontinuities (Search Methods) while methods that use derivative information are more effective when the function is continuous in the first and second derivatives (Gradient Methods).

For nominal design it is assumed that the design parameters are not subject to statistical fluctuations. When uncertain parameters are considered, one of the most common deterministic design methodology, used also in microelectronic industry, is named "Nominal Over-Design". The Nominal Over-Design fixes every objective to a secure value with regard to the nominal target specifications. The design specifications are increased of a certain percentage in the case of minimum thresholds, while in cases of maximum thresholds they are decreased. The main drawback of this scheme is often to point deterministically at unfeasible over-designs which could have the opposite effect blocking the optimization process at initial stages.

The methodology introduced in this study allows to combine optimization and uncertainty analysis for target satisfaction design problem. The treatment of uncertainty is an unavoidable step because the lack of precision in the model and among its parameter values can invalidate the results. The target satisfaction under uncertainty gives a robust framework to use optimization in a design context.

3 Biological Design Optimization under Uncertainty

In this work a methodology was formulated to perform a target satisfaction in the framework of biological systems design that incorporates reliable feedback. The proposed methodology is based on the scheme described in [Sciacca *et al.*, 2007] which was applied to a sizing problem of an electronic circuit. The biological system was investigated in terms of performances or main key features of the system. For example, in autonomously oscillating biochemical systems, the key features of oscillatory trajectories of the species concentrations can be taken into consideration. Once the performances of the system are chosen, the system can be designed in a reliable manner with respect to the uncertainties related to the estimation of the parameters. Our methodology combines an analysis of the system parameters and their

relationships with the uncertainties arising from the model describing the biological system.

The analysis of the biological system is based on the idea that the set of parameters involved in the model can be classified into different typologies. Referring for instance to the case study of the circadian oscillations of the PER gene product in Drosophila (introduced in Sect. 8) two different classes of parameters were identified:

- The uncertain kinetic parameters K_U, that are known in terms of confidence intervals. Those parameters are for example the Michaelis constants or the constant rates for the kinases and the phosphatases.
- The control design parameters K_{CD}, that can be defined in a reliable way and determine the behavior of the biological system. Those are for example the maximum rate of degradation of an enzyme or the first order transportation rate parameters.

Since this biological design problem includes elements of the input data in a real-valued confidence interval, we deal with an optimization problem under uncertainty [Lodwick and Jamison, 2007]. In particular, the following programming constraint satisfaction problem is considered:

$$g_i(K_{CD}, K_U) \leq t_i \quad i = 1, \ldots, n \tag{2}$$

The constraint set is denoted as $\Omega = \{K_{CD} | g_i(K_{CD}, K_U) \leq t_i \;\; i = 1, \ldots, n, K_{CD} \in X\}$. The values of K_U are input parameters of the programming problem and are subject to uncertainty arising from different sources. Depending on the nature of the uncertainty, they may be probability distributions, intervals, fuzzy sets, or possibilistic distributions. In our case, these parameters are intervals which are particular cases of fuzzy numbers. The values t_i are the maximum or minimum thresholds for the constraints and can also be considered as uncertain.

In order to design a robust and reliable biological system with respect to a target performance, the design parameters were set up to balance the uncertainty of the kinetic parameters. To take into account these uncertainties arising from the estimations of the kinetic parameters, the function representing the feedback of the system was fuzzified. Finally, a measure of failure of the designed biological circuit to reach the required performance was analyzed by means of the possibility theory. The possibility measure is a consistent alternative to the statistical and probabilistic hypothesis. In fact, especially in fields such as the biological one, statistic and probabilistic assumptions are difficult to justify and moreover they could not include all the possible phenomena involved in biological processes. For such problems, little information regarding the uncertainty is known, and the uncertainty is typically modelled depending on expert opinions and assumptions made by the biologist. Fuzzy set theory is able to compensate the fact that uncertainty is modelled based on subjective opinions and assumptions. In contrast, probabilistic methods require large amount of data and the results obtained are, in same cases, sensitive to both the accuracy of the data as well as the assumptions made during the design process.

The interpretation of the possibility failure stands for a measure for the worst case design in the uncertainty context of the system. An optimization methodology adopted to minimize the possibility failure leads to a reliable configuration which guarantee the desired design.

4 Fuzzification of the Objective Function Representing the Performance

Fuzzy sets have been introduced by [Zadeh, 1965] as an extension of the classical notion of set. In classical set theory, the membership of elements in a set is assessed in binary terms according to a bivalent condition in which an element either belongs or does not belong to the set. By contrast, fuzzy set theory permits the gradual assessment of the membership of elements in a set; this is described with the aid of a membership function valued in the real unit interval $[0,1]$.

A fuzzy set is a pair (F,m) where F is a set and $m : F \to [0,1]$. For each $x \in F$, $m(x)$ is the grade of membership of x. An element mapping to the value 0 means that the member is not included in the fuzzy set while the mapping value 1 describes a fully included member in the fuzzy set. Values strictly between 0 and 1 characterize the fuzzy members. The set $\{x \in F \mid m(x) > 0\}$ is called the support of the fuzzy set (F,m) and the set $\{x \in A \mid m(x) = 1\}$ is called the Core [Klir and Yuan, 1995] of the fuzzy set (F,m).

A fuzzy number is a convex, normalized fuzzy set $\tilde{F} \subseteq \mathbb{R}$ whose membership function is at least segmentally continuous and has the functional value $\mu_F(x) = 1$ at precisely one element [Klir and Yuan, 1995].

In order to model with fuzzy numbers [Zadeh, 1968] the uncertainty arising from simulation design, a response surface of the function representing the feedback of the system was used as suitable approximation. The response surface was fitted with respect to the uncertain parameters sampled using a Latin Hypercube methodology.

Designing for uncertainty is computationally intensive and typically requires at least an order of magnitude more computational cost as compared to a corresponding deterministic design. Response surface approximations reduce the high computational cost associated with designing for uncertainty by using approximations that are accurate over the entire design space [Venter and Haftka, 1999] to replace costly stiff ODE system integrations. In the present work, the scheme to generate the response surface [Gavin and Yau, 2008] approximates the function with an approximate polynomial response surface of arbitrary order:

$$\tilde{g}(X) = a + \sum_{i=1}^{n} \sum_{j=1}^{h_i} (b_{ij} X_i^j) + \sum_{q=1}^{m} c_q \prod_{i=1}^{n} X_i^{p_{iq}} \tag{3}$$

where the coefficients b_{ij} correspond to terms involving only one random variable, and the coefficients c_q correspond to mixed terms, involving the product of two or more random variables. The polynomial order, h_i, the total number of mixed terms, m, and the order of a random variable in a mixed term, p_{iq}, are determined in the

algorithm described below. The algorithm used in this methodology, makes use of the last three stages of the High Order Stochastic Response Surface Method (HO-SRSM) [Gavin and Yau, 2008]. In the first stage, the number and types of mixed terms are determined. This stage results in the formulation of the higher order polynomial to be used for the response surface. After the formulation of the higher order polynomial, the coefficients of the higher order response surface polynomial are estimated in the second stage, using singular value decomposition to perform least squares on Latin Hypercube samples in the uncertain parameters space. Finally, from a Monte Carlo Simulation, the fuzzification of the performances are carried out with an acceptable approximation of the objective function given by the response surface.

The fuzzy representation of the performance is constructed enveloping the fitted data by intervals. The fuzzy map is built by α-level considering the minimum median interval which envelopes a fraction $(1-\alpha)$ of the performance values [Spinella and Anile, 2004]. In formal terms:

Definition 1 (Median Interval I_α). *Given n samples $X_1,\ldots,X_n$, and a reorder of them $X_{i_1},\ldots,X_{i_n}$, then the median interval at level α is:*

$$I_\alpha = [X_{i_j}, X_{i_k}] \tag{4}$$

where $j = \lfloor \frac{\alpha}{2} \rfloor + 1$ and $k = n - \lfloor \frac{\alpha}{2} \rfloor$.

The pseudocode in Algorithm 1 describes the procedure to fuzzify the performances. The method $RespSurfBuild$ generates the response surface using equation 3 and return the coefficients C. Then, a set of samples s of uncertain parameters K_U is generated using a Latin Hypercube technique. Finally, the response surface is evaluated on the set s and its output, the performance approximation, is fuzzified.

Algorithm 1. $Fuzzify(\mathbf{K_{CD}}, \mathbf{inf_u}, \mathbf{sup_u}, sys_{sim})$

$\mathbf{C} := RespSurfBuild(\mathbf{K_{CD}}, \mathbf{inf_u}, \mathbf{sup_u}, model_{sim})$
$\mathbf{s} := LatinHypercube(\mathbf{inf_u}, \mathbf{sup_u}, N_{samples})$ {random variables between **inf** and **sup** generated using a Latin Hypercube sampling}
$\mathbf{F_s} = RespSurfEval(\mathbf{s}, \mathbf{C})$;
$\tilde{F} := FuzzyfyFun(\mathbf{F_s})$

5 Possibility of Failure

One way to deal with design specifications is to compare the fuzzy numbers representing the performances with the crisp numbers representing a reasonable threshold and give a measure of satisfaction of these constraints. For this purpose the possibility measure of failure with respect to the specification constraint can give a useful information to improve the design . Note that a fuzzy number may also be considered as the trace of a possibility measure Π on the singletons (single elements) x of the universal set X [Zadeh, 1978].

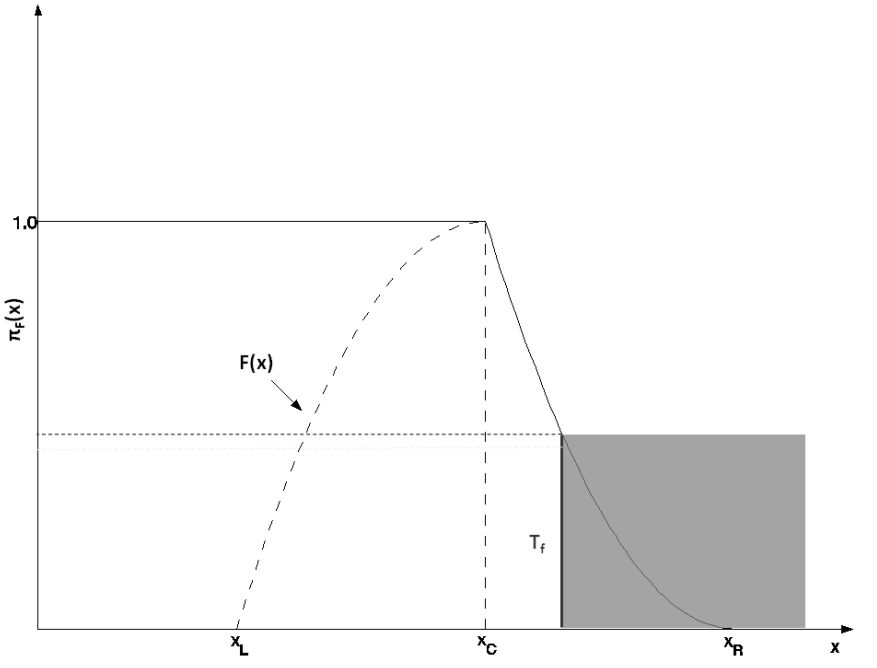


Fig. 1 Possibility distribution of $\tilde{F} \geq x$ for the nonlinear membership function $\mathbf{F}(x)$, with support in (x_L, x_R) and graphical description for the event of possibility measure of failure when $F > T_f$

When a possibility measure is considered, its possibility distribution π is then interpreted as the membership function of a fuzzy number $\tilde{F}$ describing the event that Π focuses on, as follows:

$$\Pi(\{x\}) = \pi(x) = \tilde{F}(x), \quad \forall x \in X \tag{5}$$

The possibility measure of a crisp number being smaller or equal to a fuzzy number $\tilde{F}$ is then defined [Dubois and Prade, 1988] as follows:

$$\Pi_{\tilde{F}}([x, +\infty)) = \sup_{y \geq x} \tilde{F}(y), \quad \forall x \tag{6}$$

The possibility distribution function $\Pi_{\tilde{F}}$ of the possibility measure $\pi_{\tilde{F}}$ can be seen in Fig. 1 for the general case of a nonlinear membership function.

Based on equations 5 and 6, given a representation of the function F and a maximal failure threshold T_f then the possibility measure of failure of this function is a measure for the event $F > T_f$, hence

$$\Pi_{\tilde{F}-T_f}([0, +\infty)) = \sup_{y \geq 0} (\tilde{F} - T_f)(y), \quad \forall x \tag{7}$$

Fig. 1 provides a graphical description of the event for the possibility measure of failure in the case when $F > T_f$. From each possibility measure related to the specification performances, it is possible to deduce a vector of measures $(p_1, \ldots, p_n)$ for a given design. A suitable metric that summarizes all can be used as objective for an optimization process. In this work the chosen function to optimize was:

$$\Sigma_{i \in I} \Pi_{\tilde{F}_i - T_{f_i}}([0, +\infty]) \tag{8}$$

where I is the set of performance $\tilde{F}_i$ to guarantee and T_{f_i} are the respective specification failures. This formulation of the function to optimize allows one to define the design goals through a single function that summarize all of them. Moreover, the sum of possibilities of failure avoids a conventional ideal design often based on unfeasible attainments while it identifies a more realistic design keeping the events of failure at a reasonable distance. From the mathematical point of view, the sum of possibilities is the L_1 norm inside the space of the problem objectives possibilities. This choice allows the characterization of convex regions of the multi-objective problem with a suitable merit function.

6 Methodology

Our methodology can be schematized in the diagram in Fig. 2. The two main interacting components are identified as Analyzer and Optimizer. The Analyzer

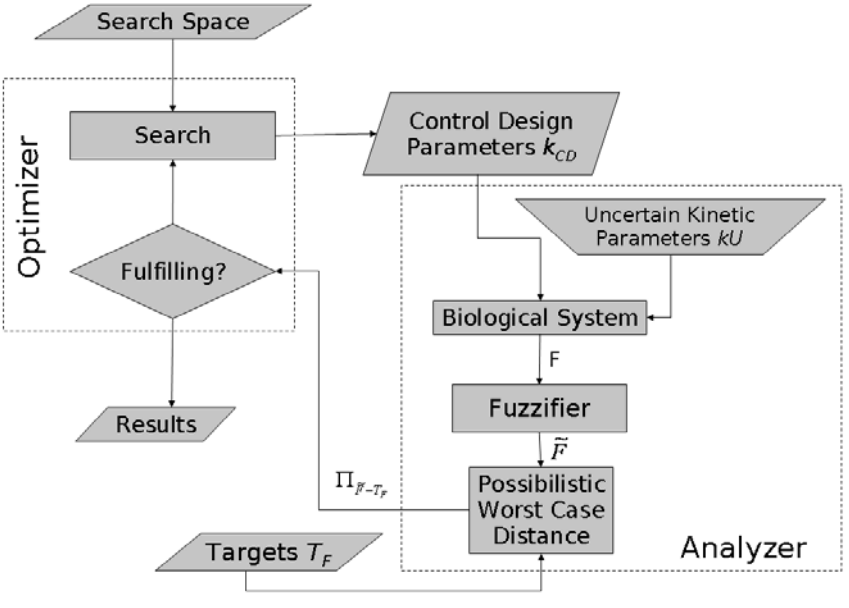


Fig. 2 Schema of the methodology

component contains the model describing the biological system, the Fuzzifier module and the Possibilistic Worst Case Distance (PWCD) module. This component can be seen as a black box taking as inputs the control design parameters k_{CD} and the chosen targets of the performances T_F of the biological system.

The Fuzzifier module samples the space of the uncertain kinetic parameters k_U through a Latin Hypercube methodology and approximates the performances F of the system using a response surface approximation. From this approximation, the fuzzification is carried out using the sampled uncertain kinetic parameters k_U while the control design parameters k_{CD} are considered as crisp numbers.

The PWCD module takes as input the fuzzified functions of the performances $\tilde{F}$ and the target of the performances T_F and gives as output the measure of failure of the biological system with respect to the given targets; this possibility measure of failure are estimated with the equation 7.

The Optimizer component searches for a set of control design parameters k_{CD} that minimizes the objective function (see equation 8) sum of all possibilities of failure for the biological system. The k_{CD} parameters will be given as new input for the Analyzer component. The optimization will run until a stop criterion is fulfilled (the sum of all possibilities of failure equals to zero), or alternatively until the unsatisfiability for the target is detected. Generally, the satisfiability of this stop criterion is not guaranteed for all the problems and a compromise between the targets should be chosen by the designer.

7 Biochemical Modeling

7.1 Oscillatory Biological Networks

Among the numerous biological networks models, autonomously oscillating systems are the most investigated. Such systems underly many of the periodic phenomena which have been identified in biology such as in processes describing glycolytic oscillations, the cell cycle, circadian rhythms, periodic neuronal signals [Goldbeter, 1996].

A first attempt at providing a mathematical model of the mechanism underlying these oscillations was presented by [Goldbeter, 1995] describing the circadian oscillations of the PER gene product in Drosophila. The PER gene was shown to play a role in circadian rhythms. This model was built in the absence of detailed molecular descriptions of the reactions involved, and was proposed as a minimal model which was able to reproduce experimental observations of wild type and mutant behavior. This model suffices for investigations of the core behavior of the system, and its simplicity recommends it for an illustrative example. Further analysis on how the kinetic parameters influence the extreme and the period of the oscillations can be found in [Ingalls, 2004, Bagheri *et al.*, 2007] where a detailed sensitivity analysis was presented. In comparison with the methods described in [Ingalls, 2004, Bagheri *et al.*, 2007], our methodology allows us to find a set of parameters that satisfy a fixed target without making use of an analytical approach.

Quantitative mechanism-based models could allow researchers to predict the comprehensive behavior of the specified system over time and to track its dynamics for each set of fixed system parameters [van Riel, 2006, Zi *et al*., 2005]. However, all of the parameters including rate constants and initial component concentrations in the mathematical models must be experimentally measured or inferred to specify the model. Even for those models with experimentally estimated parameters, it is still uncertain whether the particular set of parameters closely approximates the corresponding biological system because some of the kinetic parameters are usually taken or estimated from measurements reported by different laboratories using different in vitro models and conditions.

7.2 Kinetic of Biochemical Systems

A typical chemical or biochemical rate function relates the temporal change in a chemical compound to the concentration itself. In a simple first order degradation reaction which does not involve any enzyme, the change in concentration of the substrate over time is directly proportional to its concentration. This rate function can be mathematically formulated as follows:

$$\frac{dX}{dt} = -kX \tag{9}$$

where k (which is positive by definition) is the turnover per time unit and X denotes the concentration of the metabolite X. The negative sign indicates that material X is actually lost from existing pool. This mathematical form of the equation results from considerations of statistical thermodynamics.

When two substrates, say X_1 and X_2, go in to reaction to yield product X_3, the change of concentration of X_3 with respect to time can be written as:

$$\frac{dX_3}{dt} = kX_1X_2 \quad \frac{dX_1}{dt} = -kX_1X_2 \quad \frac{dX_2}{dt} = -kX_1X_2 \tag{10}$$

These formulas are written as the product of the concentrations of the species X_1 and X_2. This product form of the rate law is partly due to thermodynamics and partly due to the fact that the two metabolites have to come to the physical contact which is a matter of probability. That implies the simplest formulation as a product.

Michaelis-Menten rate equations describe the kinetics of many enzymes. This kinetic model is relevant in situations where very simple kinetics can be assumed. A substrate S and its catalyzing enzyme E form an intermediate complex (ES) in a reversible reaction. This complex can either break apart and return into substrate and enzyme or generate product P while releasing enzyme E unchanged.

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftarrows}} ES \overset{k_2}{\rightarrow} E + P \tag{11}$$

In order to obtain the Michaelis-Menten rate equations, some assumptions about the system should be made. The first assumption states that the total enzyme

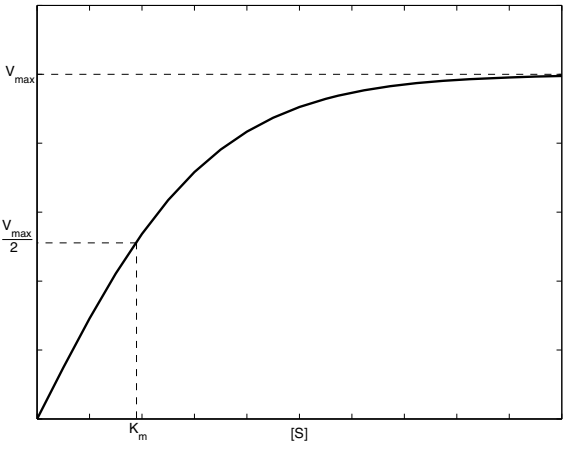


Fig. 3 Michaelis-Menten plot relating the reaction rate $\frac{dP}{dt}$ to the substrate concentration [S]

concentration can be divided into free enzyme and enzyme bound in the intermediate complex. The second assumption indicates that the total substrate concentration is much larger than the total enzyme concentration. The last assumption, which is called quasi-steady-state assumption, states that no enzyme is formed or lost during the reaction and the concentration of the intermediate complex is constant. Then the Michaelis-Menten rate is defined with the following equation:

$$\frac{dP}{dt} = \frac{V_{max}S}{K_m + S} \tag{12}$$

where the value of Michaelis-Menten rate constant is $K_m = \frac{k_{-1}+k_2}{k_1}$ and the maximum velocity $V_{max} = k_2 E_T$ occurs when the enzyme is saturated, i.e., when all enzyme molecules are tied up with S, or $(ES) = E_T$. Fig. 3 displays the variation of the product with respect to the concentration of the substrate.

The binding of a ligand to a macromolecule is often enhanced if there are already other ligands present on the same macromolecule. The Hill rate equation, provides a way to quantify this effect. It describes the fraction of the macromolecule saturated by ligand as a function of the ligand concentration; it is used in determining the degree of cooperativity of the ligand binding to the enzyme or receptor.

$$\theta = \frac{L^n}{K_d + L^n} = \frac{L^n}{K_A^n + L^n}, \tag{13}$$

where θ is the fraction of ligand binding sites filled, L denotes ligand concentration, K_d is the apparent dissociation constant derived from the law of mass action (equilibrium constant for dissociation), K_A is the ligand concentration producing half occupation (ligand concentration occupying half of the binding sites), that is also the microscopic dissociation constant and finally n is called the Hill coefficient and it describes the cooperativity. A coefficient $n = 1$ indicates completely independent binding, regardless of how many additional ligands are already bound.

Numbers greater than one indicate positive cooperativity, while numbers less than one indicate negative cooperativity.

8 Drosophila Circadian Rhythm Case Study

8.1 The Period Protein Model

Drosophila melanogaster is a two-winged fly and it is also one of the most commonly used model organisms in biology, especially in genetics and physiology. Some reasons for the choice of Drosophila as the most studied organism in biological research are: the simplicity of its morphology, the short generation time and the high fecundity. Period proteins are central components of the Drosophila circadian clock. Circadian clock generates circadian rhythms that are 24-hour activity cycles exhibited by the organisms during their life time. The model structure introduced in [Goldbeter, 1995] for the Period Protein (PER) of Drosophila is depicted in Fig. 4.

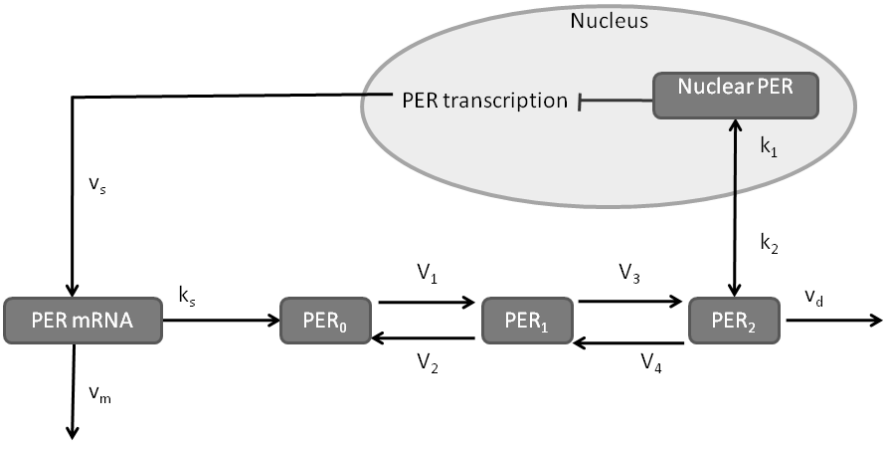


Fig. 4 Model of the period protein (PER) of Drosophila

The mechanism can be described as follows: the protein PER (P0) is produced in the cytosol at a rate determined by the concentration of PER mRNA (M). It is then reversibly phosphorylated at two sites (producing species P1 and P2). The fully phosphorylated protein can then be degraded or can migrate across the nuclear membrane. The variable PN describes the concentration of nuclear PER, which inhibits transcription of PER mRNA. This mRNA is subsequently degraded.

The quantities involved in this model are the following:

M is the PER mRNA,
P_0 unphosphorylated PER form,

P_1 monophosphorylated PER form,
P_2 bisphosphorylated PER form,
P_N is the nuclear bisphosphorylated form of PER,
v_s is the rate of mRNA production,
v_m is the rate of mRNA degradation,
n is the Hill constant which describes the cooperativity of ligand binding,
K_I is mRNA repression constant,
K_m is the Michaelis constant for the mRNA, degradation,
k_s is the rate of PER production,
V_i is the constant for the kinase,
K_i is the constant for the phosphatase,
v_d is the rate of degradation of the bisphosphorylated PER form,
k_1 is the transportation rate of the bisphosphorylated PER form in the nucleus,
k_2 is the transportation rate of bisphosphorylated nuclear PER form in the cytosol,
K_d is the Michaelis constant for the degradation of bisphosphorylated PER form

The time evolution of the specie concentrations is governed by the following kinetic equations:

$$\begin{aligned}
\frac{dM}{dt} &= v_s \frac{K_I^n}{K_I^n + P_N^n} - v_m \frac{M}{K_m + M} \\
\frac{dP_0}{dt} &= k_s M - V_1 \frac{P_0}{K_1 + P_0} + V_2 \frac{P_1}{K_2 + P_1} \\
\frac{dP_1}{dt} &= V_1 \frac{P_0}{K_1 + P_0} - V_2 \frac{P_1}{K_2 + P_1} - V_3 \frac{P_1}{K_3 + P_1} + V_4 \frac{P_2}{K_4 + P_2} \\
\frac{dP_2}{dt} &= V_3 \frac{P_1}{K_3 + P_1} - V_4 \frac{P_2}{K_4 + P_2} - k_1 P_2 + k_2 P_N - v_d \frac{P_2}{K_d + P_2} \\
\frac{dP_N}{dt} &= k_1 P_2 - k_2 P_N
\end{aligned}$$

The PER mRNA is synthesized in the nucleus and transfers to the cytosol where it accumulates at a maximum rate v_s and it is degraded by an enzyme of maximum rate v_m and Michaelis constant K_m. The rate of synthesis of the PER protein is proportional to the concentration of the PER mRNA and is characterized by a first order rate constant k_s. The reversible phosphorylation of P_0 into P_1 and P_1 into P_2 involves the parameters V_i and K_i that denote the maximum rates and Michaelis constants of the kinases and the phosphatases. The bisphosphorylated PER form P_2 is degraded by an enzyme of maximum rate v_d and Michaelis constant K_d and it is transported into the nucleus at a rate characterized by the first order rate constant k_1. The transportation of the nuclear bisphosphorylated form of PER into the cytosol is characterized by the first order rate constant k_2. The negative feedback exerted by P_N on the trascription of the PER protein is described by an equation of the Hill type where K_I is the threshold constant for repression.

8.2 The Use of Optimization: Parameters and Target Performances

The nominal values of the parameters are given in [Goldbeter, 1995] and were chosen so as to yield a period of oscillations close to 24 hours. A value of n=4 was chosen because the model can produce sustained oscillation in a larger parameter domain. But in this case, it was instead chosen a value of $n = 1$ to explore unstable configurations for the system. The new calibrated system leads to the parameter values showed in Table 1 with respect to reference concentrations of the species given in [Goldbeter, 1995]. Table 1 also shows the classification of the uncertain and control parameters.

Table 1 Estimated values and classification of the parameters in the PER model

Parameter	Estimated Value	Type
v_s	3.84	Control
v_m	3.381	Control
k_s	0.383	Control
k_1	1.747	Control
k_2	1.194	Control
v_d	0.934	Control
K_I	4.599	Uncertain
K_m	0.05	Uncertain
K_d	0.184	Uncertain
K_1	1.109	Uncertain
K_2	0.541	Uncertain
K_3	1.267	Uncertain
K_4	0.781	Uncertain
V_1	2.887	Uncertain
V_2	1.239	Uncertain
V_3	4.24	Uncertain
V_4	1.825	Uncertain

In this case study, the target performances of the required design problem are:

- the period (measured in hours), and
- the amplitude (measured in μmol)

of the concentration of the total quantity of the PER protein (P_t) which is given by:

$$P_t = P_0 + P_1 + P_2 + P_N \tag{14}$$

These performances are optimized by the methodology and they are expressed in terms of possibility of failure. In this particular test case, the minimum threshold for the period of the PER protein oscillations is fixed to 24 hours while the minimum threshold of the amplitude is fixed to a rather large value in order to guarantee significant oscillations. The optimization algorithm searches for the set of control design parameters that minimizes the objective function represented by the sum of these two possibilities of failure for the PER protein model.

An example of the sustained oscillations of the concentration of P_t generated by the model are depicted in Fig. 5.

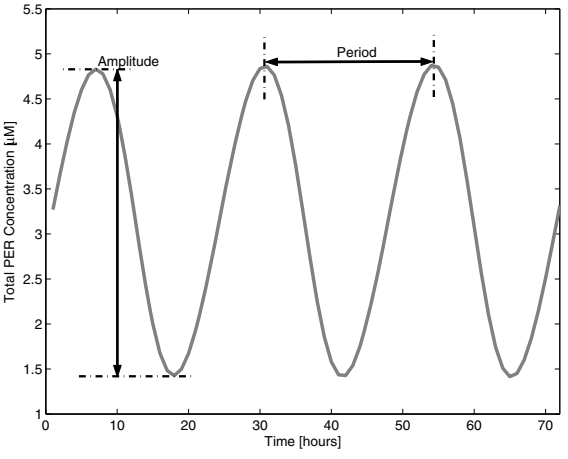


Fig. 5 Characteristics (period and amplitude) of the oscillations of the temporal variations of the total amount of PER protein

The initial conditions to solve the ODE system in this example are $\{M = 2.81, P_0 = 1.44, P_1 = 0.72, P_2 = 0.48, P_N = 0.63\}$ and the parameters are taken from Table 1. Starting from this set of initial conditions the system reaches a unique, closed curve, in the (M, P_t) plane, characterized by a period and amplitude that are fixed for the given set of parameter values (see Fig. 6).

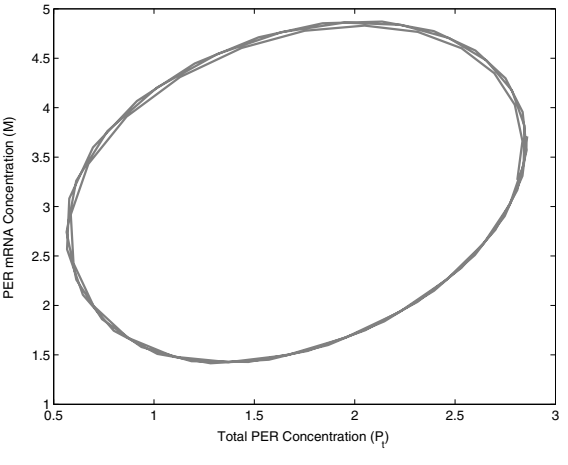


Fig. 6 The sustained oscillations of the temporal variations of the total amount of PER protein and of the per mRNA correspond to a system that reaches a unique, closed curve, in the (M, P_t) plane

9 Results

The fuzzified performances were approximated using 50 samples computed by means of a response surface of the second order. An example to explain the application of the methodology on the PER protein model can be seen in Fig. 7 and 8, where an uncertainty of a 10% in the kinetic parameters k_U was considered.

The starting point of the parameters (in Table 1) of the PER model led to a period of 23.7 hours and to an amplitude of $3.45\mu mol$ and the initial fuzzy numbers are shown in Fig. 7. The two graphs in Fig. 8 display the fuzzy numbers of the period and the amplitude of the oscillations of the PER protein concentration after the optimization process fixing the failure threshold of the period to the value of 24 hours (figure on the left) and the failure threshold of the amplitude to the value of $5\mu mol$ (figure on the right).

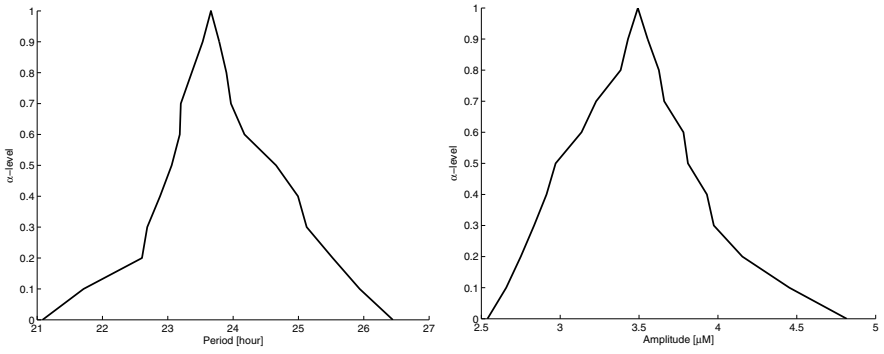


Fig. 7 Fuzzy numbers representing the period and the amplitude of the oscillations of the PER protein concentration before the optimization process

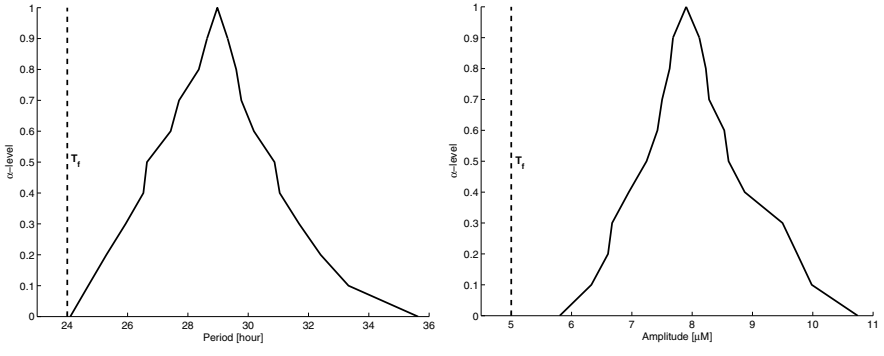


Fig. 8 Fuzzy numbers representing the period and the amplitude of the oscillations of the PER protein concentration after the optimization process

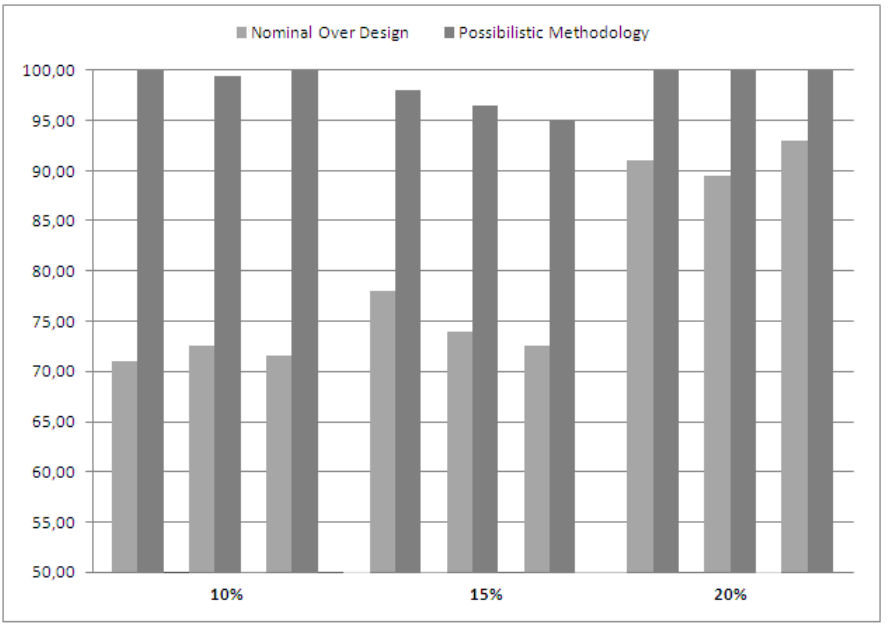


Fig. 9 Comparison between Nominal Over-Design and the Possibilistic methodology in terms of acceptable circuit varying the uncertainty of the kinetic parameters of 10%,15%,20% and 25% with regard to their nominal default value

The histogram in Fig. 9 describes in detail the comparison between the presented methodology and the "Nominal Over-Design", one of the most used deterministic design methodology. For the Nominal Over-Design, every objective was fixed to a given factor of safety with regard to the nominal target specifications. In this test case the target of the period and the amplitude were increased of a 10%, 15% and 20% in regard to the minimum threshold of 24 hours and $5\mu mol$ respectively, in order to counterbalance the uncertainty of the kinetic parameters.

Both methodologies used the Nelder and Mead Simplex [Nelder and Mead, 1965] optimization algorithm. Even thought the convergence of the algorithm is not always guaranteed, in this case the shrink steps were rare and when the Nelder-Mead iteration stagnated, a restart with the same best point and a different set of directions helped the convergence. The choice of this algorithm is due to the fact that it requires only function evaluations without derivatives estimation and it turns out to be the most appropriate in this application where the merit function is well shaped.

After the optimization process, an estimation of the percentage of "acceptable" biological circuits value was computed by means of Montecarlo simulations. The biological circuit is classified as "acceptable" if the performance specification is met i.e. if the resulting period is at least 24 hours and the amplitude of the oscillations was at least $5\mu mol$. Three independent tests were carried out considering a statistical

distribution of the uncertain kinetic parameters of 10%,15% and 20% with regard to the nominal default value. Every test performed 200 simulations.

This comparison points out that the possibilistic methodology always guarantees a better outcome of performance with respect to the Nominal Over-Design methodology. In fact, with regards to the Nominal Over-Design, it turns out to be difficult to set a suitable over-achievement which identify the best parametrization towards the uncertainty of the kinetic parameters. Note that there exist a fundamental difference between the deterministic approach and the fuzzy set based approach. The deterministic approach tends to equalize the failure load of each failure criterion, while the fuzzy set based design tends to equalize the possibility of failure of each failure criterion.

10 Conclusions

This study proposes the use of fuzzy numbers and possibility theory to perform a target satisfaction in the framework of future biological systems design that will demand reliable feedback. In order to assure a certain performance of a given biological circuit the presented methodology set the values of the design parameters of the model to balance the uncertainty of the kinetic parameters. The feedbacks of the system are represented as fuzzy numbers. By means of the possibility theory a failure value, which represents the worst case design of the system in the uncertainty context, was measured. This possibilistic failure value finally was minimized using an optimization methodology.

The proposed methodology was tested on the model describing the circadian oscillations of the PER gene product in Drosophila. The application over the oscillatory circuit of the PER gene has shown that the proposed methodology guarantees always a reliable outcome of the period and the amplitude of the oscillations of the PER gene. Moreover, a comparison of the methodology with a more diffused methodology named "Nominal Over-Design", demonstrated an higher percentage of acceptable biological circuit in the possibilistic methodology.

Future works will apply this methodology on more then two performances of a biological system including also the phase [Bagheri *et al*., 2007] for example. In fact, usually, a biological system is characterize by multiple features that can also be in conflict with each other. The methodology could allow to design the biological system to assure the defined performance including the uncertainty of the parameters into the model. To further improve the proposed methodology other optimization algorithms will be tested in the framework, such as genetic algorithms or simulated annealing. Finally, more biological systems, even non oscillatory networks, will be tested and analyzed.

Acknowledgements. This work came from an idea given by our mentor Prof. A.M.Anile some years ago. We always keep his real paragon of scientist and his helpful encouragements in our memory.

References

[Bagheri *et al*., 2007] Bagheri, N., Stelling, J., Doyle III, F.J.: Quantitative performance metrics for robustness in circadian rhythms. Bioinformatics 23(3), 358–364 (2007)

[Banga, 2008] Banga, J.R.: Optimization in computational systems biology. BMC Systems Biology 2, 47 (2008)

[Dubois and Prade, 1988] Dubois, D., Prade, H.: Possibility theory: An approach to computerized processing of uncertainty. Plenum Press, New York (1988)

[Gavin and Yau, 2008] Gavin, H.P., Yau, S.C.: High-order limit state functions in the response surface method for structural reliability analysis. Structural Safety 2(30), 162–179 (2008)

[Goldbeter, 1995] Goldbeter, A.: A model for circadian oscillations in the Drosophila period protein (PER). Proc. R. Soc. Lond. B 261, 319–324 (1995)

[Goldbeter, 1996] Goldbeter, A.: Biochemical Oscillations and Cellular Rhythms. Cambridge University Press, Cambridge (1996)

[Heinrich and Schuster, 1998] Heinrich, R., Schuster, S.: The modelling of metabolic systems: Structure, control and optimality. Biosystems 12(47), 61–77 (1998)

[Ingalls, 2004] Ingalls, B.P.: Autonomously Oscillating Biochemical Systems: Parametric Sensitivity of Extrema and Period. Systems Biology 1(1), 62–70 (2004)

[Kholodenko *et al*., 1999] Kholodenko, B.N., Demin, O.V., Moehren, G., Hoek, J.B.: Quantification of short term signaling by the epidermal growth factor receptor. J. Biol. Chem. 274, 30169–30181 (1999)

[Klir and Yuan, 1995] Klir, G., Yuan, B.: Fuzzy sets and fuzzy logic: theory and applications. Prentice Hall, USA (1995)

[Lodwick and Jamison, 2007] Lodwick, W.A., Jamison, K.D.: Theoretical and semantic distinctions of fuzzy, possibilistic, and mixed fuzzy/possibilistic optimization. Fuzzy Sets Syst. 158(17), 1861–1872 (2007)

[Nelder and Mead, 1965] Nelder, J., Mead, R.: A simplex method for function minimization. Computer Journal 7, 308–313 (1965)

[Sciacca *et al*., 2007] Sciacca, E., Spinella, S., Anile, A.M.: Possibilistic Worst Case Distance and Applications to Circuit Sizing. Theoretical Advances and Applications of Fuzzy Logic and Soft Computing 42, 287–295 (2007)

[Segre *et al*., 2002] Segre, D., Vitkup, D., Church, G.M.: Analysis of optimality in natural and perturbed metabolic networks. Proc. Natl. Acad. Sci. USA 23(99), 15112–15117 (2002)

[Spinella and Anile, 2004] Spinella, S., Anile, A.M.: Modeling uncertain sparse data with fuzzy b-splines. Reliable Computing 10(5), 335–355 (2004)

[Stelling *et al*., 2004] Stelling, J., Sauer, U., Szallasi, Z., Doyle, F.J., Doyle, J.: Robustness of cellular functions. Cell 6(118), 675–685 (2004)

[Torres and Voit, 2002] Torres, N.V., Voit, E.O.: Pathway analysis and optimization in metabolic engineering. Cambridge University Press, New York (2002)

[van Riel, 2006] van Riel, N.A.W.: Dynamic modelling and analysis of biochemical networks: mechanism-based models and model-based experiments. Briefings in Bioinformatics 7(4), 364–374 (2006)

[Venter and Haftka, 1999] Venter, G., Haftka, R.T.: Using response surface approximations in fuzzy set based design optimization. Structural and Multidisciplinary Optimization 18(4), 218–227 (1999)

[Wang *et al*., 2007] Wang, R.S., Wang, Y., Zhang, X.S., Chen, L.: Inferring transcriptional regulatory networks from high-throughput data. Bioinformatics 22(23), 3056–3064 (2007)

[Zadeh, 1965] Zadeh, L.A.: Fuzzy sets. Information and Control 8(3), 338–353 (1965)
[Zadeh, 1968] Zadeh, L.A.: Fuzzy algorithms. Information and Control 12(2), 94–102 (1968)
[Zadeh, 1978] Zadeh, L.A.: Fuzzy sets as the basis for a theory of possibility. Fuzzy Sets and Systems 1, 3–28 (1978)
[Zi *et al*., 2005] Zi, Z., Cho, K.H., Sung, M.H., et al.: In silico identification of the key components and steps in IFN-λ induced JAK-STAT signaling pathway. FEBS Letters 579, 1101–1108 (2005)

Fuzzy Optimal Algorithms for Multiple Target Convergence

Zach Richards

Abstract. The proposed fuzzy application is the use of fuzzy algorithms for a networked swarm of autonomous vehicles, such as those used in planet exploration, and to be used in target location determination and convergence. In particular, an algorithm of this type could be used in an Autonomous Stratospheric Aircraft (ASA), thus having the possibility of being used for the exploration of a planet as well as many other applications. Upon finding an unknown location of a specified target, the algorithm would then swarm and eventually converge upon the location. There are two similar, but fundamentally different algorithms proposed in this presentation. These algorithms are capable of locating and converging upon multiple targeted locations simultaneously. This project is inspired by the current thought of NASA in the search of life on Mars, which is "Follow the Water" [17], where the targeted location would be a water source. These algorithms make use of combining a modified Particle Swarm Optimization algorithm combined with fuzzy variables for added intelligence.

1 Introduction

Target location has been a very important subject in military and scientific study. Radar is used to range and altitude of an object. Sonar is used to locate the range and depth of an object in the water. Global Positioning System is used to keep precise measurements of a target on the ground once it has been located. National Aeronautics and Space Administration (NASA) has flown a variety of missions to Mars to determine if water can be located, as it was confirmed with the Phoenix Lander [16]. The satellites and landers that

Zach Richards
United Launch Alliance
Littleton, CO 80127 USA
e-mail: zachary.d.richards@ulalaunch.com

W.A. Lodwick & J. Kacprzyk (Eds.): Fuzzy Optimization, STUDFUZZ 254, pp. 479–501.
springerlink.com

were flown from Earth to Mars were large, heavy, and bulky; while creative engineering solutions and recent advancement in technology give birth to other possible missions.

With the creation of nano-technology and the use of optimization in the design process, it may be possible to make a much smaller, lighter and more compact system to search for a target. Nano-technology has made it possible for many *target searching tools* to be launched simultaneously. The target searching tools are autonomous aircraft and/or ground vehicle launched for target location. This chapter investigates multiple newly developed algorithms that could be used for a task such as target location and convergence using multiple target searching tools. The notion of swarm intelligence is developed as well as new methodologies for convergence upon multiple targets by method of a swarm.

2 Problem Statement

Assume there is some space in $\mathbb{R}^2$, where there exists multiple physical locations, such as water, however their locations are unknown. In order to determine the location of these targets as quickly as possible, it would be best if the targets were located simultaneously, thus requiring multiple agents to perform the search. Is it possible to control multiple agents or a swarm of autonomous vehicles simultaneously? Is it possible to control the swarm to converge upon multiple locations simultaneously? Is it possible to control the swarm with a single equation? We will be introducing two new algorithms to answer these questions as well new methodology in swarm search methods.

3 Particle Swarm Theory

Particle Swarm Optimization (PSO) is an evolutionary algorithm developed by R. Eberhart and J. Kennedy in 1995 [3]. Many variations of the Particle Swarm have been proposed since, but they all tend to have the same formulation, presented below. The varied algorithms consist of the difference between a particles personal best location and current location, in addition to the difference between the global or neighborhood best and a particles current location. This is further discussed below.

The PSO algorithm is similar to the Genetic Algorithm (GA)[5], but has one very distinct difference. Both algorithms use a randomization to initialize the algorithms populations. The distinct difference between PSO and GA lies in the movement of the populations, upon the initial positions being established. PSO uses a randomized velocity function of the current position to determine the next position of all particles, whereas the Genetic Algorithm calls for a reproduction of the best possible positions.

The PSO algorithm is widely sought for a variety of economic and engineering problems, because its reliability and simplicity to implement[19]. The

PSO algorithm is reliable, because it performs a thorough search of the design space and the communication between the particles allows the particles to converge upon an optimal solution. However, there is no proof demonstrating that it will always locate the global optimal solution. The simplicity lies in the lack of parameters to initialize and manipulate at each iteration of algorithm. There are two main parameters, position, x_{id}, and velocity, v_{id}. The x_{id} parameter gives the current position of the *particle*, and then the particles are "flown" through the problem space at velocity, v_{id}, for one time increment per iteration. In contrast, a GA must recompute the whole "genetic" structure.

The following general PSO algorithm, consisting of two equations, effectively demonstrates the simplicity of the optimization technique [10].

$$v_{id} = w * v_{id} + c_1 * rand() * (p_{id} - x_{id}) + c_2 * Rand() * (p_{gd} - x_{id}) \quad (1)$$

$$x_{id} = x_{id} + v_{id} \quad (2)$$

Equations (1) and (2) describe the updating behavior of the PSO algorithm. If there are n variables in x_{id}, then x_{id} is a $n \times 1$ vector, and as a result there are n elements in v_{id}, which, also is a $n \times 1$ vector. The PSO algorithm requires that the particles remember their personal best, p_{id}, as well as the local best, p_{ld}, or "global" best, p_{gd}. The local best may also be referred to as the neighborhood best. A particle's personal best is the best position determined thus far by each individual particle. The swarm best is the best position determined thus far by the overall swarm.

The PSO algorithm uses a randomized population. In Equation (1), $rand()$ denotes a randomized number for multiplication by a particle's personal performance, whereas $Rand()$ denotes a randomized number for multiplication by the performance of the best in the local or global swarm. The current position and personal best values change with each variable, for each particle, and for each iteration. For example, if the design space was of 5 variables, then the vector for the position of each individual particle would consist of 5 elements. Each element in the vector would denote a position with respect one control variable.

The two coefficients c_1 and c_2 are used to weight the importance of the personal versus the swarm best. These coefficients can drastically effect the performance and reliability of the algorithm as shown by [19], who suggest the values of $c_1 = 2.0$ and $c_2 = 1.0$ or $c_1 = 2.5$ and $c_2 = .5$. If $c_1 >> c_2$ then the result is slower converging algorithm, and runs the risk of not converging. However, if $c_1 << c_2$, then this produces a much faster converging algorithm, because the velocity will stay large even if the difference between the swarm best and current position is not very large. There is a risk that a true non-robust global minima may be over shot, and a more robust local minima is found. A robust solution is a solution that is stable, where if any one or more variables in the solution was to be slightly perturbed, the solution would still

only be slightly different. Another words, a robust solution is not sensitive to one or more variables, becoming slightly perturbed. This is very important in applying optimization to engineering applications. These two cases both have benefits and downfalls, and for further discussion refer to [19].

The last parameter is the inertia weight, w, which was not in the original particle swarm paper, [3]. The addition of the inertia weight parameters has shown an increase in the performance for a variety of applications. According to [4], w was originally developed and is often decreased linearly as a function of iterations from 0.9 to 0.4. Many other methods of using w have been proposed, and are thoroughly discussed by Ruben Perez of the University of Toronto and Kamran Behdinan of Ryerson Polytechnic Institute in [19].

4 Fuzzy Variables

Fuzzy variables are used when traditional two-valued logic standards do not suffice and what is being modeled is inherently transitional in its nature. Traditional logic has to be either true or false, but not both, and is commonly viewed as black and white with no grey area. The power of fuzzy variables is that it introduces a calculable grey area [18].

For example, when asked, "Is the car on or off?" The car must be on or off, where "on" is represented by the value 1 and "off" is represented by the value 0. We know the car must be in one of these states, because no other logical states exists.

However, in another example, if asked, "Is the car temperature hot?" there are many possible answers, because there has to be a pre-understanding to what *hot* is. It may be *hot* when compared to ice, but cold when compared to the temperature of the sun. However, it is possible for the temperature to be in an in between state of hot and cold, such as warm or cool. Therefore the value is not 0 nor 1, but it is in between 1 and 0. Therefore the variable, temperature, is a fuzzy variable in this example.

The value for fuzzy variables is usually determined by taking a measurable value such as temperature and evaluating on a predetermined curve, known as a *membership function*, to find the corresponding value of the fuzzy variable. In context of the proposed problem with the ASA's searching for multiple targets simultaneously, A, B, and C are fuzzy, because their value is representative of the physical size of the target, being area. Let targets A, B, and C have a surface area value of $120ft^2$, $70ft^2$, and $170ft^2$, respectively. By using following membership function denoting "larger target area" depicted in 1 we are able to determine the value for the fuzzy value for the size of each target. The values for targets A, B, and C are .5, .25 and .75, respectively. By prioritizing targets by size we are able to develop a relationship of importance that allows similar size targets to be valued as nearly the same.

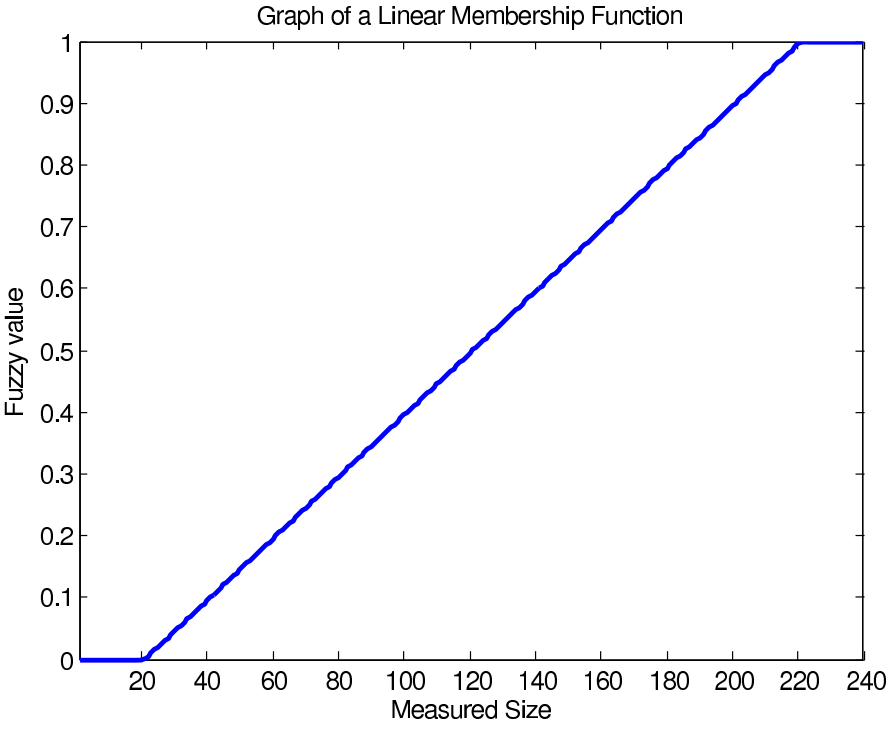


Fig. 1 Linear Membership Function

5 Methodology

By using a decrease of w, the weight coefficient, the PSO algorithm uses a very primitive form of intelligence, because it causes the the effect of the previous iteration velocity vector to decrease with increasing iterations. For example, a particle may begin searching by making large sporadic steps, because the information for the current step is so heavily dependent upon the ratio of the previous step and current information. As the algorithm iterates the ratio of previous step to current information decreases. Thus, causing the particles to take more direct step in the direction of the swarm best solution, which demonstrates that the decrease of w to be a primitive form of intelligence.

There are a few different methods that have been proposed on how to efficiently use w. The first method is a fixed weight method, where w is not increased or decreased. Thus, removing the primitive form of intelligence. The second method is a linear decrease of w to allow the particles to more thoroughly search without large velocity steps as the search progresses. The third method is a dynamic decrease that takes into account the current projection, like the linear method it allows a more though search as the search

progresses. Each of these approaches have been studied by Ruben Perez and Kamran Behdinan in [19]. They empirically found that a dynamic decrease was the most efficient approach, however using a linear decrease did not cause much, if any, noticeable change in the time to determine an "optimal" solution. A linear decrease will be used in the following algorithms, because it is the most common approach used in particle swarm optimization. By using the most common approach it is easier to understand the effect of fuzzy variables to the algorithm. The linear modification of w is a primitive form of intelligence, because the modification, increase or decrease, of w is predetermined prior to performing the algorithm.

By introducing fuzzy variables a higher level of algorithm intelligence can be reached. Fuzzy variables are dependent upon membership functions to determine the fuzzy value. Membership functions can be adaptive and non-adaptive. A non-adaptive fuzzy membership function uses a fixed membership function for the determination of their fuzzy values. An adaptive membership function will modify the shape of the membership function based on the experience of the system and what the system has learned. Generally adaptive membership functions are chosen for systems that are highly dynamic as well as have a pattern that can become recognized by the system. The problem being proposed is not dynamic, because the targets are to remain stationary both in position and size over the course of the search. A dynamic problem would consist of moving or non-stationary targets. In addition, we are assuming that there is no pattern to finding the targets. To continue with the working example of the target being water, this assumption isn't valid, because there is a pattern to finding the target. Typically a lake is surrounded by heavier vegetation than the surrounding area. However, we are making the assumption of no fixed pattern to be found, to increase the usability of the developed algorithms.

The proposed problem raises multiple questions which need to be evaluated prior to developing the fuzzy membership functions or the fuzzy velocity algorithm. Some of those questions are as follows: Is the overall time of the system being optimized or is the efficiency to search the given space in a given time being optimized? How does the effectiveness of an algorithm get measured?

Given this system is being designed for a space application or predominately a time sensitive application, the efficiency of the swarm algorithm for this application measures optimal time, in iterations, to locate all targets in a given space. The reason this is predominately a time sensitive application is driven by the power needed for the exploration. A solar panel would not be sufficient, because there lack to efficiency to effectively collect power and the search area of the aircraft would be limited by keeping the sun in position of the solar panels. We would like the aircraft's search space to not be limited by the optimal solar panel position, therefore a different power source would be needed. The aircrafts would then need to be powered by Lithium-Ion batteries, nuclear power, and/or a paper battery, [13].

It would be best allow the swarm to locate multiple targets simultaneously to ensure that all aircrafts are being used to their maximal potential as well as to optimize efficiency for a given fixed time of battery life. When we allow the swarm to investigate multiple targets simultaneously, therefore the following search methodologies were newly developed as part of the research for the problem stated above.

1. There is the **Sectioning Method**, which is explained as follows. Upon a target being located the n-closest aircrafts (particles) to a target converge upon that target with minimal search for other targets. Meanwhile, all other particles continue to search without knowledge of a target being located. The drawback of a system with this method is that the target may be over/under populated. The over/under population of the individual targets would result from the preset value of the n-closest particles to a target. The benefit of a system like this is that all possible targets would be located unless there were an under-population of particles to targets.
2. There is the **Search and Converge Method**, which is explained as follows. When a target is located all particles would use an algorithm to continue to search while moving in the direction of the determined target as effectively and efficiently as possible. If another target is located, the particles would choose by some means which target they were to converge upon. In conclusion, each and every located target would be converged upon. This is much like the Particle Swarm Optimization algorithm, because all particles are converging upon the global best, while trying to improve upon their own position and determine a new personal best and/or global best. The drawback of a method such as this is also a benefit and that is defined in the terms, effectively and efficiently, because they are user defined. The drawback is that this may not have the capability to converge as quickly as the user may choose. The benefit of this method is that it would be a mixture of efficiency and effectiveness, and that mixture is defined or developed by the user.
3. There is the **Random Selection Method**, which is explained as follows. Upon a target being located n particles in the swarm would be randomly selected to converge upon the targeted location using an algorithm. If one of those particles was to locate another target in route to converge upon the targeted location, then this would activate another random selection of n particles that were not previously chosen, to converge upon the new location. Meanwhile, the previously selected particles in the swarm would continue to converge upon their target with no knowledge of where the new target was located. This process would continue to repeat and would most likely conclude in two different scenarios. If an under-population existed, then all particles would have converged upon a target without knowledge of all targets being found. If an over-population existed, then some particles would continue to search the region although all targets to be found have been found. The drawback of this method is either being ineffective or too effective. To resolve this drawback knowledge of the feasible region would

need to be known, however this is against the assumption of characteristics about the feasible region. The effectiveness of this method is a large benefit.

The Search and Converge Method provides the most opportunity given the assumptions and the need to have the method work independently of the number of particles in the swarm and possible targets. Primarily, the assumption of not knowing how many targets exists. The Sectioning Method and Random Selection Method have the risk of too many or too little particles.

6 Initialization of Swarm

Particle Swarm theory requires the particles to be initialized to a particular position. There are two possible methods on how to initialize these particles. The first method is randomly generated positions, where the particles will be initialized by a computer, therefore, the random appearance of grouping may occur. In context of the proposed problem for planet exploration, we are not concerned with how the type of initialization is accomplished.

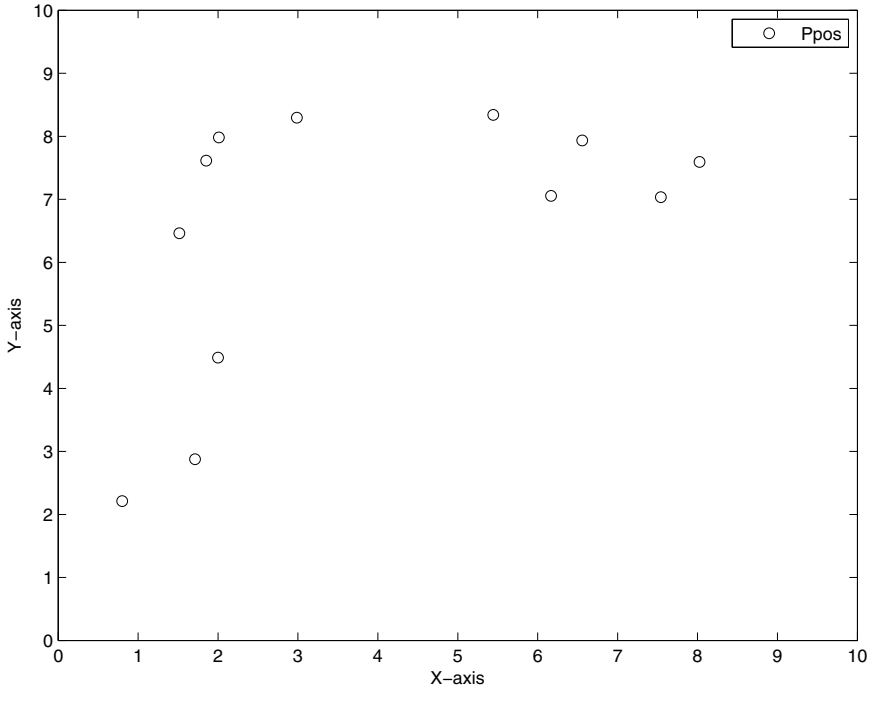


Fig. 2 Example of Computer Random

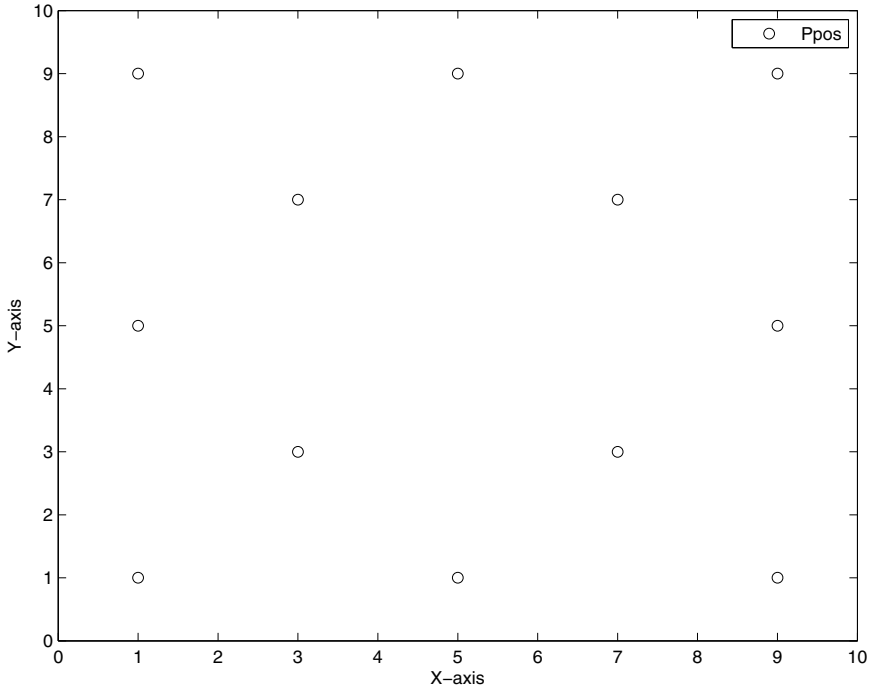


Fig. 3 Example of Regularly Spaced Grid

The second method is an evenly Regularly Spaced Grid or human random, where the particles are evenly spaced and distributed across some region. According to Summer Ann Armstrong in [2], humans automatically associate pattern with being non-random, and a grouping of particles or numbers listed from 0 to 100 is thought to not be random. However, humans automatically inject a pattern by spacing numbers imprecisely evenly if they were asked to list 20 numbers from 0 to 100. A visual example of this phenomenon is seen below.

Both methods are valid for the initialization of the particles, but Particle Swarm Theory tends to use computer randomization to initialize the particles within the swarm. However, there have been no results to demonstrate that one initialization method is better than the other. The particle initialization in the project will be by the computer randomization method. In a real application the mission designer would need to account for how to place the particles on the surface to be the primary reason for the type of initialization to be used.

7 Introduction of Algorithms

An algorithm is be used, because an algorithm has the capability to control a collection of points or swarm of particles by using a few general rules that apply to all particles. Otherwise, each individual particle would need to have a collection of rules that specifically applied to its individual situation. Because the complexity of the proposed problem, each individual situation would be very difficult to account for, therefore an algorithm can account for all situations with a general set of rules or guiding principles.

The first algorithm to efficiently and effectively control the convergence of the population to one or multiple targets, is to be referred to as the Single Fuzzy Parameter Method (SFPM). It consist of a single fuzzy parameter, t_n, where n is the number assigned to the located target. The *fuzzy size parameter*, t_n, denotes the value determined from a fuzzy membership function based upon the size of the target. The size is only determined upon a particle determining the location of the target.

The second algorithm, referred to as the Double Fuzzy Parameter Method (DFPM), consists of two fuzzy parameters, t_n and d_{id}. Here t_n, again, denotes the size of the target, and the second fuzzy parameter, d_{id}, denotes the *fuzzy distance parameter*. The fuzzy distance parameter is calculated by a comparison to the average distance from all other particles to a targeted location.

A third algorithm exist as a baseline to compare the first two algorithms, to be referred to as the Baseline Algorithm. It will contain no fuzzy variables and will consists of the general equation that causes the first two algorithms to search and converge upon a target being located.

8 Single Fuzzy Parameter Method (SFPM)

The velocity vector equation for the Single Fuzzy Parameter Method is presented below. The equation is used to calculate v_{id}. Every time a previously unlocated target is located, n of t_n is incremented by one. This allows the algorithm to identify the size of the target to the order in which the targets were determined. It is assumed that the particle or autonomous aircraft would be able to determine the size of the target upon locating the target.

$$v_{id} = w * v_{id} + c_1 * rand * (G_{id} - x_{id}) * t_n \tag{3}$$

Note: If $(G_{id} - x_{id}) = 0$, then $v_{id} = 0$. This implies that when a particle has located a target, the velocity is set to 0, therefore causing the particle to be stationary, and is used in both the SFPM and the DFPM.

The size of the target controls how fast or slow the particles converge on the located target. For example, if an aircraft is near a targeted location, A, but the A is small, then the Fuzzy Size Variable, t_n, would be small. Therefore the calculated velocity, v_{id}, would be small, and the convergence of

the aircrafts upon the targeted location would be slow. The opposite is also true if the located targets size is large.

In the velocity equation, Equation 3, the size of the target is multiplied by the difference of the current position and the best swarm position. The best swarm position is the position of the first particle that located the target. The center of the target is not used, because I am assuming that the particle or aircraft is unable to search the whole target and determine where the center is located. As a result of the Fuzzy Size being multiplied and not added, the velocity vector can be positive or negative. A calculation that resulted in a negative velocity would be a step in the negative direction with respect to a predetermined positive step. For example, if a calculation was positive it would result in movement to North, as a negative calculation would result in movement to the South or vice versa and the same is for East and West.

The following equation is used to determine the proper velocity vector, if no target has been located.

$$v_{id} = -1 + 2 * rand() \tag{4}$$

This velocity vector is a random distribution of [-1,1], and as a result the velocity vector can have a positive or negative direction.

When a particle has located a target, all particles begin to swarm to the location of the locating particle. If another target is located, then each particle selects the closest targeted location as its swarm best. This repeats every time a new target is discovered or until all particles are at the location of a target. The above concept is important, because it does cause the all particle to converge to a targeted location as quickly as possible by the particles selecting the closest target. The downfall is that there is no limit on the number of particles allowed at a target, and as a result the algorithm may or may not find all possible targets.

Upon the first target being found, all particles begin to swarm to the target using the SFPM for the calculation of the velocity vector. As a result of all particles swarming to the targeted location, the act of any other targets being located is an act of randomness of chance. The following figures demonstrate the initial positions of the particles denoted $Ppos$, where the target boundaries are denoted by $Tlim$ in the x and y direction and the center of a targets are denoted by $Tpos$.

The algorithm terminates upon all particles effectively converging upon a targeted location. The particles are given a viewable radius of .25 from their current location. This implies that particle can only view a target or a target boundary, $Tlim$, if that boundary is less then or equal to the .25 viewable radius.

In hindsight, from the above example, we see that there are two particles randomly initialized to the location of a target, because the allowed viewable range of a particle. The target approximately centered at $(6.25, 5.75)$ is

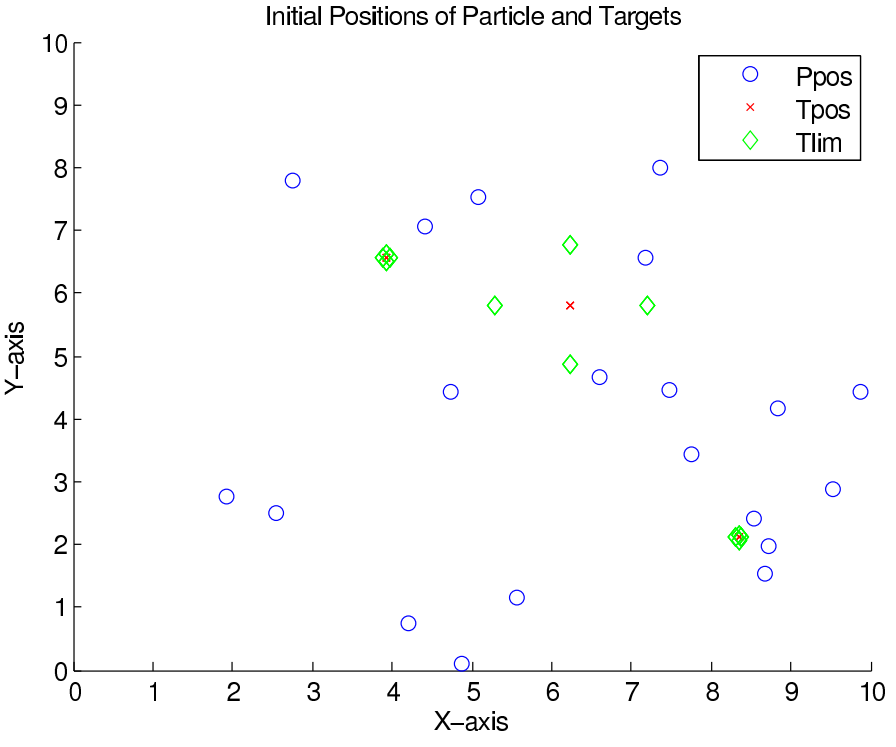


Fig. 4 Example of the Initial Positions prior to performing the Single Fuzzy Parameter Method

located by a particle located approximately at $(6.5, 4.75)$ and a second particle located approximately at $(6.75, 7.25)$.

It can also be seen that the largest target had the most particles swarm to that location, whereas there was one target which was not located and observed. However, by random chance a second target was located and two of the twenty particles swarmed to its location.

9 Double Fuzzy Parameter Method (DFPM)

The velocity vector equation for the Double Fuzzy Parameter Method is presented below. The equation is used to calculate v_{id}. Every time a previously unlocated target is located t_n is incremented to t_{n+1}, which is the same methodology that was used in the SFPM. As a result the algorithm is capable of determining the size of the target to the order in which the targets were determined.

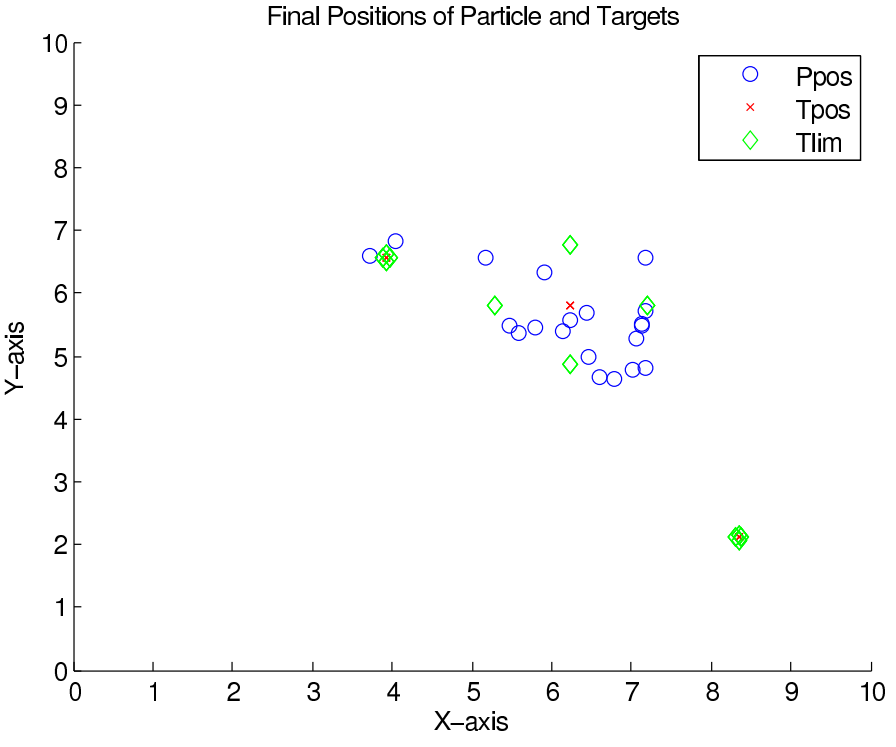


Fig. 5 Example Continued of Final Positioning for Single Fuzzy Parameter Method

$$v_{id} = w * v_{id} + c_1 * rand * (G_{id} - x_{id}) * t_n * d_{id} \tag{5}$$

The benefit of this method is that the velocity can more easily be controlled by the use of the "correct" membership function for d_{id}, without limiting the intelligence of the system, because there is no hard rules implemented. Where, the intelligence of the system is a non-measurable concept denoting how well the system is capable of adapting to new scenarios and demonstrating scientific reasoning. In addition, hard rules are rules which have no flexibility and always give the same result. There are no hard rules implemented in DFPM, because the DFPM makes use of an adaptive membership function. For example, if d_{id} has a triangular membership function, shown below, where the peak is the average of distances between all particles and a particular target, then the search velocity will be refined at distances much greater than and much less than the average distance. The fuzzy distance value is determined by fuzzy logic rules, which remain the same, but the definitions within the rules changes. This implies an adaptive fuzzy membership function. Fuzzy logic rules are used to build the membership function, and the fuzzy logic

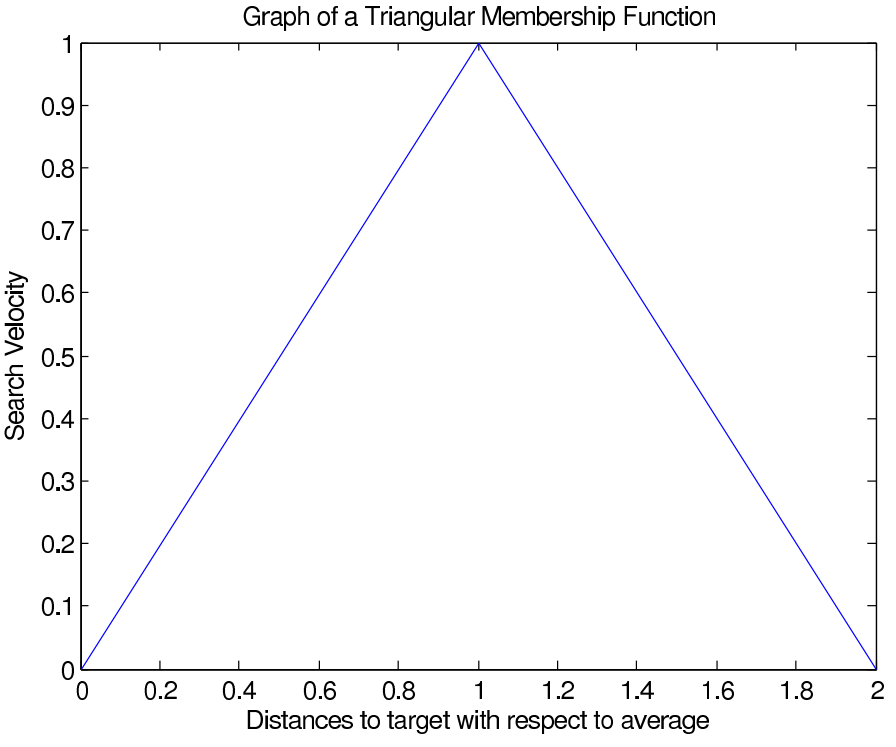


Fig. 6 Triangular Membership Function

rules presented below build the previously described triangular membership function.

The fuzzy logic rules applying to d_{id} are as follows:

1. *If the* ***distance to the target is much larger than the average****, then the* ***velocity is low*** *for a refined search.*
2. *If the* ***distance to the target is average****, then the* ***velocity is highest*** *for a non-refined search.*
3. *If the* ***distance to the target is much lower then the average****, then the* ***velocity is low*** *for a refined search.*

The benefit of using an adaptive membership function for the fuzzy distance parameter, d_{id}, is that it allows the DFPM to have increased *swarm intelligence*, to be discussed later in Section 11. The adaptive membership function for the fuzzy distance parameter, makes use of a Triangular Membership Function as seen in Figure 6.

Like the SFPM, if no target has been located the velocity function is defined by the following equation to guide a random search. The velocity

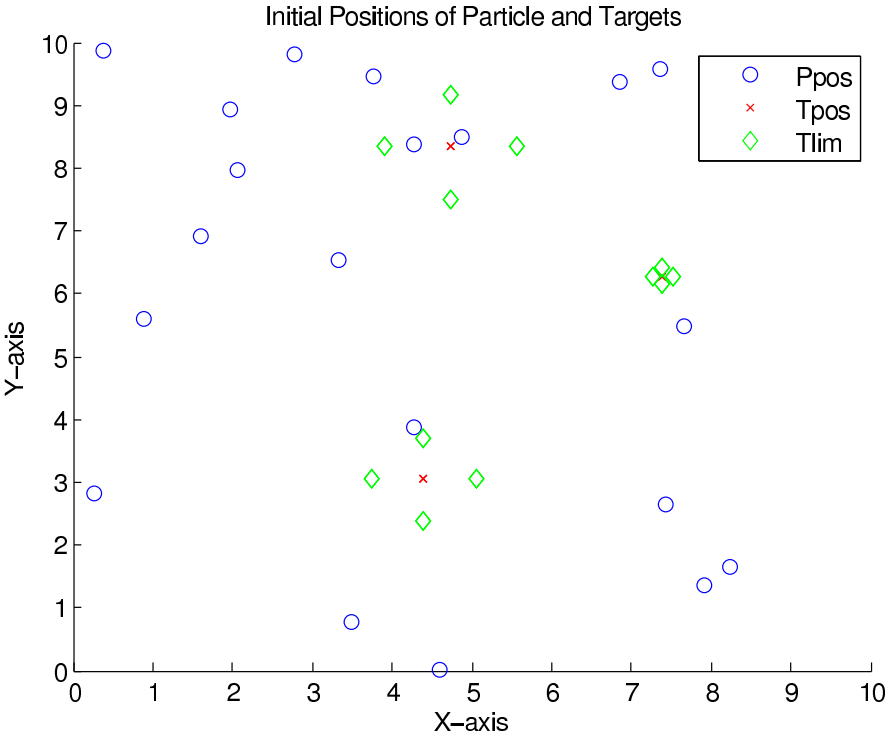


Fig. 7 Example of the Initial Positions prior to performing the Double Fuzzy Parameter Method

function presented below uses a uniform distribution for the random number generator.

$$v_{id} = -1 + 2 * rand() \tag{6}$$

As in the SFPM simulation, $Tpos$ is the center location of the target, where $Tlim$ is the boundaries of the target in the x and y direction. In the following simulation example, it is easily seen that there are two particles located within the boundary of the target approximately located at $(4.9, 8.25)$. Then there is another particle located along the boundary of the target approximately located at $(4.5, 3.0)$.

The current example demonstrates that all particles have effectively swarmed to a target location. Again, the particles are given a viewable radius of .25 from there current location. Thus, implying a particle can only view a target or a target boundary, $Tlim$, if that boundary is less then or equal to the .25 viewable radius. The smallest target centered approximately at $(7.5, 6.1)$ had three particles locate and converge upon it, which exemplifies

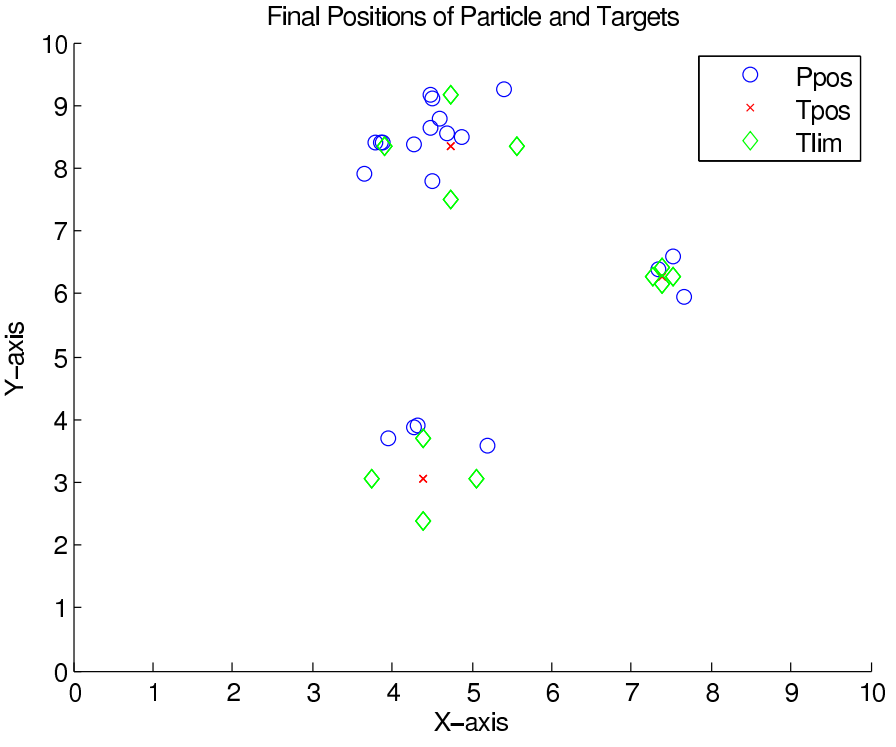


Fig. 8 Example Continued of Final Positioning for Double Fuzzy Parameter Method

that all targets were located. From the initial plot, it was known from the particles initializations that the other two targets would be located and converged upon, because particles were initialized within the boundary of the targets. It is seen that four particles converged upon the target located approximately at $(4.5, 3.0)$, and the remainder of the particles converged upon the target approximately located at $(4.9, 8.25)$.

To optimize the efficiency of the algorithm it would be best if each particle did not view ground area which was seen in the previous iteration. From analysis of (5) it can be seen that if a particle further away or closer then one standard deviation the value for d_{id} becomes ϵ, which $= .1$. The value for ϵ is developed to demonstrate a minimum value of outside one standard deviation and not derived by any method and could change as needed per application of the DFPM algorithm. This may allow the velocity vector calculation to become very small and as a result the particle would view much of the same area that was previously seen. As a result, a correction to the DFPM was

prompted and is corrected by (7), below. This corrected measure we call, Double Fuzzy Correction (DFC).

$$\texttt{if } v_{id} < .5 \ \forall x \in X \ \texttt{then } v_{id} = .5 \tag{7}$$

10 Double Fuzzy Corrected (DFC)

The fuzzy logic rules allow v_{id+1} to be very small, because the fuzzy distance variable, d_{id} may be ϵ, thus allowing particles to possibly view the same region as seen in the previous iteration. The objective of this study is to determine an optimal algorithm for efficiency and effectivity to be used in target determination and convergence. Since we are trying to obtain efficiency in algorithmic time each particle should converge as quickly as possible and this is accomplish by not allowing a particle to view any of the region as seen in the previous iteration. Therefore DFVM is modified with the above equation, (7). Each particle has a viewable radius of .25, hence each particle must move a minimum of .5 to remove the possibility of viewing any of the region which was previously seen. Therefore, (7) exists to make DFPM more efficient.

11 Swarm Intelligence

The above problem statement has also produced a better understanding of *swarm intelligence*. New to the development of swarm theory is the following understanding of swarm intelligence. Swarm intelligence refers to the overall intelligence of the swarm to compute an optimal solution, as opposed to *particle intelligence*. Particle intelligence is intelligence of the individual particles to determine a solution. It is based upon how efficiently an algorithm guides the particles to a solution. In swarm intelligence each individual particle is working for the swarm towards the swarm determining the best solution the swarm, by sharing and relating information. Whereas, particle intelligence the individual particles working towards a solution individually without knowledge of other particles.

An increased intelligence of the swarm does provide many advantages, because it allows an increase in adaptability as well as efficiency. The reason for an increase in adaptability extends from the swarm being more diverse to gather information learned by the swarm as opposed to information learned by the singular swarm best particle. The efficiency is increased because the swarm capable simultaneously seeking, determining and converging upon multiple targets.

According to John Nash, [12], the best result of a group is not accomplished by each individual particle in the group doing what is best for them, but making a compromises and doing what is best for the group, [8]. This

is very similar to the behavior of the fuzzy distance parameter, d_{id}, because one particle will take a small step in order for the majority of particles to take a large step, which does benefit the swarm as a whole. The benefit is the majority of the particle in the swarm will be taking a large step, and the few that do take a small step by comparison of d_{id} will be attempting to locate other targets. The drawback of this method is that not all particles are attempting to converge as quickly as they could. The small step is accomplished by the value of ϵ, in conjunction with (7).

The traditional PSO algorithm proposed by Eberhart and Kennedy demonstrates more particle intelligence than swarm intelligence, because each particle is only concerned about its current position in comparison with the best global position. The lack of swarm intelligence works sufficiently well, because there exists one global optimal solution for mono-objective optimization problem/model. The importance of the information preset by the swarm in c_1 and c_2, the cognitive and social parameters respectively, as a method of relating the importance of information by itself and the swarm best particle. This is a primitive form of intelligence, because it is preset. The problem proposed in the problem statement requires the swarm to be more aware of the surrounding particles, because each particle has multiple distances to target, because there may be more than one located target. To minimize the number of iterations for a solution to be found, it would be ideal for each particle to move in the direction of the closest target(s) and only the closest target(s). In conclusion, efficient and effective convergence for multiple targets simultaneously can only be accomplished if there is an increase in swarm intelligence by comparison to the traditional PSO algorithm by Eberhart and Kennedy, [3].

12 Results

The preset parameters: w, c_2, size of swarm population, and size of target population, remained constant throughout each of the three methods tested: Baseline, SFPM and DFPM. Each method was simulated ten times with a swarm population of 20, and a target population of 3. The weight coefficient, w, was set to .9 and linearly decreased by .1 with a minimum value of .3 after each iteration upon finding a target. The social parameter, c_2, was set to 2. The population of the swarm is 20, and the target population is set at 3 for every test in each of the three methods. From analysis of implementing (3) and (5), we know the larger c_2 is, the faster the convergence, and the smaller c_2, the slower the convergence. However a slower convergence would allow more locations to be search, therefore leading to an increase in possibility to locate a target. As seen in the figures for the SFPM and DFPM examples, the searchable region is a ten by ten region.

12.1 Baseline Method

The results for the Baseline Method demonstrate the average number of iterations until convergence was 8.5 iterations and the average number of targets found was 1.9 targets per simulation.

Table 1 Results of Baseline Method simulations, with Human Random

		Target Size(Particles at Target)			
Run	Targets Located	1	2	3	Iteration Count
1	2	**.4480**(11)	**.4788**(9)	.8794	14
2	2	**.9535**(15)	**.1676**(5)	.7767	10
3	3	**.3413**(3)	**.4866**(12)	**.3311**(5)	9
4	1	**.4588**(20)	.3541	.5596	10
5	2	**.7015**(18)	.2751	**.1990**(2)	12
6	1	.0493	**.7599**(20)	.0352	8
7	2	**.3839**(7)	.4033	**.7125**(13)	11
8	2	**.5706**(9)	.0886	**.7255**(11)	7
9	1	.1101	**.9805**(20)	.2835	8
10	2	**.1993**(6)	**.7696**(14)	.4034	7
Bold target size denotes the target was located.					

12.2 SFPM

The results for the SFPM demonstrate convergence was achieved in 7.4 iterations and an average of 1.7 targets were located.

Table 2 Results of SFPM simulations, with Human Random

		Target Size(Particles at Target)			
Run	Targets Located	1	2	3	Iteration Count
1	3	**.9650**(6)	**.9794**(6)	**.5891**(8)	7
2	1	.0590	**.7823**(20)	.0187	14
3	3	**.8070**(7)	**.3174**(2)	**.2988**(11)	9
4	2	.0023	**.2385**(5)	**.6473**(15)	9
5	1	.3971	.3774	**.7225**(20)	8
6	3	**.5006**(3)	**.9223**(8)	**.9228**(9)	6
7	1	.2384	.0019	**.0748**(20)	18
8	1	.3413	**.9252**(20)	.0748	10
9	2	.1443	**.4932**(18)	**.1210**(2)	13
10	2	**.5899**(7)	**.4497**(13)	.0951	10
Bold target size denotes the target was located.					

12.3 DFPM

The DFPM was performed with the Double Fuzzy Correction, DFC.

Table 3 Results of DFPM simulations, with Human Random

		Target Size(Particles at Target)			
Run	Targets Located	1	2	3	Iterations
1	2	**.4293**(2)	.2081	**.6283**(18)	21
2	3	**.9425**(9)	**.6612**(6)	**.5663**(5)	7
3	3	**.2496**(5)	**.4293**(3)	**.4529**(12)	15
4	2	.1167	**.9646**(13)	**.2057**(7)	21
5	1	**.9344**(20)	.0998	.2539	8
6	3	**.6739**(4)	**.7979**(10)	**.4550**(6)	10
7	1	**.7995**(20)	.7168	.0690	11
8	2	.4683	**.3091**(11)	**.3539**(9)	13
9	2	**.4561**(12)	**.7930**(8)	.1394	11
10	2	**.6745**(8)	**.5890**(12)	.0379	13
Bold target size denotes the target was located.					

The results for the DFPM demonstrate the average number of iterations to convergence was 13.0, and the average number of targets located was 2.1.

13 Comparison

From the results, shown above, it is easily seen that the DFPM has the highest occurrence rate of determining more targets per simulation than the SFPM and Baseline Method algorithms. The higher occurrence rate is due to the slower convergence of the particles throughout the swarm. Hence, particle will have slow convergence if it is not approximately the median distance to the target as the other particles in the swarm. Many of the particles were outside of the membership function, and therefore those particles were moving the minimum allowed due to the Double Fuzzy Correction (DFC). Thus, by allowing the particles to search the region more precisely more targets were located. The drawback of this method is that the convergence rate was more than double that of the SFPM, and over one and half times that of the Baseline Method.

In the DFPM, the velocity step calculation is approximately equivalent to the velocity step calculation used by all particles in the SFPM, when the particles are near the median of the *fuzzy distance parameter*. The random value in the velocity function is thought to not effect the convergence of the particles, because it is uniformly random, meaning there is no preference for a higher or lower value. When comparing the two algorithms, it can be thought of as null or non-influential. Therefore, the particle's personal position with

respect to the median distance will change, being the only difference factors between the SFPM and DFPM algorithms.

14 Conclusion

In conclusion, it is shown on average that the DFPM did take the most iterations to converge, however it did show that it did locate more targets. Therefore each algorithm could be used dependent upon which objective was of higher priority. If convergence time was of more importance then the SFPM should be used and vice versa for number of targets with the DFPM. When comparing the SFPM with the Baseline Method it can be seen that the addition of a fuzzy variable did increase the iteration time, with little change to the number of targets located. In the SFPM velocity function the fuzzy size parameter is used, thus making the convergence to located target more dynamic and the likelihood of determining another target was not high therefore a faster convergence could be accepted with minimal loss to the number of located targets upon convergence.

We feel that a large-scale real-world problem could be solved using this methodology. However, further studying and enhancements would be needed to understand the algorithms's advantages and drawbacks. The concept of swarm intelligence is an important aspect of this research problem, because it builds upon the idea of a collection of particles being agents working toward a common goal with the possibility of independent goals.

References

[1] Saber, A., et al.: Unit commitment computation - a novel fuzzy adaptive particle swarm optimization approach. In: Proceedings of the Power Systems Conference and Exposition, October 2006. IEEE, Los Alamitos (2006)

[2] Armstrong, S.A.: A meta-analysis of randomness in human behavoiral research. Master's thesis, Louisianna State University, Baton Rouge, Louisiana (August 2004)

[3] Eberhart, R., Kennedy, J.: A new optimizer using tarticle swarm theory. In: Proceedings of the Sixth International Symposium on Micro Machine and Human Science, pp. 39–43. IEEE, Los Alamitos (1995)

[4] Eberhart, R., Shi, Y.: Tracking and optimizing dynamic systems with particle swarms. In: Proceedings of the IEEE International Congress on Evolutionary Computation. IEEE, Los Alamitos (2001)

[5] Goldberg, D.: Genetic Algorithm *in Search Optimization, and Machine Learning.* Addison-Welsey, Inc., Reading (1989)

[6] Goodhew, I., Hutt, B., Warwick, K.: Control and experimentation of a personal robot tracking system. International Journal of Modelling, Identification and Control 1(1), 4–12 (2006)

[7] Halder, A., Machadevan, S.: Probability, Reliability and Statistical Methods in Engineering Design. John Wiley and Sons, Inc., New York (2000)

[8] Holt, C., Roth, A.: The nash equilibrium: a perspective. 101(12), 3999–4002 (2004)
[9] Howard, C., Hanson, C., Zander, A.: Vibro-acoustic noise control treatment for payload bays of launch vehicels: distrete to fuzzy systems (June 2005)
[10] Hu, X., Eberhart, R.: Solving constrained nonlinear optimization problems with particle swarm optimization. In: Proceedings of the Sixth World Multiconference on Systemics, Cybernetics and Informatics 2002 (SCI 2002), pp. 203–206 (2002)
[11] Hu, X., Eberhart, R., Shi, Y.: Engineering optimization with particle swarm. In: Proceedings of the IEEE Swarm Intelligence Symposium, pp. 53–57. IEEE, Los Alamitos (2003)
[12] Nash Jr., J.F.: Equilibrium points in an n-person games. Proceedings of the National Academy of Sciences 36, 48–49 (1950)
[13] Linhardt, R., Ajayan, P., Nalamasu, O.: Flexible energy storage devices based on nanocomposite paper. Proceedings of the National Academy of Sciences (2007)
[14] Nguyen, H.T., Kreinovich, V.: Classical-logic analogue of a fuzzy "paradox". In: Proceedings of the Fifth IEEE International Conference on Fuzzy Systems, September 1996, vol. 1, pp. 428–431 (1996)
[15] Nise, N.: Controls System Engineering, 4th edn. John Wiley and Sons, Inc., New York (2004)
[16] Media Relations Office of NASA. Bright chunks at Phoenix Lander's Mars site must have been ice (June 19, 2008)
[17] Media Relations Office of NASA. NASA press release on Mars exploration strategy (October 26, 2000)
[18] Passino, K.M., Yurkovich, S.: Fuzzy Control. Addison-Wesley Longman, Inc., Amsterdam (1998)
[19] Perez, R., Behdinan, K.: Swarm Intelligence, Focus on Ant and Particle Swarm Optimization. In: Particle Swarm Optimization in Structural Design, ch. 21, pp. 373–394. Tech Education and Publishing, Vienna (2007)
[20] Richards, Z., Salters, R.: Fuzzy optimal swarm for an optimal swarm of autonomous aircrafts. In: Proceedings of the Twelfth AIAA Conference on Multidisciplinary Optimization and Analysis, September 2008. American Institute of Aeronautics and Astronautics (2008)
[21] Shi, Y., Eberhart, R.: Fuzzy adaptive particle swarm optimization. In: Proceedings of the IEEE Congress on Evolutionary Computation. Institute of Electrical and Electronics Engineers, Los Alamitos (2001)
[22] Venkataraman, P.: Applied Optimization with MATLAB Programming. John Wiley and Sons, Inc., New York (2002)
[23] Wang, L., Kang, Q., Wu, Q.: Grouped-and-delayed broadcasting mechanism for optimum information in particle swarm optimization. In: Proceedings of the Conference on Cybernetics and Intelligent Systems, December 2004. Institute of Electrical and Electronics Engineers (IEEE), Los Alamitos (2004)
[24] Watada, J.: History of fuzzy multivariant analaysis in 40 years. Fuzzy Logic, Soft Computing, and Computational Intelligence 3, 1509–1514 (2005)
[25] Winfield, A., Nembrini, J.: Safety in numbers: fault tolerance in robot swarms. International Journal of Modelling, Identification and Control 1(1), 30–37 (2006)

[26] Zhang, Q., Li, X., Tran, Q.: A modified particle swarm optimization algorithm. In: Proceedings of the Fourth International Conference on Machine Learning and Cybernetics (August 2005)
[27] Zitzler, E., Laumanns, M., Bleuler, S.: A tutorial on evolutionary multiobjective optimization. In: Kenneth, T'kindt, V. (eds.) Metaheuristics for Multiobjective Optimisation. LNEMS, vol. 535, pp. 3–37 (2004)
[28] Zurada, J., Marks II, R., Robinson, C.: Computational Intelligence: Imitating Life. Institute of Electrical and Electronics Engineers (IEEE), Los Alamitos (1994)

Fuzzy Linear Programming in Practice: An Application to the Spanish Football League

J.M. Cadenas, V. Liern, R. Sala, and J.L. Verdegay

Abstract. FLP problems are perhaps one of the most and best studied topics of Soft Computing, and are among the most fruitful in applications and in theoretical and practical results. Areas of application of FLP problems are many and varied and in fact suppose an extraordinary example of technology transfer in action. In this paper, Fuzzy Linear Programming models are applied to the European football game in which the inherent uncertainty of the parameters relating to the football teams in the Spanish Football League serve to establish these models and so optimize the returns of the investments made to maintain a high quality competition. In this context, fuzzy DEA models are established which provide teams predictions as to their efficiency score. At the end of the study we offer some experiments using data from the Spanish Football League 2006/07.

1 Introducction

The broad area of Soft Computing rests on two large scientific fields: Approximate Reasoning and Functional Approximation/Randomized Search, which are directly related to the two most important activities associated with Intelligent

J.M. Cadenas
Dept. Ingeniería de la Información y las Comunicaciones
University of Murcia, Murcia, Spain
e-mail: jcadenas@um.es

V. Liern · R. Sala
Dept. de Economía Financiera y Matemàtica
University of Valencia, Valencia, Spain
e-mail: vicente.liern@uv.es, ramon.sala@uv.es

J.L. Verdegay
Dept. of Computer Science and Artificial Intelligence
University of Granada, Granada, Spain
e-mail: verdegay@decsai.ugr.es

W.A. Lodwick & J. Kacprzyk (Eds.): Fuzzy Optimization, STUDFUZZ 254, pp. 503–528.
springerlink.com

Systems - decision making and search processes (classification/optimization). Both areas also give rise to the four basic constituents of Soft Computing: Probabilistic Models, Fuzzy Sets and Systems, Neural Networks and Metaheuristics [45], the latter being understood as the set of optimization/search methods which can solve real world problems. It is in this context that we find Fuzzy Linear Programming (FLP) problems, which are the fruit of the hybridizing of models and techniques hailing from Fuzzy Sets and Systems and optimization methods. FLP problems (along with the fuzzy control) are perhaps one of the most and best studied topics of Soft Computing, and are among the most fruitful in applications and in theoretical and practical results. Areas of application of FLP problems are many and varied and in fact suppose an extraordinary example of technology transfer in action.

Yet despite the many applications, the relevance of the results (theoretical and applied) and the wide number of related fields, to the best of our knowledge FLP models have never been applied to the sphere of the European football (or soccer) game. This is done here, where the uncertainty inherently associated to the parameters relating to soccer teams in the Spanish league finds the best possible theoretical context from the concepts pertaining to Fuzzy Sets and Systems. These serve to establish FLP models which optimize the returns on investments made to maintain a high quality competition, which is finally given in an efficiency measure of the different teams, which can be classified. With this aim, initially we tackle the basic theoretical elements necessary to this article. While not descending to the trivial level, we present the most elementary ideas on fuzzy sets and numbers. Below, we put forward (within the fuzzy context) the most typical problems and methods in FLP. Next, we show an application of FLP in a specific domain - the problem of the Spanish professional football league. We give the model for that problem and many ways of solving it Finally, we show the most important conclusions of the study.

2 Basic Concepts of Fuzzy Sets

One basic concept is that of the fuzzy number. From the point of view of a fuzzy number's being a fuzzy set in $\mathbb{R}$, it can be stated that the notion of a fuzzy number appears in 1965 with the appearance of L.A. Zadeh's famous paper [47].

Nevertheless, fuzzy numbers really appear on the scene around 1978, with the papers by S. Nahmias on fuzzy variables, and D. Dubois and H. Prade on handling imprecise quantities. Since then, the study of possible definitions of fuzzy numbers and, in particular, how to manage and compare them, has aroused a lot of interest within the field of fuzzy sets, [46].

This section introduces the elementary notions and operations of fuzzy sets leading to the concept of fuzzy number. Once these have been established, the remaining part of this section is devoted to the comparison of two fuzzy numbers. This is a complex problem since, given the imprecise nature of the quantities considered, e.g. A and B, it cannot be guaranteed a priori that $A \leq B$ or that $B \leq A$. Instead, these properties will be verified simultaneously and with certain degrees of fulfilment. This means that there are many ways of comparing two fuzzy numbers, which

in the specialist literature has been developed through the so-called comparison indices.

2.1 Introduction to the Fuzzy Set Concept

Let X be a set, whose elements are we will denote by x, and A a subset of X. The membership of one element x of X to the subset A is given by the characteristic function:

$$\mu_A(x) = \begin{cases} 1 & \text{iff } x \in A \\ 0 & \text{iff } x \notin A \end{cases}$$

where $\{0,1\}$ is the so-called valuation set.

If the valuation set is the real interval $[0,1]$, A is called a fuzzy set ([47]) and μ_A measures the degree of membership of element x in A. A is characterised by the the set of pairs $\{(x, \mu_A(x)),\ x \in X\}$.

Two fuzzy sets, A and B are considered equal iff: $\forall x \in X,\ \mu_A(x) = \mu_B(x)$.

Definition 1. *([47]) Given a fuzzy set $A = \{(x, \mu_A(x))\}$, its support is defined as the ordinary set $Supp(A) = \{x \in X \ / \ \mu_A(x) > 0\}$.* □

Definition 2. *([47]) Given a fuzzy set A, we give the name α-cut of that set to the ordinary set $A_\alpha = \{x \in X \ / \ \mu_A(x) \geq \alpha\}$ con $\alpha \in [0,1]$.* □

It is clearly seen that the sets A_α, $\alpha \in [0,1]$ constitute a decreasing succession. If $\alpha_1 \geq \alpha_2 \Leftrightarrow A_{\alpha_1} \subseteq A_{\alpha_2},\ \alpha_1, \alpha_2 \in [0,1]$.

Theorem 1. *(Representation Theorem) If A is a fuzzy set and A_α its α-cuts, $\alpha \in [0,1]$, it is verified that*

$$A = \bigcup_{\alpha \in [0,1]} \alpha A_\alpha$$

taking this formal notational as the equality between the membership functions of both sets. If $\mu_{A_\alpha}(x)$ denote the characteristic function of A, a particular case of the membership function:

$$\mu_{A_\alpha}(x) = \begin{cases} 1 & \textit{iff } x \in A_\alpha \\ 0 & \textit{otherwise} \end{cases}$$

membership function of the fuzzy set A can be expressed in terms of the characteristic functions of its α-cuts, according to the formula

$$\mu_A(x) = \sup_{\alpha \in [0,1]} \min(\alpha, \mu_{A_\alpha}(x))$$ □

Definition 3. *([47]) A fuzzy set is convex iff its α-cuts are convex.* □

A definition equivalent to convexity is that A is convex iff $\forall x_1, x_2 \in X, \forall \lambda \in [0,1]$, $\mu_A(\lambda x_1 + (1-\lambda)x_2) \geq \min(\mu_A(x_1), \mu_A(x_2))$.

Definition 4. *The height of a fuzzy set $hgt(A) = \sup_{x \in X} \mu_A(x)$.* □

Definition 5. *A fuzzy set is said to be normalized iff $\exists x \in X$ in which $\mu_A(x) = 1$.* □

2.2 Fuzzy numbers

Definition 6. *[18] A fuzzy number A is a μ_A set of the real straight, it is convex and normalized such that*

a) $\exists x_0 \in \mathbb{R} i \ / \ \mu_A(x_0) = 1$, *which is generally called mode, and*
b) μ_A *is in parts continuous.* □

Any fuzzy number is therefore characterised by a membership function $\mu_A : \mathbb{R} \rightarrow [0,1]$ and any function like the above gives a fuzzy number where $\forall x \in \mathbb{R}$, $\mu_A(x)$ is the degree of membership of x to the fuzzy number A.

We will denote by $F(\mathbb{R})$ the set of membership functions on $\mathbb{R}$. Thus, when talking about the fuzzy number we can refer both to the element $A \in F(\mathbb{R})$ and to $\mu_A \in F(\mathbb{R})$.

A fuzzy number A is said to be of the type $\mathscr{L} - \mathscr{R}$, if and only if its membership function μ_A is of the form

$$\mu_A(x) = \begin{cases} \mathscr{L}\left[\frac{(m-x)}{\alpha}\right] & \text{for } x \leq m,\ \alpha > 0 \\ \mathscr{R}\left[\frac{(x-m)}{\beta}\right] & \text{for } x \geq m,\ \beta > 0 \end{cases}$$

where m is the mode of A and α (β) is the width on the left (right), $\mathscr{L}$ and $\mathscr{R}$ represent a function on the left or right of m, $\mathscr{L}$ is non decreasing and $\mathscr{R}$ is not increasing. We will abbreviate the fuzzy number A by $A = (m-\alpha, m, m+\beta)_{\mathscr{L}R}$.

Definition 7. *[19] A plane fuzzy number is an A fuzzy number such that*

$$\exists (m_1, m_2) \in \mathbb{R} i,\ \ m_1 < m_2 \text{ and } \mu_A(x) = 1,\ \ \forall x \in [m_1, m_2]$$ □

A plane fuzzy number can model a fuzzy interval. An A plane fuzzy number of type $\mathscr{L} - \mathscr{R}$ is defined as

$$\mu_A(x) = \begin{cases} \mathscr{L}\left[\frac{(m_1-x)}{\alpha}\right] & \text{for } \ \leq m_1,\ \alpha > 0 \\ \mathscr{R}\left[\frac{(x-m_2)}{\beta}\right] & \text{for } x \geq m_2,\ \beta > 0 \\ 1 & \text{for } m_1 \leq x \leq m_2 \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

This will be more briefly denoted by $(m_1 - \alpha, m_1, m_2, m_2 + \beta)_{\mathscr{L}R}$.

It is clear that depending on the $\mathscr{L}$ and $\mathscr{R}$ functions, we will obtain different types of fuzzy numbers.

We will consider numbers as fuzzy, plane, linear and normalised, those whose analytical membership function is as follows.

A plane fuzzy number, which we will denote by $\tilde{u}_j = (r_j, \underline{u}_j, \overline{u}_j, R_j)$ will have the membership function

$$\forall v \in \mathbb{R},\ \ \mu_{\tilde{u}_j}(v) = \begin{cases} \frac{(v-r_j)}{(\underline{u}_j - r_j)} & \text{if } r_j \le v \le \underline{u}_j \\ \\ \frac{(R_j - v)}{(R_j - \overline{u}_j)} & \text{if } \overline{u}_j \le v \le R_j \\ \\ 1 & \text{if } \underline{u}_j \le v \le \overline{u}_j \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

From now on we will frequently use fuzzy numbers expressed as linear combinations $\tilde{y} = \sum_j \tilde{u}_j x_j$ with $x_j \in \mathbb{R},\ j = 1, \ldots, n$.

In [42] we find the membership function of those numbers, which we express below.

Proposition 1. *If $\tilde{y} = \sum_j \tilde{u}_j x_j = \tilde{u}x$ is a linear expression in which the $\tilde{u}_j,\ j = 1, \ldots, n$, are fuzzy numbers linear membership functions given by $\tilde{u}_j = (r_j, \underline{u}_j, \overline{u}_j, R_j)$ and $x_j \ge 0,\ j = 1, \ldots, n$, then the membership function of $\tilde{y}$ is*

$$\mu(z) = \begin{cases} \frac{(z - rx)}{(\underline{u}x - rx)} & \text{if } x > 0 \text{ and } rx \le z \le \underline{u}x \\ \\ \frac{(Rx - z)}{(Rx - \overline{u}x)} & \text{if } x > 0 \text{ and } \overline{u}x \le z \le Rx \\ \\ 1 & \text{if } \underline{u}x \le z \le \overline{u}x \\ 0 & \text{otherwise} \end{cases}$$

where $r = (r_1, \ldots, r_n)$, $\underline{u} = (\underline{u}_1, \ldots, \underline{u}_n)$, $\overline{u} = (\overline{u}_1, \ldots, \overline{u}_n)$ and $R = (R_1, \ldots, R_n)$. □

2.3 Methods for Comparing Fuzzy Numbers

A constant problem over the last years has been that of the distribution of imprecise quantities, and hence the comparison of fuzzy numbers. The many and varied approaches to the problem mean that a wide range of methods exist to make the comparison in question. An excellent collection of techniques, methods and approaches can be found in [46, 51].

We will use the ways of comparing fuzzy numbers exclusively to analyse the repercussion of using various methods of comparison in a Fuzzy Linear Programming problem. Thus, it is not our aim here to review all the possible ways of comparing.

The solution to the problem can be shortened in either of the following ways, depending on whether the method used is based on the definition of an ordering function or on the comparison of alternatives.

2.3.1 Methods Based on the Definition of an Ordering Function

We will consider $A, B \in F(\mathbb{R})$. A simple method to compare these lies in the definition of a function $g : F(\mathbb{R}) \to \mathbb{R}$. If the function $g(\cdot)$ is known, then

$$g(A) < g(B) \Leftrightarrow \text{ A is less than B}$$
$$g(A) > g(B) \Leftrightarrow \text{ A is greater than B}$$
$$g(A) = g(B) \Leftrightarrow \text{ A is equal to B}$$

g is usually called a linear ordering function if

1) $\forall A, B \in F(\mathbb{R})$, $g(A+B) = g(A) + g(B)$
2) $\forall r \in \mathbb{R}$, $r > 0$, $g(rA) = rg(A)$, $\forall A \in F(\mathbb{R})$

In this case, the indices can be classified according to whether the ordering function is linear or not.

2.3.2 Methods Based on the Comparison of Alternatives

These methods consist of obtaining the fuzzy set of the optimal alternatives:

$$\tilde{O} = \{i, \mu_{\tilde{O}}(i)\},\ \mu_{\tilde{O}}(i) = \mu_{\tilde{O}}(A^i),\ A^i \in F(\mathbb{R}i)$$

where $\mu_{\tilde{O}}(i)$ represents the degree to which the ith alternative can be considered the best.

Finally, we underline in spite of the huge wealth of methods for comparing fuzzy numbers, as yet few indices have been studied since it is perfectly justifiable for each human decision taker to use their own method of comparison independently of any method described in the literature. A detailed study in this respect can be found in [32] where an artificial neuronal network is used which learns the ordering criteria of each decision taker considered.

3 Methods and Models of Fuzzy Linear Programming

An LP problem is generally set out as Max $\{z = cx / Ax \leq b, x \geq 0\}$ where A is matrix $m \times n$ of real numbers, $b \in \mathbb{R}^m$ and $c \in \mathbb{R}^n$.

Obviously, it is assumed here that the decision taker has exact information on the elements intervening in the problem. Even were this the case, the decision taker usually finds it more convenient to express his knowledge in linguistic terms, i.e. through conventional linguistic labels [48], rather than by using high precision numerical data. Thus, it makes perfect sense to talk about optimization problems from a vague predicate approach as it is understood that this vagueness arises from the way we use to express the decision taker's knowledge and not from any random event. In short, it is supposed that the imprecision of the data defining the problem is fuzzy.

The first case of optimization problems with fuzzy approach appeared in the literature more almost four decades ago [2], in an article which put forward the now classical key concepts of constraint, objective and fuzzy optimal decision.

As with LP in conventional optimization, so have FLP methods been the subject of most study in the fuzzy context. While not exhaustive, there are three main types of FLP problems, depending on the imprecision established in the constraints, on the

coefficients of the technological matrix or on the costs which define the objective function. Models and methods to solve these problems abound in the literature [4, 5, 6, 16, 17, 23, 28]. In some cases precise solutions are obtained, while in others these are fuzzy and more in line with the approach to the problem. The latter offer a set of good alternatives and encompass the more precise solutions obtained using other methods. Finally, it is the decision maker who must choose.

3.1 Several Types of Fuzzy Linear Programming Problems

3.1.1 Fuzzy Constraint

We consider the case in which the decision maker assumes that there is a certain tolerance in the fulfilment of the constraints. The associated problem is represented as follows

$$\text{Max} \ \{z = cx \ / \ Ax \lesssim \ b, \ x \geq 0\} \tag{3}$$

where $c \in \mathbb{R}^n$, $b \in \mathbb{R}^m$, A is a matrix $m \times n$ of real numbers, and each constraint can be modelled using a membership function

$$\mu_i : \mathbb{R} \to [0,1] \ \ / \ \ \mu_i(a_i x) = \begin{cases} 1 & \text{if } \ a_i x \leq b_i \\ f_i(a_i x) & \text{if } \ b_i \leq a_i x \leq b_i + t_i \\ 0 & \text{if } \ a_i x \geq b_i + t_i \end{cases}$$

This problem was addressed and solved in [41, 50], and later it was generalized in [44] to obtain a fuzzy solution involving the particular point-solutions provided by the methods shown in [41, 50].

3.1.2 Fuzzy Costs

In this case, the decision taker does not know the exact values of the coefficients c. The situation is represented by the following FLP problem.

$$\text{Max} \ \{z = \tilde{c}x \ / \ Ax \leq \ b, \ x \geq 0\} \tag{4}$$

with $\tilde{c} \in (F(\mathbb{R}))^n$ a vector of fuzzy numbers and supposing membership functions defined for each cost. There are various approaches [16, 33, 42] to solve (4). The method proposed in [16] gives a formal context to find the solution of (4) and encompasses the methods proposed in [33, 42, 28].

3.1.3 Fuzzy Numbers in the Technological Matrix

Now we consider that the coefficients in the technological matrix and the coefficients of the right hand size are represented by fuzzy numbers, with the costs that define the objective function being real.

This type of FLP problem is set out in the following terms

$$\text{Max } \{z = cx \;/\; \tilde{a}_i x \lesssim \tilde{b}_i,\ x \geq 0,\ i \in M\} \tag{5}$$

where for each $i \in M$, $\tilde{a}_i = (\tilde{a}_{i1}, \ldots, \tilde{a}_{in})$, $\tilde{a}_{ij} \in F(\mathbb{R})$, $j \in N$, $\tilde{b}_i \in F(\mathbb{R})$, $x \in X = \{x \in \mathbb{R}^n \;/\; \tilde{a}_i x \lesssim \tilde{b}_i, i \in M, x \geq 0\}$ and $c \in \mathbb{R}^n$.

The first version of this problem appeared in [42]. In order to find (5) a fuzzy solution, in [17] a method is presented which supposes first that violations in the accomplishment of its constraints are permitted up to a maximum amplitude of a fuzzy value, and second that in order to compare the left hand side to the right hand one a fuzzy comparison relation between fuzzy numbers is to be considered.

3.1.4 A General Model

A general FLP model, [4], which encompasses all the above cases, is a problem of the type:

$$\text{Max } \{z = \tilde{c}x \;/\; \tilde{a}_i x \lesssim \tilde{b}_i,\ x \geq 0,\ i \in M\} \tag{6}$$

where $\tilde{c} \in (F(\mathbb{R}))^n$, and for each $i \in M$, $\tilde{a}_i \in (F(\mathbb{R}))^n$, $\tilde{b}_i \in F(\mathbb{R})$.

One method [4] for solving this general model 6 consists of substituting the constraints set by a convex fuzzy set, in the same way as described above.

4 An Application to the Spanish Football League

Experts and managers of professional Football League are conscious that the statistical data of the results, referee decisions, etc. are insufficient. They need to know if the teams use their inputs properly and how increase their outputs. Taking into account this requirement, we present an analysis of the technical efficiency of the Spanish Football League teams (season 2006/07). Actually, we consider the following objectives:

a) Knowing the offensive and defensive (at home, away and general) of each team.
b) Obtaining the weights of the optimal combination of inputs that would make efficient to a given team.
c) Knowing, for each non-efficient team, which teams we should compare with.
d) Ranking the teams by its efficiency.

For our proposal, the use of linear fuzzy programming is useful because the models show uncertainty for several reasons: the data are not precise because they are collected by non-professional evaluators, sometimes there are plays that are recorded to one team which should be recorded to the other (own goal shot, for example), referee decisions, etc. Besides, the efficiency analysis is used to rank the teams and make future decisions.

In the next sections we describe the characteristics of the problem, the procedure of modeling the situation and some ways of solving them.

4.1 Spanish Football League

Football is a competitive sport with two teams of 11 players. The winning team must score more goals than the other team during the game. While there are various styles of competition, the national leagues involve each of the teams (20 in Spain) playing each other during the season. Each team plays every other team twice; once in its own "home" ground, and once "away" in the opponent's ground. Victories are rewarded with three points, draws receive one point, and defeats do not receive any pointstherefore it is a non-zero sum game. The team with the most points wins the league and any ties at the end of the season are broken in various ways in each league.

There are incentives for occupying the highest possible position in the league at the end of the season. Of course, if being first were the only criteria then teams would lose their incentive to win when they realise that first place had become unachievable. Therefore, the best-positioned teams are rewarded with the opportunity to play in European competitions in the following season, and the lowest teams are relegated to the league below.

The basic characteristics of football mean that it can be analysed just like any productive activity. Consequently, it is possible to use the idea of a "sports production function", as first proposed by Rottenberg [34] when discussing baseball. Later, Scully [37] offered an initial empirical estimation of production functions for the same sport. These pioneering works were followed by others centred in other sports, such Zak et al. [49], or Scott et al. [36] in basketball, Schofield [35] in league cricket, and Sloane [39, 40] in football. More recent papers, such as Carmichael and Thomas [8], or Carmichael et al. [9], or Dawson et al. [15], have been focussed on the production function and efficiency analysis applied to rugby and the English Premier football league.

The productive process in football can be modelled formally using the following production function:

$$Y_i = Y_i(X_i), \quad i = 1, 2, \ldots, n,$$

where Y_i is the football output measured for team i (usually the percentage of points or victories obtained out of the possible total, or the difference between goals scored and conceded) and X_i is a vector of inputs. Usually, the inputs in the production function are variables that measure the technical abilities of the players. Many different methods of measurement of outputs and inputs can be found in the literature, depending on the object being studied. It is possible to consider each league game as a reference unit (see [9]), where output is measured as goal difference, and input is measured as the difference between the two teams in a series of game statistics (for example, the difference between shots-on-goal, the number of balls controlled with the first touch, etc.).

Another approach consists in considering aggregated output for a team during one, or various leagues (see [15]). In this case, output is normally measured as the percentage of points or victories obtained; the inputs are the various indicators of each teams footballing ability (for example, the amount of possession, number of centres, etc.).

These studies share two common denominators, regardless of the econometric techniques used in each case (OLS, fixed effects, random effects, maximum

likelihood, etc.). Firstly, they attempt to quantify the relative importance of each of the inputs in the production of the output and this enables conclusions to be made about which technical aspects of play are most important. Secondly, they attempt to explain why, once the various productive factors have been controlled for, the productivity of each team apparently differs[1].

In the economic literature, the issue is often tackled by assuming that some firms have access to an advanced technology that enables them to obtain higher levels of output from a given volume of resources. Therefore, the standard practice consists in estimating production functions usually with parametric techniques, that allow the recovery of homogeneous output elasticities for all firms, and to control for the presence of level effects for each firm, which are interpreted as indicators of the level of technological development in each firm (for example, estimating fixed effects models).

However, two firms with the same technology may have differing productivities because the quality of their management varies, or because the efficiency of their organisation differs (meaning the efficiency of their productive processes). This second possibility offers a more realistic explanation for productivity differences between companies in homogeneous sectors, or when the technology used is well-known and established.

In this paper, we will assume that clubs have access to the same level of technology, but differ in their levels of efficiency, and that this may explain differences in productivity[2]. It seems reasonable to assume that the technology used in football (tactics, plans, physical training, etc.) is homogeneous and basically wellknown among industry professionals. For this reason, non-parametric optimisation techniques, specifically DEA[3] models, are used. Non-parametric methodology offers great flexibility and an absence of specification errors because it is not necessary to choose any particular functional form. However, it suffers the disadvantage of being technically deterministicand so atypical observations may bias the efficiency results and attribute any random shocks to inefficiency.

Additionally, our analysis of efficiency will take into account other important characteristics of the productive process in football. One of the most important problems in the interpretation of results derived from parametric estimations of the production functions of football is the following: the measurements of output combine offensive productivity (goal-scoring) with defensive efficiency (preventing goals). As a consequence, the inputs used in the measurements are a combination of indicators of each club's offensive and defensive ability, and so the signs expected in the regressions for some inputs will be positive, while others will be negative, or indeterminate. For this reason, the calculations of the frontier of efficient production presented in the following pages will explicitly distinguish between offensive

[1] In other words, they try to explain why one team obtains more points or victories than anothereven if it has a team with identical abilities (identical inputs).

[2] An interesting discussion about the advantages and disadvantages of parametric and non-parametric techniques when studying efficiency and productivity can be found in [31].

[3] An introduction to Data envelopment analysis (DEA) models can be found in [14].

and defensive production and, thus, will enable to calculate separately offensive and defensive efficiency indicators.

4.2 DEA Models

Data Envelopment Analysis (DEA) is a methodology that has been widely used to evaluate the relative efficiency of a set of decision-making units (DMUs) involved in a production process. DEA models provide efficiency scores that assess the performance of the different DMUs in terms of either the use of several inputs or the production of certain outputs (or even simultaneously). Most of DEA efficiency scores vary in $(0, 1]$, the unity value being reserved to efficient units. In the particular case of the radial models, the CCR (Charnes, Cooper and Rhodes [11]) and the BCC (Banker, Charnes and Cooper [1]) models yield efficiency scores both in input and in output orientation, although nonoriented DEA efficiency scores can also be defined (see [21] for hyperbolic measures and [3, 10] for directional measures). The radial efficiency scores represent either equiproportionate input reductions or output expansions (or both simultaneously in the nonoriented case). Nevertheless, nonradial efficiency measures are also available (see e.g., [22] for the Russell measures).

Traditionally, the coefficients of DEA models, i.e., the data of inputs and outputs of the different DMUs, are assumed to be measured with precision. However, as some authors point out (see, e.g., [26]), this is not always possible. In these cases it may be more appropriate to interpret the experts understanding of the parameters as fuzzy numerical data which can be represented by means of fuzzy numbers or fuzzy intervals. Fuzzy mathematical programming provides us a tool to deal with the natural uncertainty inherent to some production processes.

We can find several fuzzy approaches to the assessment of efficiency in the DEA literature. Sengupta [38] considers fuzzy both objective and constraints and analyzes the resulting fuzzy DEA model by using Zimmermann's method [50]. Triantis and Girod [43] use Carlsson and Korhonen method [7] in an application developed in the context of a preprint and packaging line which inserts commercial pamphlets into newspapers. Kao and Liu [29] develop a method to find the membership functions of the fuzzy efficiency scores when some observations are fuzzy numbers. The idea is based on the α-cuts and Zadeh's extension principle [47]. Hougaard's approach [27] allows the decision makers to use scores of technical efficiency in combination with other sources of information as expert opinions for instance. Entani et al. propose in [20] a DEA model with an interval efficiency consisting of efficiencies obtained from the pessimistic and the optimistic viewpoints. Their model, which is able to deal with fuzzy data, also consider inefficiency intervals. Here, we are particularly interested in the approach by Guo and Tanaka [26], which uses the possibilistic programming. In a similar manner, in the present paper we also utilize possibilistic programming techniques to approach the problem of the measurement of efficiency, but exploiting the use of the primal envelopment formulation of the DEA models instead of the dual multiplier one. Some gains are obtained with respect to both computational and interpretative aspects.

Consider that we are interested in evaluating the relative efficiency of n DMUs which use m inputs to produce s outputs. Suppose that the data of inputs and outputs cannot be precisely measured and, also, that they can be expressed as LR-fuzzy numbers with bounded support $\tilde{x}_{ij} = (x_{ij}^L - \alpha_{ij}^L, x_{ij}^L, x_{ij}^R, x_{ij}^R + \alpha_{ij}^R)_{L_{ij},R_{ij}}$, $i = 1,\ldots,m$, $j = 1,\ldots,n$, $\tilde{y}_{rj} = (y_{rj}^L - \beta_{rj}^L, y_{rj}^L, y_{rj}^R, y_{rj}^R + \beta_{rj}^R)_{L'_{rj},R'_{rj}}$, $r = 1,\ldots,s$, $j = 1,\ldots,n$ satisfying

$$\begin{array}{l} L_{i1} = \ldots = L_{in} := L_i, \quad i = 1,\ldots,m, \\ L'_{r1} = \ldots = L'_{rn} := L'_r, \quad r = 1,\ldots,s, \\ R_{i1} = \ldots = R_{in} := R_i, \quad i = 1,\ldots,m, \\ R'_{r1} = \ldots = R'_{rn} := R'_r, \quad r = 1,\ldots,s. \end{array} \tag{7}$$

Note that (7) is not too restrictive, as we are simply requiring that, for any variable (both inputs and outputs), the corresponding n data can be described by means of LR-fuzzy numbers of the same type. For instance, if these are trapezoid or triangular fuzzy numbers then (7) holds.

In both cases of offensive and defensive production, the analysis of efficiency is obtained by means n DEA models[4] with input orientation. For the DMU_0 the program is the following:

$$\begin{array}{ll} (E_0) & \text{Min } \theta \\ & \text{s.t. :} \\ & \quad \displaystyle\sum_{j=1}^{n} \lambda_j \tilde{x}_{ij} \lesssim \theta \tilde{x}_{i0}, \ i = 1,\ldots,m \\ & \quad \displaystyle\sum_{j=1}^{n} \lambda_j \tilde{y}_{rj} \gtrsim \tilde{y}_{r0}, \ r = 1,\ldots,s \\ & \quad \lambda_j \geq 0, \ j = 1,\ldots,n \end{array} \tag{8}$$

Since inputs and outputs are LR numbers, the constraints in (8) can be regarded as inequalities between LR numbers.

If we assume that the flexibility in the constraints are described by means the fuzzy number $\tilde{t}$, model (8) can be expressed as follows (see subsection 3.1.3):

$$\begin{array}{ll} (E_0^{\text{aux}}) & \text{Min } \theta \\ & \text{s.t. :} \\ & \quad \displaystyle\sum_{j=1}^{n} \lambda_j \tilde{x}_{ij} \leq_g \theta \tilde{x}_{i0} - \tilde{t}(1-\beta), \ i = 1,\ldots,m \\ & \quad \displaystyle\sum_{j=1}^{n} \lambda_j \tilde{y}_{rj} \geq_g \tilde{y}_{r0} + \tilde{t}(1-\beta), \ r = 1,\ldots,s \\ & \quad \lambda_j \geq 0, \ j = 1,\ldots,n \end{array} \tag{9}$$

[4] The proposed DEA models assume constant returns to scale. The reason is that we will analyse the league total. If the unit of analysis were each match, it would be also reasonable to use variable or non-increasing returns to scale, given that the bigger incidence of inputs may increase (in an increasing or decreasing form) the attainment of more outputs in a given match. In the league total (at home and away) this effect will be diluted by the total of games played. This is confirmed by the fact that our results are not appreciably affected when assuming returns that are non-constant.

where $\tilde{x}_{ij}$ and $\tilde{y}_{rj}$ represent inputs and output, respectively, for the team j ($j = 1,2,\ldots,n$), $\tilde{x}_{i0}$ is the input level of the team being analysed, $\tilde{y}_{r0}$ the output of the same team, and λ_j the weights of the optimal combination. Finally, θ is the relative efficiency of team j using the weights associated with the team being evaluated, and g is a ordering function.

To describe the uncertainty of inputs and outputs, we use triangular fuzzy numbers,

$$\widetilde{x}_{ij} = (x_{ij} - \alpha_{ij}^L, x_{ij}, x_{ij} + \alpha_{ij}^R) \quad \widetilde{y}_{rj} = (y_{rj} - \beta_{rj}^L, y_{rj}, y_{rj} + \beta_{rj}^R),$$

then, model (9) can be expressed, for each α-cut, as the following linear programming problem:

$$\begin{aligned}
(E_0^\alpha) \quad & \text{Min} \quad \theta \\
& \text{s.t. :} \\
& \sum_{j=1}^n \lambda_j x_{ij} - (1-\alpha)\sum_{j=1}^n \lambda_j \alpha_{ij}^L \le \theta x_{i0} - (1-\alpha)\theta\alpha_{i0}^L, \quad i = 1,\ldots,m \\
& \sum_{j=1}^n \lambda_j x_{ij} - (1-\alpha)\sum_{j=1}^n \lambda_j \alpha_{ij}^R \le \theta x_{i0} - (1-\alpha)\theta\alpha_{i0}^R, \quad i = 1,\ldots,m \\
& \sum_{j=1}^n \lambda_j x_{ij} + (1-\alpha)\sum_{j=1}^n \lambda_j \alpha_{ij}^L \le \theta x_{i0} + (1-\alpha)\theta\alpha_{i0}^L, \quad i = 1,\ldots,m \\
& \sum_{j=1}^n \lambda_j x_{ij} + (1-\alpha)\sum_{j=1}^n \lambda_j \alpha_{ij}^R \le \theta x_{i0} + (1-\alpha)\theta\alpha_{i0}^R, \quad i = 1,\ldots,m \\
& \sum_{j=1}^n \lambda_j y_{rj} - (1-\alpha)\sum_{j=1}^n \lambda_j \beta_{rj}^L \ge y_{r0} - (1-\alpha)\beta_{r0}^L, \quad r = 1,\ldots,s \\
& \sum_{j=1}^n \lambda_j y_{rj} - (1-\alpha)\sum_{j=1}^n \lambda_j \beta_{rj}^R \ge y_{r0} - (1-\alpha)\beta_{r0}^R, \quad r = 1,\ldots,s \\
& \sum_{j=1}^n \lambda_j y_{rj} + (1-\alpha)\sum_{j=1}^n \lambda_j \beta_{rj}^L \ge y_{r0} + (1-\alpha)\beta_{r0}^L, \quad r = 1,\ldots,s \\
& \sum_{j=1}^n \lambda_j y_{rj} + (1-\alpha)\sum_{j=1}^n \lambda_j \beta_{rj}^R \ge y_{r0} + (1-\alpha)\beta_{r0}^R, \quad r = 1,\ldots,s \\
& \lambda_j \ge 0, \quad j = 1,\ldots,n
\end{aligned} \tag{10}$$

Thus, the efficiency score of a given DMU_0 is a fuzzy set whose membership function is defined as follows [30]:

$$\mu_0(\theta) = \sup\{\alpha : \ \theta \text{ is an optimal value of } (E_0^\alpha)\}. \tag{11}$$

As the exact computation of $\mu_0(\theta)$ requires to obtain a supremum, if we show the decision maker a table displaying for values of α from 0 to 1 by 0.1 for each DMU, we can only compute approximately. In case that more precision is required only a few extra computational effort is necessary.

4.2.1 Other Approaches

The membership functions of the efficiency fuzzy set can be obtained by means a family of intervals. The main idea is, for each α-cut, firstly to place each DMU in the most favourable situation, $E_k^L(\alpha)$, secondly in the most unfavourable scenario, $E_k^L(\alpha)$, and then to obtain an interval $E_k(\alpha) = [E_k^L(\alpha), E_k^R(\alpha)]$ as efficiency score. For instance, Kao and Liu [29] start from the α-cuts for the inputs and the outputs,

$$x_{ij}(\alpha) = [x_{ij}^L(\alpha), x_{ij}^R(\alpha)], \quad y_{rj}(\beta) = [y_{rj}^L(\beta), y_{rj}^R(\beta)], \quad \forall i, j, r, \quad \alpha, \beta \in [0,1],$$

and they calculate the bounds of the interval $E_k(\alpha)$ making use of two mathematical programming models:[5]

$$\begin{aligned} E_{j_0}^L(\alpha) = \operatorname{Min} & \theta \\ \text{s.t.} & \sum_{j=1, j\neq j_0}^{n} \lambda_j x_{ij}^L(\alpha) + \lambda_{j_0} x_{ij_0}^R(\alpha) \leq \theta x_{io}, \forall i \\ & \sum_{j=1, j\neq j_0}^{n} \lambda_j y_{rj}^R(\alpha) + \lambda_{j_0} y_{rj_0}^L(\alpha) \geq y_{ro}, \forall r \\ & \lambda_j \geq 0, \forall j \end{aligned} \tag{12}$$

$$\begin{aligned} E_{j_0}^R(\alpha) = \operatorname{Min} & \theta \\ \text{s.t.} & \sum_{j=1, j\neq j_0}^{n} \lambda_j x_{ij}^R(\alpha) + \lambda_{j_0} x_{ij_0}^L(\alpha) \leq \theta x_{io}, \forall i \\ & \sum_{j=1, j\neq j_0}^{n} \lambda_j y_{rj}^L(\alpha) + \lambda_{j_0} y_{rj_0}^R(\alpha) \geq y_{ro}, \forall r \\ & \lambda_j \geq 0, \forall j \end{aligned} \tag{13}$$

These models are used to ascertain the efficiency for different values of α, for instance, if we take $\alpha_k = k/10, 0 \leq k \leq 10$, we obtain a family of interval valued scores of efficiency, i. e.

$$\{[E_j(k/10)^L, E_j(k/10)^R]\}_{k=0}^{10}.$$

4.2.2 Choosing a Right Model for the Teams

Clearly, the choice of the model depends on the necessities of the clubs. If the managers of the teams just want to know the efficiency and obtain a ranking based on it, we could use models that calculate interval-valued scores of efficiency (models (12), 13), and models that give us a crisp score for each α-cut (model (10)). However, in practice they prefer model (10) because models based on intervals provides intervals whose amplitude is excessive to the necessity of the clubs.

[5] In reality, Kao and Liu [29] employ the dual programmes such as those presented in this research.

The problem is more difficult if we want to know the set of teams that should be compared with a given team. For each $\alpha \in [0,1]$, the solution of models (12), (13) provides intervals to the weights λ_j,

$$[\lambda_j^L(\alpha), \lambda_j^R(\alpha)], \quad j = 1,2,\cdots,n.$$

With this approach, the number of reference teams (the one we are comparing with), strongly depends on the analysed scenario. This could create an economic difficulty to the managers of the clubs. For these reasons, in our calculations we have used model (10).

4.3 Ranking Fuzzy Efficiencies

Owing to the fuzzy data incorporated into DEA problem, the final efficiency ratio of DMU is no longer a crisp number; it is a fuzzy number. Since a fuzzy number represents many possible real numbers that have different membership values, it is not easy to determine which DMU is preferred by comparing the efficiency ratio. To solve the problem, a great deal of fuzzy ranking methods have been proposed, based on the classification by Chen and Hwang [13], these ranking method include using degree of optimality, Hamming distance, α-cut, comparison function, fuzzy mean and spread, proportion to the ideal, left and right scores, centroid index, area measurement, linguistic method etc.

Usually, the problem that we deal with DEA approach includes many evaluated DMUs, it is necessary to employ an efficient fuzzy ranking method which can handle a large quantity of fuzzy numbers. Here we cite Chen and Klein's method, based on area measurement, which use the area of the rectangle obtained by multiplying the height of the membership function by the distance between two crisp maximizing and minimizing barriers as a referential rectangle (or compare background) for each fuzzy number, and then calculates the ranking index by their difference.

The referential rectangle is used as a normalization condition and is based on the concept of more is better. That is, generally whenever fuzzy number is compared the one with more area to the right is considered better. The referential rectangle is another approach to help measure this. The Chen and Klein's ranking index is based on the difference between the compared fuzzy number and the referential rectangle.

Fortunately, in the fuzzy DEA problem, the efficiency ratio is always in the interval $[0,1]$ and the height of the membership function is 1, which forms a closed and normalized rectangle with the horizontal-axe interval being $[0,1]$. It is denoted as fuzzy number R. Hence, it is easy to transform the original membership function into a new one by subtraction from R.

In Figure 1, for the k-th DMU to say, $u_{Ck-R}(x)$ is the membership function of $C_k - R$, A^+ is the positive area of $C_k - R$, A^- is the negative area of $C_k - R$. We favor A^+ with a larger value and A^- with a smaller value. Thus an index $I_k = \frac{A^+}{A^+ + A^-}$ is defined to determine the ranking order of the k-th DMU. The subtraction between fuzzy numbers C_k and R shift to left, since the subtracted interval at α_1-cut is

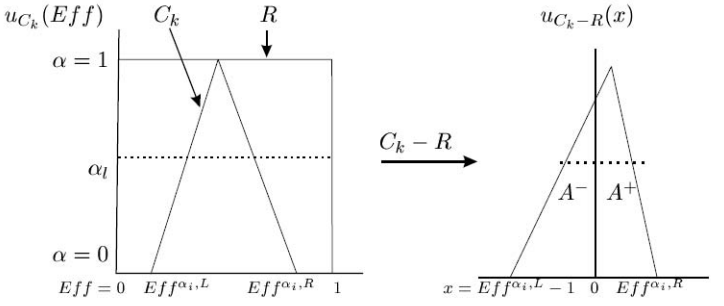


Fig. 1 Illustration of Chen and Klein's fuzzy ranking index, [25]

$[Eff^{\alpha_i,L}, Eff^{\alpha_i,R}] - [0,1] = [Eff^{\alpha_i,L} - 1, Eff^{\alpha_i,R}]$. Therefore, the ranking index is represented for k-th DMU as:

$$I_k = \frac{\sum_{i=1}^{n} Eff^{\alpha_i,R}}{\sum_{i=1}^{n} Eff^{\alpha_i,R} - \sum_{i=1}^{n} (Eff^{\alpha_i,L} - 1)}, \quad n \to \infty$$

This is a much simpler form for computing I_k, where n is the number of α-cuts and as n approaches ∞. A^+ is $\sum_{i=1}^{n} Eff^{\alpha_i,R}$ and A^- is $\sum_{i=0}^{n} (Eff^{alpha_i,L} - 1)$. The k-th DMU with a higher index I_k is considered more efficient than the DMU with a lower index. The value scope of I_k is between 0 and 1, which is consistent with the scope of efficiency ratio in crisp DEA model.

According to the discussion by Chen and Klein [12], the proposed ranking method for m fuzzy numbers uses only m comparisons to the same referential rectangle as opposed to the $\frac{m(m-1)}{2}$ comparisons needed by existing ranking methods. Hence it is an efficient, accurate and effective fuzzy ranking method, especially when a large quantity of fuzzy numbers are evaluated. In addition, Chen and Klein's method requires only three of four α-cuts and uses the summation of each α-level interval which does not require normality to measure the summation for the ranking order of the fuzzy numbers. This increases efficiency as well and allows for applicability to a wider range of fuzzy numbers.

4.3.1 A Suitable Ranking

If we choose the Kao and Liu approach, or any other based on interval-valued scores of efficiency for each α-cut, for values of α evenly spaced, $\alpha_\ell = \ell/N, \quad \ell = 0, \cdots, N$, it is useful to apply the Chen and Klein index for obtaining a rank of DMUs.

Definition 8. *The j^{th}-DMU is more efficient than the k^{th}-DMU if and only if $I_j > I_k$, i. e.*

$$\tilde{E}_j > \tilde{E}_k \qquad \text{iff} \qquad I_j > I_k. \tag{14}$$

However, model (10), for some values of α, provides an interval for each team. Then, it is easier to introduce a ranking for the intervals.

Definition 9. *Given the intervals* $A=[a_1,a_2], B=[b_1,b_2]\subset\mathbb{R}$*, we will say that* A *is bigger than* B *if and only if*

$$A>B \Leftrightarrow \begin{cases} k_1a_1+k_2a_2>k_1b_1+k_2b_2, & k_1a_1+k_2a_2\neq k_1b_1+k_2b_2 \\ a_1>b_1, & k_1a_1+k_2a_2=k_1b_1+k_2b_2 \end{cases}$$

where k_1,k_2 *are two pre-established positive constants.*

We use a rank method proposed, *ad hoc*, by Govan *et al.* [24] to crisp problems. They obtain a ratio of efficiency as the product[6] of the offensive efficiencies, $\{e_{\text{of}_j}\}_{j=1}^n$, and defensive ones, $\{e_{\text{def}_j}\}_{j=1}^n$:

$$r_j=e_{\text{of}_j}\cdot e_{\text{def}_j},\quad j=1,2,\cdots,n. \tag{15}$$

For extending this result to interval-valued scores of efficiency, $\{[E^L_{\text{of}_j},E^R_{\text{of}_j}]\}_{j=1}^n$, $\{[E^L_{\text{def}_j},E^R_{\text{def}_j}]\}_{j=1}^n$, firstly we need multiply both kinds of efficiencies:

$$[E^L_j,E^R_j]=\left[E^L_{\text{of}_j},E^R_{\text{of}_j}\right]\times\left[E^L_{\text{def}_j},E^R_{\text{def}_j}\right],\quad j=1,2,\cdots,n.$$

According to Definition 9, we rank the intervals $\{[E^L_j,E^R_j]\}_{j=1}^n$ and, by (15), we can rank the teams of the League:

Definition 10. *The* j^{th}*-DMU is more efficient than the* k^{th}*-DMU if and only if* $[E^L_j,E^R_j]>[E^L_k,E^R_k]$.

4.4 Experiments

In every experiments we work with data of the Spanish Football League 2006/07. We present results of offensive and defensive efficiencies in general, without distinguish between home and away because this provides a global idea and permits simplify the results. Besides, in all the tables, the teams are ranged according to their classification in the League.

4.4.1 The Data of the Model

We begin by defining football offensive and defensive outputs, and then justify our choice of inputs for the resolution of the DEA models. Our calculations of the frontiers of offensive production assume that teams in Spanish League are trying to maximise the most straightforward measure of football production, i.e. the number of goals scored over the course of the season (O^o). The reason to choose the number

[6] Govan *et al.* define the ratio of efficiency as a quotient between efficiencies, but our definition is equivalent because in [24] the model for determine the defensive efficiency have output orientation and, besides, the defensive inputs and outputs are the inverse of ours, $1/I_i$ y $1/O$.

of goals instead of the total number of points in the classification, for instance, is that in this way we take into account implicitly the show aspect of football. The corresponding output measure to compute the frontiers of defensive production will be the inverse[7] of the number of goals conceded ($O^d = 1/O^o$) by a team.

For the sake of simplicity and comparability we want always to use the same inputs to compute the different defensive and offensive frontiers of production. We have chosen the following four offensive inputs:

- Balls kicked into the opposing teams centres area (I_1^o),
- attacking plays made by the team (I_2^o),
- minutes of possession (I_3^o),
- shots-on-goal (I_4^o).

And we will consider four defensive inputs defined as the inverse of the above mentioned, i. e. $I_i^d = 1/I_i^o$, $i = 1,2,3,4$.

With respect to the attacking inputs, the four chosen measures approach the effort, the time or the abilities a team has employed in offensive tasks. Hence, conceivably these four proxies capture reasonably well the amount of offensive inputs employed by each team. With respect to the defending inputs, the idea is similar, although we need to make some qualifications. In this case, we are taking as proxies of the defending inputs, the inverse of game events that are performed by the opposing team, rather than actions of the team whose defending abilities we wish to capture. The data of the League 2005/07 are shown in table 1.

Table 1 Values of crisp inputs and outputs

TEAM	O1	I1	I2	I3	I4	TEAM	O1	I1	I2	I3	I4
Real Madrid	84	555	165	813	1008	Athletic C. Bilbao	40	448	96	1038	765
Villarreal	63	450	133	799	940	Espanyol	43	428	125	1119	880
F.C. Barcelona	76	566	333	981	1191	Real Betis	45	444	99	1033	843
Atl. de Madrid	66	516	154	814	905	Getafe	44	522	137	1142	919
Sevilla F.C.	75	568	151	1197	912	Real Valladolid	42	478	108	1288	914
Racing Santander	42	469	110	996	816	Recreativo Huelva	40	456	139	1084	793
R.C.D. Mallorca	69	477	161	1063	883	Osasuna	37	479	113	1081	859
Almería	42	450	134	1072	876	Real Zaragoza	50	519	159	1040	941
Dep. A Coruña	46	476	124	1127	911	Murcia	36	417	102	1048	803
Valencia C.F	48	394	141	998	956	Levante	33	413	120	1023	819

On attempting to include the uncertainty of the data, we have encountered non-uniform situations in all teams that should be illustrated by different variation percentages. A cluster analysis of the last 10 football seasons has been carried out and

[7] Using the inverse of goals conceded as the output measure means that the team with the best defensive performance displays the highest output.

clubs have been grouped under two criteria: annual income and sporting history in the Spanish League since they were founded. This did not make it possible to group the teams into five categories so that variation percentages were homogeneous within each, but little difference is appreciated among the groups themselves. Assuming this criterion, the percentages obtained for each team is expressed in Table 2.

Note that if one same variation percentage were applied to all the teams, we would not be able to compare the two indexes. Table 2 includes the tolerances (as percentages) applied to each team and the group to which it belongs.

Table 2 Percentage of tolerances for each input and output

TEAM	O1	I1	I2	I3	I4	TEAM	O1	I1	I2	I3	I4
Real Madrid	6	3	12	4	3	Athletic C. Bilbao	10	4	13	3	3
Villarreal	12	4	13	6	4	Espanyol	10	4	13	3	3
F.C. Barcelona	6	3	12	4	3	Real Betis	12	4	13	6	4
Atl. de Madrid	12	4	13	6	4	Getafe	10	5	12	5	4
Sevilla F.C.	8	5	13	5	4	Real Valladolid	10	5	12	5	4
Racing Santander	10	5	12	5	4	Recreativo Huelva	10	5	12	5	4
R.C.D. Mallorca	10	5	12	5	4	Osasuna	10	4	13	3	3
Almería	10	5	12	5	4	Real Zaragoza	12	4	13	6	4
Dep. A Coruña	8	5	13	5	4	Murcia	10	5	12	5	4
Valencia C.F.	8	5	13	5	4	Levante	10	5	12	5	4

To sum up, our model have the following inputs and outputs:

$$\begin{aligned} \tilde{x}_{ij}^{o} &= \left(I_{ij}^{o} - \alpha_{ij}^{oL}, I_{ij}^{o}, I_{ij}^{o} + \alpha_{ij}^{oR}\right), \ \tilde{y}_{j}^{o} = \left(O^{o} - \beta_{j}^{oL}, O_{j}^{o}, O_{j}^{o} + \beta_{j}^{oR}\right), \\ \tilde{x}_{ij}^{d} &= \left(I_{ij}^{d} - \alpha_{ij}^{dL}, I_{ij}^{d}, I_{ij}^{d} + \alpha_{ij}^{dR}\right), \ \tilde{y}_{j}^{d} = \left(O^{d} - \beta_{j}^{dL}, O_{j}^{d}, O_{j}^{d} + \beta_{j}^{dR}\right), \end{aligned} \tag{16}$$

where

$$\begin{aligned} &\alpha_{ij}^{oL} = \alpha_{ij}^{oR} = I_{ij}^{o} p_{ij}, && \beta_{j}^{oL} = \beta_{j}^{oR} = O_{j}^{o} p_{j} \\ &\alpha_{ij}^{dL} = \frac{(I_{ij}^{d})^2 p_{ij}}{1+I_{ij}^{d} p_{ij}}, \quad \alpha_{ij}^{dR} = \frac{(I_{ij}^{d})^2 p_{ij}}{1-I_{ij}^{d} p_{ij}}, && \beta_{j}^{dL} = \frac{(O_{j}^{o})^2 p_{ij}}{1+O_{j}^{d} p_{ij}}, \quad \beta_{j}^{dR} = \frac{(O_{j}^{o})^2 p_{ij}}{1-O_{j}^{d} p_{ij}} \end{aligned} \tag{17}$$

and $i = 1, \ldots, 4, \quad j = 1, \ldots, 20.$

4.4.2 Analizing the Efficiency

In this subsection we present the result of general offensive efficiency (Table 3) and defensive efficiency (Table 4) obtained with model (10) for some values of α. In these tables, the last column show the interval where we can find the scores of efficiency of each team.

Table 3 Scores of offensive efficiency

Teams	Score of efficiency for some values of α							Interval
	0	0.1	0.3	0.5	0.7	0.9	1	
Real Madrid	1	1	1	1	1	1	1	[1. 1]
Villarreal	0.9576	0.9562	0.9529	0.9487	0.9432	0.9355	0.9305	[0.9305. 0.9576]
F.C. Barcelona	0.8872	0.8872	0.8872	0.8872	0.8872	0.8872	0.8872	[0.8872. 0.8872]
Atl. de Madrid	0.9006	0.8993	0.8962	0.8923	0.8871	0.8799	0.8751	[0.8751. 0.9006]
Sevilla F.C.	0.9964	0.9959	0.9948	0.9933	0.9913	0.9886	0.9868	[0.9868. 0.9964]
Rac. Santander	0.7646	0.7638	0.7621	0.7598	0.7568	0.7527	0.75	[0.7500. 0.7646]
R.C.D. Mallorca	0.9743	0.9734	0.9711	0.9682	0.9645	0.9592	0.9558	[0.9558. 0.9743]
Almería	0.6286	0.628	0.6266	0.6247	0.6223	0.6189	0.6167	[0.6167. 0.6286]
Dep. A Coruña	0.7358	0.7354	0.7345	0.7334	0.732	0.73	0.7287	[0.7287. 0.7358]
Valencia C.F.	0.8127	0.8123	0.8114	0.8102	0.8086	0.8064	0.8049	[0.8049. 0.8127]
Ath. C. Bilbao	0.8343	0.8335	0.8316	0.8292	0.8259	0.8214	0.8185	[0.8185. 0.8343]
Espanyol	0.6888	0.6882	0.6866	0.6845	0.6819	0.6782	0.6757	[0.6757. 0.6888]
Real Betis	0.9189	0.9175	0.9144	0.9104	0.9051	0.8977	0.8929	[0.8929. 0.9189]
Getafe	0.6431	0.6425	0.641	0.6391	0.6366	0.6331	0.6309	[0.6309. 0.6431]
Real Valladolid	0.7787	0.778	0.7762	0.7739	0.7708	0.7667	0.7639	[0.7639. 0.7787]
Rec. Huelva	0.617	0.6164	0.615	0.6132	0.6108	0.6075	0.6053	[0.6053. 0.6170]
Osasuna	0.6557	0.655	0.6535	0.6516	0.649	0.6455	0.6432	[0.6432. 0.6557]
Real Zaragoza	0.6562	0.6552	0.653	0.6501	0.6463	0.6411	0.6376	[0.6562. 0.6376]
Murcia	0.7067	0.7061	0.7044	0.7023	0.6996	0.6958	0.6933	[0.6933. 0.7067]
Levante	0.5507	0.5501	0.5489	0.5472	0.5451	0.5421	0.5402	[0.5402. 0.5507]

Table 4 Scores of defensive efficiency

Teams	Score of efficiency for some values of α							Interval
	0	0.1	0.3	0.5	0.7	0.9	1	
Real Madrid	1	1	1	1	1	1	1	[1, 1]
Villarreal	0.9574	0.9561	0.9528	0.9486	0.9431	0.9354	0.9303	[0.9303, 0.9574]
F.C. Barcelona	0.7841	0.7841	0.7841	0.7841	0.7841	0.7841	0.7841	[0.7841, 0.7841]
Atl. de Madrid	0.7883	0.7871	0.7844	0.781	0.7764	0.7701	0.766	[0.7660 0.7883]
Sevilla F.C.	0.7177	0.7173	0.7165	0.7154	0.714	0.7121	0.7108	[0.7108, 0.7177]
Rac. Santander	0.995	0.994	0.9918	0.9888	0.985	0.9796	0.9761	[0.9761. 0.9950]
R.C.D. Mallorca	0.7658	0.765	0.7632	0.761	0.758	0.7539	0.7512	[0.7512, 0.7658]
Almería	0.8089	0.8081	0.8062	0.8039	0.8007	0.7964	0.7935	[0.7935, 0.8089]
Dep. A Coruña	0.8534	0.853	0.852	0.8507	0.8491	0.8468	0.8452	[0.8452, 0.8534]
Valencia C.F.	0.604	0.6037	0.603	0.6021	0.6009	0.5993	0.5982	[0.5982, 0.6040]
Ath. C. Bilbao	1	0.8675	0.8655	0.8629	0.8596	0.8549	0.8518	[0.8518, 1]
Espanyol	0.8875	0.8866	0.8846	0.882	0.8785	0.8737	0.8706	[0.8706, 0.8875]
Real Betis	0.8498	0.8485	0.8456	0.8419	0.837	0.8302	0.8257	[0.8257, 0.8498]
Getafe	0.7948	0.794	0.7922	0.7898	0.7867	0.7825	0.7797	[0.7797, 0.7948]
Real Valladolid	0.6393	0.6387	0.6372	0.6353	0.6328	0.6294	0.6271	[0.6271, 0.6393]
Rec. Huelva	0.6586	0.6579	0.6564	0.6545	0.6519	0.6484	0.646	[0.6460, 0.6586]
Osasuna	0.9969	0.9154	0.9133	0.9106	0.907	0.9021	0.8988	[0.8988, 0.9969]
Real Zaragoza	0.6886	0.6876	0.6853	0.6822	0.6783	0.6728	0.6691	[0.6691, 0.6886]
Murcia	0.6046	0.6041	0.6027	0.6009	0.5985	0.5953	0.5931	[0.5931, 0.6046]
Levante	0.572	0.5715	0.5701	0.5685	0.5662	0.5632	0.5611	[0.5611, 0.5720]

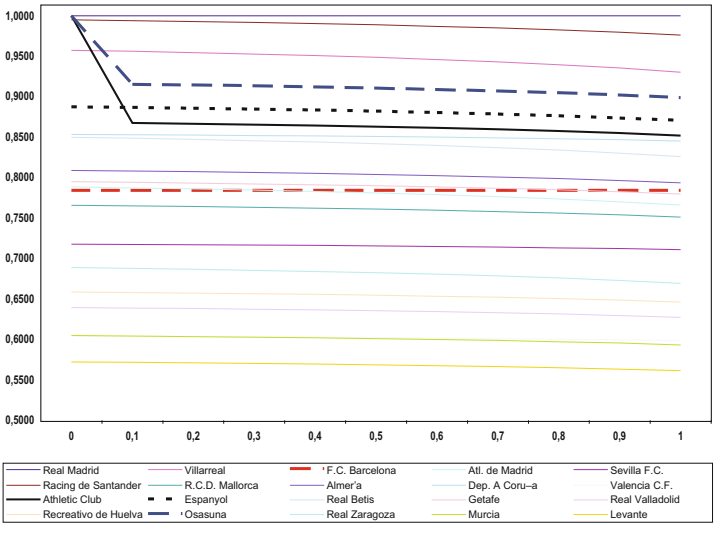


Fig. 2 Defensive efficiency for some values of α in model (10)

Table 5 Ranking of the teams

TEAM	Ratio	Position	TEAM	Ratio	Position
Real Madrid	1	1	Athletic C. Bilbao	0.752038	4
Villarreal	0.886108	2	Espanyol	0.597482	12
F.C. Barcelona	0.695653	8	Real Betis	0.754713	3
Atl. de Madrid	0.686173	9	Getafe	0.499602	13
Sevilla F.C.	0.706896	7	Real Valladolid	0.486554	15
Racing Santander	0.743556	5	Recreativo Huelva	0.397156	19
R.C.D. Mallorca	0.729245	6	Osasuna	0.609144	11
Almería	0.497001	14	Real Zaragoza	0.439058	17
Dep. A Coruña	0.620711	10	Murcia	0.4176261	18
Valencia C.F	0.485243	16	Levante	0.307863	20

Figure 4 shows how the defensive score of efficiency of some teams (for instance, Espanyol and Athletic Club Bilbao) strongly depends on the values of α.

4.4.3 Ranking the Teams According to their Efficiency

From the obtained intervals in sections above, by using Definition 10 we can rank the teams (see Table 5).

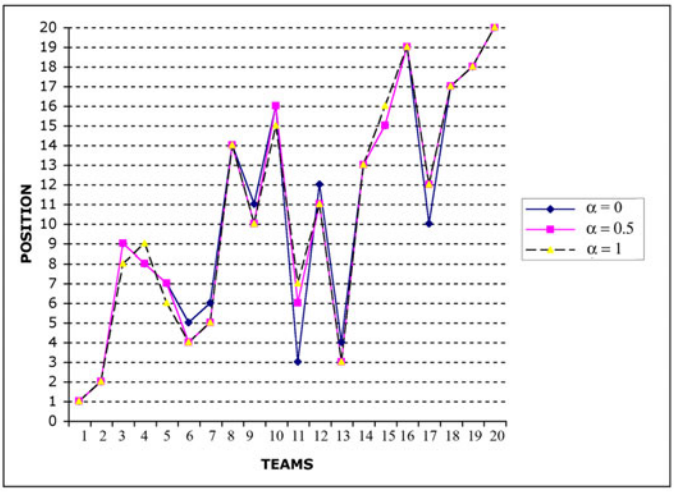


Fig. 3 Illustration of the ranking for some values of α

This information can be completed analyzing the *stability* of position of each team in the ranking. For this, it is useful to provide to the managers different scenarios for some values of $\alpha \in [0, 1]$. For instance, Figure 3 illustrates as slight modifications of the value of α produce modifications in the defensive efficiency score.

4.4.4 Which Teams Should Be Compared to a Club

The managers find valuable to know the teams they should compare with to increase its efficiency, specially if the league is not finished and they still have possibilities to change the use of their inputs.

As an example, we will only comment that the Osasuna team for the defensive efficiency compares to *Real Madrid* and *Athletic Club de Bilbao* with percentages 65% and 21.6%, respectively ($\lambda_1 = 0.65, \lambda_{11} = 0.216$).

In order to give better advices to sports managers, sometimes it is convenient to include in model (10) a convexity constraint

$$\sum_{j=1}^{n} \lambda_j = 1. \tag{18}$$

In this way, the efficient frontier is closer to data set. For instance, if we consider defensive efficiency and take $\alpha = 0$, the comparison among the different teams can be found in Table 6.

Table 6 Defensive efficiency. Optimal values of the weights λ_j (for $\alpha = 0$) adding in the model (E_0) the convexity constraint

Teams	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	.42	0	0	0	0	.51	0	0	0	0	0	0	.08	0	0	0	0	0	0	0
3	.49	0	0	0	0	0	0	0	0	0	0	.51	0	0	0	0	0	0	0	0
4	.46	0	0	0	0	0	0	0	0	0	0	0	.27	0	0	0	0	0	0	.27
5	0	0	0	0	0	.17	0	0	0	0	0	.83	0	0	0	0	0	0	0	0
6	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0	0	0	0	0	.68	.23	0	0	0	0	0	0	.09
8	.35	0	0	0	0	0	0	0	0	0	0	.21	.43	0	0	0	0	0	0	0
9	0	0	0	0	0	.35	0	0	0	0	0	.65	0	0	0	0	0	0	0	0
10	.66	0	0	0	0	0	0	0	0	0	0	.34	0	0	0	0	0	0	0	0
11	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
12	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
13	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
14	.15	0	0	0	0	0	0	0	0	0	0	0	.85	0	0	0	0	0	0	0
15	.44	0	0	0	0	0	0	0	0	0	0	0	.28	0	0	0	0	0	0	.28
16	0	0	0	0	0	0	0	0	0	0	0	0	.65	0	0	0	0	0	0	.35
17	.45	0	0	0	0	0	0	0	0	0	0	.48	.07	0	0	0	0	0	0	0
18	0	0	0	0	0	0	0	0	0	0	0	.8	.2	0	0	0	0	0	0	0
19	0	0	0	0	0	0	0	0	0	0	0	0	.5	0	0	0	0	0	0	.5
20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1

[a] The teams $1, 2, \cdots, 20$ are ranged according to their classification in the League.

5 Conclusions

The clubs which make up the professional Spanish Football League officially presented income and expenses of over 1,200 million euros for the season 2004/05. The direct and indirect impact of the league on the Spanish economy may account for over 0.25% of the GDP. Similar figures can be deduced for the all the major European Leagues, e.g. the Italian and English Leagues. In these countries football has become the main leisure activity, apart from its huge social and economic importance, and yet a large part of the mathematical techniques used in the economics and finances of other companies have not been applied to football.

In this paper we show some methods to combine two necessities: the handling of data which present inherent uncertainty and the need of the clubs (our firms) to know whether they are managing their inputs efficiently. There is no doubt that modelling the League through fuzzy sets has caused some scenarios to appear that would have passed unnoticed in a crisp context, the FLP has provided very useful tools for decision making in the context of the League. For example, a crisp analysis of efficiency usually leads to inefficient clubs being compared with very few efficient ones. Fuzzy treatment, however, means that for some values of the α-cuts there are more efficient teams, thus increasing the number of teams to be compared. We have shown objective ways or ordering the teams in the League on the basis of their technical/sporting efficiency. This classification, which does not have to coincide

with necessarily with that of the league, allows managers of clubs to know the aspirations of their teams and so adjust their strategies to achieve a better classification. Moreover, although the calculation of efficiency with DEA models is not conceived for application as prediction, but to perform analyses at the end of a period, working with fuzzy DEA models gives teams predictions as to their efficiency score. This result is without doubt of high value for the clubs because if they can be informed before the end of the league, they can maintain their capacity to modify how they use the inputs.

Acknowledgements. We would like to express our thanks to GECA Sport who have generously supplied the data used in this study. This work has been partially supported by the Ministerio de Ciencia e Innovación of Spain, TIN2008-06872-C04-02, TIN2008-06872-C04-03, TIN2008-06872-C04-04 and TIC-02970-JA.

References

1. Banker, R.D., Charnes, A., Cooper, W.W.: Some models for estimating technical and scale inefficiencies in data envelopment analysis. Management Sci. 30(9), 1078–1092 (1984)
2. Bellman, R.E., Zadeh, L.A.: Decision Making in a Fuzzy Environment. Management Science 17(B)4, 141–164 (1970)
3. Briec, W.: A graph-type extension off arrell technical efficiency measure. J. Prod. Anal. 8(1), 95–110 (1997)
4. Cadenas, J.M., Verdegay, J.L.: Using Fuzzy Numbers in Linear Programming. IEEE Transactions on Systems, Man, and Cybernetics 27(B)6, 1017–1022 (1997)
5. Cadenas, J.M., Verdegay, J.L.: Modelos de Optimización con Datos Imprecisos, Univesidad de Murcia, Servicio de Publicaciones (1999)
6. Campos, L., Verdegay, J.L.: Linear Programming Problems and Ranking of Fyzzy Numbers. Fuzzy Sets and Systems 32, 1–11 (1989)
7. Carlsson, C., Korhonen, P.: A parametric approach to fuzzy linear programming. Fuzzy Sets and Systems 20, 17–30 (1986)
8. Carmichael, F., Thomas, D.: Production and efficiency in team sports: an investigation of rugby league football. Applied Economics 27, 859–869 (1995)
9. Carmichael, F., Thomas, D., Ward, R.: Team performance: the case of English premiership football. Managerial and Decision Economics 21, 31–45 (2000)
10. Chambers, R.G., Chung, Y., FWare, R.: BeneHt and distance functions. J. Econom. Theory 70(2), 407–419 (1996)
11. Charnes, A., Cooper, W.W., Rhodes, E.: Measuring the efficiency of decision making units. European J. Oper. Res. 2(6), 429–444 (1978)
12. Chen, C.B., Klein, C.M.: A simple approach to ranking a group of aggregated fuzzy utilities. IEEE Transaction on System, Man and Cybernetics 27(1), 26–35 (1997)
13. Chen, S., Hwang, C.L.: Fuzzy multiple attribute decision making. Spring-Veriag, New York (1992)
14. Cooper, W.W., Seiford, L.M., Tone, K.: Data envelopment analysis. Kluwer, Boston (1999)
15. Dawson, P., Dobson, S., Gerrar, B.: Stochastic frontiers and the temporal structure of managerial efficiency in English soccer. Journal of Sports Economics 1(4), 341–362 (2000)

16. Delgado, M., Verdegay, J.L., Vila, M.A.: Imprecise costs in mathematical programming problems. Control and Cybernet. 16(2), 113–121 (1987)
17. Delgado, M., Verdegay, J.L., Vila, M.A.: A general model for fuzzy linear programming. Fuzzy Sets and systems 29, 21–29 (1989)
18. Dubois, D., Prade, H.: Operations on Fuzzy Numbers. International Journal Systems Science 9, 613–626 (1978)
19. Dubois, D., Prade, H.: Fuzzy Sets and Systems. Theory and Applications. Academic Press, London (1980)
20. Entani, T., Maeda, Y., Tanaka, H.: Dual models ofinterval DEA and its extension to interval data. European J. Oper. Res 136, 32–45 (2002)
21. FWare, R., Grosskopf, S., Lovell, C.A.K.: The Measurement of Efficiency of Production. Kluwer-Nijho Publishing, Dordrecht-Hague (1985)
22. FWare, R., Lovell, C.A.K.: Measuring the technical efficiency of production. J. Econom. Theory 19, 150–162 (1978)
23. Fedrizzi, M., Kacprzyk, J., Verdegay, J.L.: A Survey of Fuzzy Optimization and Fuzzy Mathematical Programming. In: Fedrizzi, M., Kacprzyk, J., Roubens, M. (eds.) Interactive Fuzzy Optimization. Springer, Berlin (1991)
24. Govan, A.Y., Langville, A.N., Meyer, C.D.: Offense-Defense Approach to Ranking Team Sports. Journal of Quantitative Analysis in Sports 5(1) (2009), http://www.bepress.com/jqas/vol5/iss1/4
25. Guh, Y.-Y.: Data Envelopment Analysis in Fuzzy Environment. Information and Management Sciences 12(2), 51–65 (2001)
26. Guo, P., Tanaka, H.: Fuzzy DEA: a perceptual evaluation method. Fuzzy Sets and Systems 119, 149–160 (2001)
27. Hougaard, J.L.: Fuzzy scores of technical efficiency. European J. Oper. Res. 115, 529–541 (1999)
28. Ishibuchi, H., Tanaka, H.: Multipleobjective programming in optimization of the interval objective function. EJOR 48, 219–225 (1990)
29. Kao, C., Liu, S.T.: Fuzzy efficiency measures in data envelopment analysis. Fuzzy Sets and Systems 119, 149–160 (2000)
30. León, T., Liern, V., Ruiz, J., Sirvent, I.: A Possibilistic Programming Approach to the Assessment of Efficiency with DEA Models. Fuzzy Sets and Systems 139, 407–419 (2003)
31. Maudos, J., Pastor, J.M., Serrano, L.: Convergencia en las regiones españolas: cambio técnico, eficiencia y productividad. Revista Española de Economía 15(12), 235–264 (1998)
32. Requena, I.: Redes Neuronales en problemas de decisión con ambiente difuso, Tesis doctoral, Universidad de Granada (1992)
33. Rommelfanger, H., Hanuscheck, R., Wolf, J.: Linear programming with fuzzy objectives. Fuzzy Sets and Systems 29, 31–48 (1990)
34. Rottenberg, S.: The baseball players labor-market. Journal of Political Economy 64, 242–258 (1956)
35. Schofield, J.A.: Production functions in the sports industry: an empirical analysis of professional cricket. Applied Economics 20, 177–193 (1988)
36. Scott, F.A., Long, J.E., Somppi, K.: Salary vs marginal revenue product under monopsony and competition: the case of professional basketball. Atlantic Economic Journal 13, 50–59 (1985)
37. Scully, G.W.: Pay and performance in major league baseball. American Economic Review 64, 915–930 (1974)

38. Sengupta, J.K.: A fuzzy systems approach in data envelopment analysis. Comput. Math. Appl. 24, 259–266 (1992)
39. Sloane, P.J.: The economics of professional football, the football club as a utility maximiser. Scottish Journal of Political Economy 8, 121–146 (1971)
40. Sloane, P.J.: The economic of sport. Economic Affairs 17 (1997) (special issue)
41. Tanaka, H., Okuda, T., Asai, K.: On Fuzzy Mathematical Programming. Journal of Cybernetics 3(4), 37–46 (1974)
42. Tanaka, H., Ichihashi, H., Asai, F.: A formulation of fuzzy linear programming problems based a comparison of fuzzy numbers. Control and Cybernet. 13, 185–194 (1984)
43. Triantis, K., Girod, O.: A mathematical programming approach for measuring technical efficiency in a fuzzy environment. J. Prod. Anal. 10, 85–102 (1998)
44. Verdegay, J.L.: Fuzzy Mathematical Programming. In: Gupta, M.M., Sanchez, E. (eds.) Fuzzy Information and Decisión Processes (1982)
45. Verdegay, J.L., Yager, R.R., Bonissone, P.P.: On heuristics as a fundamental constituent of Soft Computing. Fuzzy Sets and Systems 159, 846–855 (2008)
46. Wang, X., Kerre, E.E.: Reasonable properties for the ordering of fuzzy quantities (I) and (II). Fuzzy Sets and Systems 118, 375–385, 387–405 (2001)
47. Zadeh, L.A.: Fuzzy Sets. Information and Control 8, 338–353 (1965)
48. Zadeh, L.A.: The Concept of a Linguistic Variable and its Applications to Approximate Reasoning. Part I. Information Sciences 8, 199–249 (1975); Part II. Information Sciences 8, 301–357 (1975); Part III. Information Sciences 9, 43–80 (1975)
49. Zak, T.A., Huang, C.J., Siegfried, J.J.: Production efficiency: the case of professional basketball. Journal of Business 52, 379–392 (1979)
50. Zimmermann, H.J.: Description and Optimization of Fuzzy Systems. International Journal of General Systems 2, 209–215 (1976)
51. Zhu, Q., Lee, E.S.: Comparison and Ranking of Fuzzy Numbers. In: Kacprzyk, J., Fedrizzi, M. (eds.) Fuzzy Regression Analysis. Onmitech Press Warsan/Physica-Verlag (1992)

Author Index

Beliakov, Gleb 77

Cadenas, J.M. 503
Cruz, Carlos 127

Dompere, Kofi Kissi 3

Fang, Shu-Cherng 109

Ge, Yue 313
Gil-Lafuente, Anna M. 389

Hasuike, Takashi 407

Inuiguchi, Masahiro 145
Ishii, Hiroaki 313, 407

Kasperski, Adam 287
Kaymak, Uzay 343
Kuchta, Dorota 323

Li, Pingke 109
Liern, V. 503
Lodwick, Weldon A. 33

Mareš, Milan 271
Merigó, José M. 389

Ogryczak, Włodzimierz 209

Richards, Zach 479

Saito, Seiji 365
Sala, R. 503
Sciacca, Eva 457
Silva, Ricardo C. 127
Sirbiladze, Gia 231
Śliwiński, Tomasz 209
Soldatenko, Ilia 163
Sousa, João M.C. 343
Spinella, Salvatore 457

Torra, Vicenç 65
Tsuda, Hiroshi 365

Untiedt, Elizabeth 33, 193

Verdegay, J.L. 503
Verdegay, José L. 127
Vieira, Susana M. 343

Wang, Shuming 425
Watada, Junzo 425

Yamakami, Akebo 127
Yazenin, Alexander 163

Zieliński, Paweł 287