Universidad de Las Palmas de Gran Canaria
Departamento de Informática y Sistemas

Programa de Doctorado:

Sistemas Inteligentes y Aplicaciones Numéricas en Ingeniería

Tesis Doctoral

On Uncertainty and Robustness in Evolutionary Optimization-based Multi-Criterion Decision-Making

(Acerca de la incertidumbre y la robustez en toma de decisiones multicriterio basada en optimización evolutiva)

Daniel E. Salazar Aponte

2008

Universidad de Las Palmas de Gran Canaria
Departamento de Informática y Sistemas

Programa de Doctorado:

Sistemas Inteligentes y Aplicaciones Numéricas en Ingeniería

Tesis Doctoral

On Uncertainty and Robustness in Evolutionary Optimization-based Multi-Criterion Decision-Making

(Acerca de la incertidumbre y la robustez en toma de decisiones multicriterio basada en optimización evolutiva)

El Director	El Codirector	El Autor
Blas José Galván González	Claudio Miguel Rocco Sanseverino	Daniel Enrique Salazar Aponte

Las Palmas de Gran Canaria, junio de 2008

Abstract

In general, decision-making problems consist in finding the subset of solutions that optimize the level of satisfaction of the Decision-Makers according to some figure of merit or criterion, amongst the whole set of feasible alternatives. When the quality of the alternatives is compared against more than one criterion, the problem is said to be of multiple-criteria decision-making.

In order to identify the subset of multiple-criteria optimal solutions or Pareto frontier, many analysts make use of metaheuristics like Multiple-Objective Evolutionary Algorithms (MOEA), due to their ability to perform well over non-continuous and non-linear domains as well as on black-box functions. This type of metaheuristics comprises specific as well as all-purpose algorithms and in general, its use turns out a discrete approximation of the Pareto frontier.

This thesis addresses the negative effects and problems risen by uncertainty on multiple-criteria decision-making based on evolutionary algorithms, making methodological and algorithmic contributions to cope with such uncertainty, regarding the formulation of the decision-making problems into a mathematical programs as well as structural and operational issues of MOEA to solve such programs.

In the first part of this thesis, the meaning of uncertainty, its sources, types and the possible instances of a system subject to uncertainty are analyzed. Additionally, the main theoretical frameworks to characterize uncertainty are surveyed. On the other hand, the structure and operation of all-purpose MOEA, from the most general structure of an evolutionary algorithm to the state of the art in MOEA design, is analyzed. From such analyses, the ranking and the archiving procedures are identified in this work as the features of MOEA to be tailored in order to cope with uncertainty. The structure of such procedures are compared against the non-probabilistic, probabilistic, possibilistic, fuzzy and imprecise probabilistic characterizations of uncertainty, whereas conclusions are drawn around the possible alternatives to address such types of uncertainty.

A MOEA named IP-MOEA is presented in this thesis as an alternative for dealing with non-probabilistic uncertain outcomes. Other contributions to the MOEA design are envisaged thorugh the analytical comparison of the ϵ-dominance with the minimal regret criterion as an alternative to cope with non-probabilistic uncertain outcomes. Likewise, the archiving of probabilistic uncertain outcomes is considered by reinterpreting the hyperbox's density of the hypergrid-based archiving procedures in terms of probability.

The notion of robustness, particularly with respect to optimization is thoroughly explored. The state of the art in robustness regarding decision-making and MOEA is surveyed. Afterwards a new taxonomy of classes of robustness is proposed in this thesis in terms of the available information. The possible

interpretations of such classes with respect to the frameworks for characterizing uncertainty surveyed in this thesis are also addressed.

The first part concludes with the introduction of a new framework for analyzing the decision-making problem with respect to the uncertainty attached to it and to the MOEA-based solving method. This framework named *Analysis of Uncertainty and Robustness in Evolutionary Optimization* or AUREO, heavily relies upon the taxonomy of robustness just introduced. A comprehensive analysis of the contributions presented in the literature as well as original contributions of this thesis in the field of robustness, along with its application to MOEA, underpin the analytical framework proposed. The AUREO is designed to help the proper identification of the mathematical program derived from the decision-making problem as well as the sources of uncertainty associated with the MOEA and the objective functions. This way, the analyst can select or design the corresponding MOEA as well as to evaluate strategies to improve its efficiency.

In the second part of this thesis, the application of the AUREO to three dependability problems related to reliability, scheduling and vulnerability analysis respectively is presented. In the reliability problem the use of AUREO leads to focus on improving the efficiency of the MOEA, which is done by means of an encoding technique named *percentage representation*. The scheduling problem, on the other hand, required to redefine the mathematical program as well as the metaheuristic in order to produce robust solutions. Finally, the use of the AUREO in the vulnerability analysis problem, along with making possible to reformulate the original single-objective problem like one of two objectives with uncertainty handling, evidenced new challenges in the Evolutionary Multiple-objective Optimization field. The problems studied are of interest to the evolutionary multiple-objective optimization as well as to the dependability communities.

Resumen

En términos generales, la toma de decisiones consiste en hallar el subconjunto de soluciones que optimizan el grado de satisfacción de la Unidad de Decisión, con respecto a algún criterio de calidad, dentro del dominio factible de alternativas. Cuando se emplea más de un criterio para medir la calidad de las soluciones, entonces se habla de un problema de toma de decisiones multicriterio.

A fin de determinar el subconjunto de soluciones óptimas multicriterio, también llamado frontera de Pareto, muchos analistas emplean métodos heurísticos o evolutivos como los Algoritmos Evolutivos de Optimización Mulitobjetivo (MOEA en inglés), dada la capacidad de los mismos para resolver problemas con dominios discontinuos y/o no lineales, así como también para emplear funciones de tipo caja negra.

En esta tesis se abordan los diferentes problemas causados por la presencia de incertidumbre en procesos de toma de decisiones multicriterio, apoyados en métodos evolutivos. A tal fin, se hacen contribuciones metodológicas y algorítmicas relativas al manejo de incertidumbre, en lo referente a la formulación de los programas matemáticos derivados de los problemas de toma de decisiones, así como también acerca de la estructura y funcionamiento de los MOEA empleados para solucionar dichos programas.

En la primera parte de esta tesis se analizan el significado de la incertidumbre, sus fuentes, sus tipos, y las diferentes partes de un sistema que pueden ser afectadas por la incertidumbre. Adicionalmente, se presenta una revisión de los principales marcos teóricos existentes para representar la incertidumbre. Por otro lado, la estructura y operación de los MOEA de propósito general son analizados, considerando desde aspectos básicos de algoritmos evolutivos hasta el estado del arte en diseño de MOEA. De dicho estudio se desprende que los procedimientos de jerarquización (*ranking*) y almacenamiento (*archiving*) de soluciones que conforman los MOEA, fueron identificados como los elementos clave que deben ser modificados con el fin de manejar la incertidumbre. La estructura de dichos procedimientos es estudiada respecto a las formas de caracterización de incertidumbre no probabilística, probabilística y posibilística, al igual que las basadas en lógica difusa y probabilidades imprecisas. Como resultado, se ofrecen conclusiones en torno a las posibles alternativas para el manejo de incertidumbre con los métodos de caracterización mencionados.

Entre las contribuciones de la tesis, se presenta un MOEA diseñado para trabajar con incertidumbre no probabilística, llamado IP-MOEA. Por otra parte, se contemplan futuros desarrollos en el área de diseño de MOEA a través de la comparación del criterio de mínimo arrepentimiento (*minimal regret*) con el concepto de ϵ-dominancia, como alternativa para el manejo de incertidumbre no probabilística. De igual modo, se sugiere una reinterpretación probabilística

del concepto de densidad de un hipercubo, en el almacenamiento mediante *hypergrid*, para el almacenamiento de soluciones con incertidumbre probabilística.

La idea de robustez es cuidadosamente estudiado, sobre todo en lo tocante a optimización. Se ofrece un estado del arte en el concepto de robustez en toma de decisiones. Posteriormente se introduce una nueva taxonomía de clases de robustez, fundamentadas en la información disponible. También se abordan las posibles interpretaciones de dichas clases de robustez, respecto a los métodos de caracterización de incertidumbre considerados en esta tesis.

La primera parte del trabajo concluye con la introducción de un nuevo marco metodológico para el análisis de problemas de toma de decisiones, con respecto a las incertidumbres asociadas tanto al problema en sí como al MOEA empleado para solucionarlo. Este marco teórico, llamado *Análisis de incertidumbre y robustez en Optimización Evolutiva* o AUREO, se apoya en la taxonomía de clases de robustez propuesta en esta tesis. Un profuso análisis de las contribuciones presentes en la literatura, junto con las contribuciones introducidas en esta tesis en materia de robustez, fundamenta el marco teórico propuesto.

La metodología AUREO está diseñada para ayudar en la correcta identificación del programa matemático derivado del problema de toma de decisiones, así como también las fuentes de incertidumbre asociadas al MOEA y a las funciones objetivo. De este modo, el analista puede seleccionar o diseñar el MOEA correspondiente a la vez que evaluar las estrategias para mejorar su eficiencia.

En la segunda parte de la tesis, se presenta la aplicación de AUREO a tres problemas de seguridad de funcionamiento, relacionados con fiabilidad de sistemas, planificación y análisis de vulnerabilidad. En el problema de fiabilidad, el uso de AUREO permite concentrar los esfuerzos en mejorar la eficiencia del MOEA, lo cual es posible por medio de una técnica de representación llamada *representación porcentual*. Por otro lado, en el problema de planificación se requirió redefinir tanto el programa matemático como el método heurístico empleado, a fin de producir soluciones robustas. Finalmente, el uso de AUREO en el problema de análisis de vulnerabilidad, además de hacer posible la reformulación del problema original de un objetivo, en uno bi-objetivo, permitió identificar nuevos retos en el área de Optimización Evolutiva Multiobjetivo. Los resultados obtenidos son de interés, no solamente para el área de optimización evolutiva multiobjetivo, sino también para la el área de seguridad de funcionamiento.

A Dios, mi Padre y mi Amigo:
ojalá me regales la eternidad para darte gracias
por la hermosa vida que me das.

"Y Daniel dijo: «Te has acordado de mí, Dios mío, y no has abandonado a los que te aman.»"
Dan 14,28.

A mi hermano Gustavo y a Vero,
quienes seguramente dirán que mi tesis es incomprensible, pero que tanto han hecho a lo largo de los años para que yo pueda alcanzar las cosas que me hacen feliz, como este trabajo.

Acknowledgements

Faced with the possibility of pursuing my PhD in my country or leaving to study abroad, I was advised to live in another country because 'the lessons that one learns far from home nobody gets at the university'. I'll always be grateful to my friend for sharing with me such a nice piece of wisdom.

Since the very beginning when I learnt about a PhD offer at SIANI Institute (formerly IUSIANI), and all along these years, two people have supported me in all possible ways. They are my advisers and primarily my friends: Blas Galván and Claudio Rocco. To them I owe all goals I have reached during my PhD, and many others yet to come. I'll be eternally in debt with you for your help and friendship.

Many others have contributed immensely to my academical and personal development. Above all I would like to thank Prof. Gabriel Winter and all the team of CEANI for their help and selfless support. I would also like to express my gratitude with no special order to Dr. Marc Sevaux, Dr. Enrico Zio, Dr. Gregory Levitin, Dr. Sebastián Martorell, Dr. Jürgen Branke, Dr. Billan M. Ayyub, Prof. Bernard Roy and Dr. Máximo Méndez for collaborating on the research that underpinned this dissertation. I would also like to recognize Philipp Limbourg for coauthoring the proposal described in section 4.2.1.

Prof. Xavier Gandibleux from LINA, University of Nantes in France, deserves a special mention who generously offered me the chance of making a research visit at LINA under his supervision. This stay was granted by the 'Consejería de Educación' of the Government of the Canary Islands (B.O.C. N-148§1093 29.07.05). Further research related to this period was supported by 'Fundación Universitaria Las Palmas', through grant 'Beca Innova 2005'.

The human, spiritual and sometimes economical support of the following people have played a crucial role in my life during these years. To all of them my eternal gratitude:

To the community 'Nuestra Señora del Atlántico' of Las Palmas, for their openness, friendship and caring attention.

To Martine and Bérénice Poher for sharing their lives and being a like a family for me in France: vous m'avez accueilli dans votre maison et surtout dans vos cœurs, merci.

To my beloved family, for giving up the joy of being together in exchange for a better life for me.

To Irina, for filling up all these years with all kinds of beautiful sounds.

Contents

1 Introduction **1**
1.1 Outline 3
1.2 List of publications 3

I Background and fundamentals for the AUREO **7**

2 Uncertainty **9**
2.1 Uncertainty: an opening discussion 9
2.1.1 *What?–Information* 10
2.1.2 *Why?–Purpose* 13
2.1.3 *Obtainable?–Nature* 15
2.1.4 *Reducible?–Price* 15
2.1.5 *Where?–Sources and Influence* 16
2.2 Main Theories of Uncertainty 17
2.2.1 Mathematics of Uncertainty 17
2.2.2 Fuzzy Logic 21
2.2.3 Possibility Theory 24
2.2.4 Probability Theory 25
2.2.5 Dempster-Shafer Evidence Theory 28
2.2.6 Imprecise Probabilities 31
2.2.7 Probability Boxes 32
2.2.8 Info-gap Models 33
2.3 Generalized Theoretical Frameworks 33
2.4 Summary of the chapter 34

3 Evolutionary Optimization-based MCDM **35**
3.1 Multiple-Criteria Decision-Making 36
3.1.1 Basics of Decision-Making 36
3.1.2 Multiple-Objetive Optimization 38
3.2 Evolutionary Algorithms 39
3.2.1 Fundamentals 39
3.2.2 Structure of EA 41
3.2.3 Representation and Sampling Operators in EA 42
3.2.4 Representation and Recombination in our EA-based Research 44
3.2.5 A Proposal: Percentage Representation 45

3.2.6 Search Space Analysis and Percentage Representation: An Industrial Application Example . . . 46
3.3 Multiple-Objective Evolutionary Algorithms . . . 48
3.3.1 Historic Outline . . . 49
3.3.2 Algorithmic Structure . . . 51
3.3.3 ϵ-Dominance . . . 55
3.3.4 Some Representative MOEA . . . 55
3.3.5 Performance Measures . . . 59
3.4 Uncertainty Handling in EMO . . . 61
3.4.1 Probabilistic Dominance-based Procedures . . . 62
3.4.2 Quality Indicator-based Procedures . . . 63
3.4.3 Robust Optimization Procedures . . . 63
3.4.4 Probability bounding approaches . . . 64
3.4.5 Non-Probabilistic Procedures . . . 64
3.4.6 Metrics and Quality Assessment under Uncertainty . . . 64
3.5 Summary of the chapter . . . 65

4 Analysis of Uncertainty and Robustness in Evol. Opt. 67
4.1 An insight into uncertainty in decision-making . . . 67
4.1.1 Uncertainty induces comparisons of sets . . . 67
4.1.2 Criteria to compare sets . . . 70
4.1.3 Comparison of sets with no extra information . . . 71
4.1.4 Comparison of sets with extra information . . . 75
4.2 Proposals for Uncertainty Handling in EMO . . . 79
4.2.1 An Optimization Algorithm for Imprecise Multiple-Objective Problem Functions: IP-MOEA . . . 79
4.2.2 The ϵ-indicator revisited through the concept of minimal regret . . . 82
4.2.3 A probabilistic interpretation of the hypercube's density . 84
4.3 Robustness: from concepts to classes . . . 86
4.3.1 What do we mean when we say that something is robust? 87
4.3.2 Robustness concept: A survey . . . 90
4.3.3 Robustness: many concepts and a few classes . . . 92
4.4 Information based Perspective for Robustness Analysis . . . 101
4.4.1 AUREO Stage 1: From the analysis of interactions to the formulation of a new program . . . 102
4.4.2 AUREO Stage 2: From the uncertainty-handling program to an efficient implementation . . . 104
4.5 Summary of the chapter and concluding remarks . . . 107

II Applications of the AUREO + MCDM in Dependability: Theoretical and Algorithmic Issues 109

5 AUREO + MCDM in Reliability and Scheduling 111
5.1 Introduction . . . 111
5.2 AUREO in Reliability . . . 112
5.2.1 Basics of Reliability Engineering . . . 112
5.2.2 Statement of the problem . . . 112
5.2.3 AUREO: Stage 1 . . . 113

5.2.4 AUREO: Stage 2 and results 115
5.3 AUREO in Scheduling 119
5.3.1 Basics of scheduling theory 119
5.3.2 Statement of the problem: The RCb hybrid flowshop bi-objective scheduling problem 120
5.3.3 AUREO: Stage 1 122
5.3.4 AUREO: Stage 2 and results 125
5.4 Summary of the chapter and concluding remarks 128

6 AUREO + MCDM in Vulnerability Analysis 129
6.1 Introduction 129
6.2 Statement of the problem 130
6.2.1 The decision-making problem 131
6.2.2 Instances analyzed 133
6.2.3 The simulation tool 133
6.3 Application of the AUREO 134
6.3.1 Description of the problems 135
6.3.2 AUREO: Stage 1 137
6.3.3 AUREO: Stage 2 and results 139
6.3.4 What if the information changes? 144
6.4 Summary of the chapter and concluding remarks 145

7 Conclusions and further research 147

A Robustness concept: A survey 151
A.1 Forewords 151
A.2 Questionnaire 151
A.3 General comments about the questionnaire 154
A.4 Answers 154
A.4.1 B. M. Ayyub's answers 155
A.4.2 J. Branke's answers 156
A.4.3 B. Roy's answers 157

B Incertidumbre y la robustez en MCDM basada en EA 159
B.1 Motivación y objetivos de la tesis 159
B.2 Conceptos fundamentales de incertidumbre 161
B.2.1 Discusión inicial 161
B.2.2 Marcos teóricos para la representación de la incertidumbre 163
B.2.3 Lógica difusa 165
B.2.4 Teoría de la posibilidad 166
B.2.5 Teoría de la probabilidad 167
B.2.6 Teoría de la evidencia de Dempster-Shafer 169
B.2.7 Probabilidades imprecisas y p-box 170
B.2.8 Modelos Info-gap 171
B.3 Conceptos fundamentales de Optimización Evolutiva Multiobjetivo171
B.3.1 Fundamentos de toma de decisiones y optimización multiobjetivo 171
B.3.2 Algoritmo evolutivos de optimización multiobjetivo 173
B.3.3 Manejo de incertidumbre en EMO 180
B.4 Contribuciones al análisis de incertidumbre y robustez en EMO . 182

B.4.1 Análisis de efectos de la incertidumbre en MOEA 183
B.4.2 Propuestas para incorporar manejo de incertidumbre en MOEA 189
B.4.3 Contribuciones acerca del concepto de robustez 195
B.4.4 AUREO: Análisis de robustez desde una perspectiva basada en la información disponible 200
B.5 Aplicación y validación de AUREO en problemas de Seguridad global 206
B.5.1 Aplicación de AUREO a un problema de fiabilidad 206
B.5.2 Aplicación de AUREO a un problema de planificación . . 211
B.5.3 Aplicación de AUREO a problemas de análisis de vulnerabilidad 215
B.5.4 AUREO Etapa 2 218
B.6 Conclusiones y trabajo futuro 219

Bibliography **241**

List of Figures

2.1 Smithson's taxonomy of ignorance. 11
2.2 Adaptation of Smeth's thesaurus on imperfection. 12
2.3 Klir's taxonomy of uncertainty. 12
2.4 Model building and resolution stages in decision-making. 14
2.5 Instances and sources of uncertainty in a systemic view. 16
2.6 Example of rough sets. 20
2.7 Fuzzy sets representing 'cold' and 'very cold'. 22
2.8 Example of membership functions 'around 2' of three fuzzy sets. 22
2.9 Example of bpa, Bel and Pls in Dempster-Shafer Theory. 31
2.10 An example of a p-box. 32

3.1 Algorithm 1: Evolutionary Algorithm. 41
3.2 Flowchart of an EA. 41
3.3 Example of the crossover mechanism for binary chromosomes. . . 43
3.4 Example of the crossover mechanism for mixed chromosomes. . . 45
3.5 Triangular mutation operator. 46
3.6 High Pressure Injection System. 47
3.7 A model for MCDM and EMO. 49
3.8 Historic development of EMO. 50
3.9 Flowchart of a 2nd MOEA. 51
3.10 Algorithm 2: Infinite Size Archiving Algorithm. 53
3.11 Algorithm 3: Finite Size Archiving Algorithm. 54
3.12 Hypergrid archiving strategy. 54
3.13 Quality assessment in NSGA-II. 56
3.14 Algorithm 4: NSGA-II 57
3.15 Dominance rank + count strategy in SPEA and SPEA2. 58
3.16 Algorithm 5: SPEA2. 58
3.17 Algorithm 6: MOSA 59
3.18 Example of the hypervolume or S metric. 60

4.1 Different spaces of decisional processes. 68
4.2 Coarse mapping from reality onto decisional space. 68
4.3 Propagation of aleatory uncertainty. 69
4.4 Effect of epistemic uncertainty in decision-making. 69
4.5 Five possible cases of intervals comparisons in MOEA ranking under uncertainty. 73
4.6 Examples of possible cases of bi-dimensional intervals comparisons in MOEA ranking. 75

4.7 Examples of intervals comparisons. 80
4.8 Examples of multiple-objective intervals comparisons. 81
4.9 Comparison of the ϵ-dominance and the minimal regret concepts. 84
4.10 Comparison of the ϵ-dominance and the minimal regret concepts for uncertain objective vectors. 84
4.11 Example of an hypergrid with decision vectors with uncertain objectives. 85
4.12 Image, expected value and standard deviation of $f(x)$ (example 4.1). 89
4.13 Trade-off between standard deviation and mean (example 4.1). . 89
4.14 Trade-off between COV and mean (example 4.1). 89
4.15 Examples of different inner hyper-boxes for specified and unspecified centres 98
4.16 AUREO: The two levels or stages of analysis. 102
4.17 AUREO Stage 1: Analysis of interactions between model and uncertainties. 103
4.18 AUREO Stage 2: Elements of model formulation for uncertainty handling. 105

5.1 Schematic of the Residual Heat Removal system (RHR) of a Nuclear Power Plant. 114
5.2 Schematic of a Residual Heat Removal system. 114
5.3 Pareto approximation frontier robustness-cost for a design centred in $c_i = 0.90$. 116
5.4 Pareto approximation frontiers robustness-cost for non-specified centre using a double-loop approach. 117
5.5 Pareto approximation frontiers robustness-cost for non-specified centre using percentage representation. 119
5.6 RCb scheduling problem in a waste treatment plant (two-stage example). 120
5.7 A real instance of ready and processing times for each truck. . . 121
5.8 Example of using MOSA and MNEH over a real instance. 122
5.9 Example of ϵ-dominance used as threshold of indifference. 126
5.10 Results filtering the optimal solutions in two stages. 127
5.11 Results of solving the problem as a three-objective program. . . . 127

6.1 Example of propagation through networks. 131
6.2 Algorithm 7: Simulation tool based on cellular automata and Monte Carlo simulation. 135
6.3 Topology of Network 1: a real telephone network. 135
6.4 Topology of Network 2: the US airports network. 136
6.5 Attacker's efficient frontier for Network 1 without protections . . 140
6.6 Selected antagonist's efficient frontiers in function of the protected node of Network 1. 142
6.7 Antagonist's efficient frontier provided two attacks on Network 1. 142
6.8 Antagonist's efficient frontiers up to three simultaneous attacks on Network 1. 142
6.9 Antagonist's efficient frontiers when node 24 of Networks 1 is protected. 143

6.10 Selected antagonist's efficient frontiers if function the protected node of Network 2. 144
6.11 Cost and impact of attacking nodes of Network 1 from the antagonist perspective. 145
6.12 Selected min-max frontiers for risk reduction of Network 1 against single attacks. 146

B.1 Visión sistémica de las instancias y fuentes de incertidumbre. . . 163
B.2 Algoritmo B.1: Algoritmo evolutivo. 175
B.3 Diagrama de un MOEA de 2ª generación. 176
B.4 Algoritmo B.2: Algoritmo de almacenamiento en archivo finito. . 176
B.5 Almacenamiento mediante *Hypergrid*. 177
B.6 Evaluación de la calidad en NSGA-II. 178
B.7 Algoritmo B.3: NSGA-II 178
B.8 Algoritmo B.4: MOSA 179
B.9 Diferentes espacios del proceso de decisión. 183
B.10 Aleatoriedad, preferencias y espacio de decisión. 183
B.11 Propagación de incertidumbre aleatoria. 184
B.12 Efecto de la incertidumbre epistémica en toma de decisiones. . . 184
B.13 Cinco posibles casos de comparación de intervalos en jerarquización de soluciones bajo incertidumbre usando MOEA. 185
B.14 Ejemplos de posibles casos de comparación de intervalos bidimensionales en jerarquización de soluciones usando MOEA. 187
B.15 Ejemplo de comparación de intervalos. 189
B.16 Ejemplos de comparación de intervalos multiobjetivo. 190
B.17 Comparición de la ϵ-dominancia y el arrepentimiento mínimo. . . 192
B.18 Comparición de la ϵ-dominancia y el arrepentimiento mínimo para vectores objetivo inciertos. 193
B.19 Ejemplo de almacenamiento mediante *hypergrid* con vectores objetivo inciertos. 193
B.20 Ejemplos de distintas cajas internas para centroides especificados y no especificados. 200
B.21 AUREO: Dos niveles de análisis. 201
B.22 AUREO Fase 1: Análisis de interacciones modelo-incertidumbre. 202
B.23 AUREO Fase 2: Elementos de la formulación de modelos para el manejo de incertidumbre. 204
B.24 Frente de aproximación de Pareto robustez-costo con centroide $c_i = 0.90$. 208
B.25 Frentes de aproximación de Pareto robustez-coste para centroides no especificado, usando el método de doble lazo. 210
B.26 Fronteras de aproximación de Pareto robustez-coste para centroide no especificado usando representación porcentual. 211
B.27 Resultados filtrando los óptimos en dos etapas. 214
B.28 Resultados resolviendo el problema como un programa de 3 objetivos. 215
B.29 Selección de fronteras efficientes para el atacante en función del nodo protegido para la Red 1. 219
B.30 Selección de fronteras efficientes para el atacante en función del nodo protegido para la Red 2. 220

List of Tables

2.1 Keywords to analyze uncertainty. 10
2.2 Operations on Information. 11
2.3 Operations on classical sets. 19
2.4 Common Probability Distribution Functions for continuous random variables employed in engineering. 29

3.1 Basics of Decision-Making. 38
3.2 The abstraction of Darwinian postulates in EA 40
3.3 State of the art in MOEA design for optimization under uncertainty. 62

4.1 Example of the minimum regret criterion for discrete problems. . 73
4.2 Conclusions drawn by five non-probabilistic criteria in MOEA ranking with intervals. 74
4.3 Conclusions drawn by five non-probabilistic criteria in MOEA ranking with multiple-objective intervals. 74
4.4 Imprecise-valued test functions ZDT_I (adapted from ZDT1, ZDT2, ZDT4, ZDT6). 82
4.5 Hypervolume values after 100 generations for imprecise-valued test functions ZDT_I. 83
4.6 Answers to the questionnaire about robustness. 91
4.7 Some examples of Class 1 $R(\cdot)$ realizations in the literature. . . . 95
4.8 Some contributions to the calculation of the mean and variance of fuzzy or imprecise probability quantities that can help towards the resolution of Class 1 $R(\cdot)$ functions. 96
4.9 AUREO Stage 1: Classes of uncertainty-handling formulations according to the available information. 104
4.10 AUREO Stage 2: Existing alternatives and possible extensions in EMO to solve the uncertainty-handling formulations prescribed in tab. 4.9. 106

6.1 Parameters for the number of people affected by an attack to a single node in the network shown in fig. 6.3. 136

B.1 Palabras claves para analizar la incertidumbre. 162
B.2 Operaciones relativas a la información. 162
B.3 Fundamentos de toma de decisiones. 173
B.4 Abstracciones de los postulados Darwinistas en EA 174

B.5 Estado del arte en diseño de MOEA para optimización con incertidumbre. 181
B.6 Conclusiones de la aplicación de cinco criterios no probabilísticos en jerarquización bajo incertidumbre usando MOEA. 186
B.7 Hipervolumen luego de 100 generaciones para los valores inciertos de las funciones test ZDT_I. 191
B.8 Algunos ejemplos de funciones $R(\cdot)$ Clase 1 en la literatura. . . . 198
B.9 Algunas contribuciones al cálculo de media y varianza de cantidades difusas o probabilísticas imprecisas, que podrían ayudar a resolver funciones $R(\cdot)$ Clase 1. 199
B.10 AUREO Fase 1: Clases de formulaciones para manejar incertidumbre en función de la información disponible. 203
B.11 AUREO Fase 2: Alternativas existentes y posibles ampliaciones en EMO para resolver las formulaciones de manejo de incertidumbre prescritas en la tabla B.10. 205

Acronyms

The main acronyms used in this work are presented below. Plurals are spelled the same.

AUREO	Analysis of Uncertainty and Robustness in Evolutionary Optimization *Análisis de incertidumbre y robustez en Optimización evolutiva*
CDF	Cumulative Distribution Function *Función de distribución acumulada*
DM	Decision-maker Unidad de decisión
DST	Dempster-Shafer Evidence Theory *Teoría de la evidencia de Dempster-Shafer*
EA	Evolutionary Algorithm *Algoritmo evolutivo*
EMO	Evolutionary Multiple-objective Optimization *Optimización evolutiva multiobjetivo*
GA	Genetic Algorithm *Algoritmo genético*
GP	Genetic Programming *Programación genética*
IB	Inner Box *Caja interna*
iff	If and only if *Si y solo si*
MCDM	Multi-criteria Decision-Making *Toma de decisiones multicriterio*
MIB	Maximal Inner Box *Caja interna máxima*
MO	Multiple-Objective *Multiojetivo*
MOEA	Multiple-Objective Evolutionary Algorithm *Algoritmo evolutivo de optimización multiobjetivo*
MOSA	Multiobjective Simulated Annealing *Recocido simulado multiobjetivo*
NSGA-II	Nondominated Sorting Genetic Algorithm II *Algoritmo genético de ordenamiento de no dominadas II*
PDF	Probability Density Function *Función de densidad de probabilidad*
PMF	Probability Mass Function *Función de masa de probabilidad*

Chapter 1

Introduction

> *"I have no data yet. It is a capital mistake to theorize before one has data. Insensibly one begins to twist facts to suit theories, instead of theories to suit facts".*
>
> *A Scandal in Bohemia*
> by Sir Arthur Conan Doyle.

People have problems in every domain of their lives. Many of such problems are strongly related to satisfaction of needs and fulfilment of aspirations. When an entity (namely person or group of persons) faces a problem that can be solved in multiple ways, it is compelled to choose one final alternative, i.e. to make a decision.

Rather than making an arbitrary choice, any rational decision-maker (DM) would try to maximize its level of satisfaction by selecting that alternative that produces a higher return. The identification of such an alternative is, however, dependent on the way the problem is modelled.

By modelling we mean the way reality is represented. Modelling entails building a simplified description of a system of interest, which is only possible after a complete realization of the facets that are relevant to such system. If a problem can be properly modelled by a unique criterion, it means that only one aspect of the reality seems pertinent to measure the level of satisfaction, to guide the search towards the best solution. According to B. Roy [167] such a problem or system spontaneously tends to an extreme value of its single criterion. In consequence, optimization constitutes a main activity in the solution of *single-criterion paradigm* problems.

Nevertheless, the richness of possibilities and perceptions that human mind is able to conceive cannot be always reduced to nor reflected upon a single criterion. Instead, a more complex conception of reality characterized by multiple -and *conflicting*- criteria is an alternative way for modelling problems according to the *multiple-criterion paradigm* [167]. In this view different aspects of reality (criteria) are considered relevant to assess the level of satisfaction that any possible alternative may provide, the direction that a system or a process should follow. This paradigm constitutes the first concern of this thesis.

In a multiple-criteria problem there is not a unique alternative capable of maximizing the satisfaction of all the criteria, but a subset of alternatives, called *efficient* or *non-dominated*, each of which representing a consensus, a trade-off

[31]. Hence, the whole multiple-criterion decision-making process comprises three tasks: the modelling of the problem, the identification of the set of non-dominated alternatives and the final selection of one of them. The success of the process depends on the success in performing each one of the stages. Conversely, any hindrance to accomplish any step can affect the whole decision-making process.

In that sense, uncertainty rises as one of the major obstacles to decision-making since it hampers the identification of the efficient set. Such effect is due to many factors that are in close connection with the sources of uncertainty. Amongst others, some of these sources are: the model itself, the parameters of the model, the variables, and the procedures for assessing the model [146, 145].

In the first part of this thesis we are concerned about the effect of uncertainty in multiple-criteria decision making problems where evolutionary algorithms serve as solving tools. In particular, we aim at investigating two fundamental issues:

1. What are the options to classify the quality of the solving alternatives for multiple-objective problem when their performance measures are subjected to uncertainty? and,

2. What can be done to design new algorithms or to adapt the existing ones to conform with the answers to the previous question?

The premise that rules the whole investigation is that the available information determines the options to be considered in any case. Therefore, every analyst should start from analyzing what is known, what can be known and what cannot before selecting the solving tool. Afterwards, it should be determined the type of problem induced by uncertainty and the criteria to handle it. Finally, the evolutionary tool should be designed or tailored according with the chosen criteria. This methodology is called '*Analysis of Uncertainty and Robustness in Evolutionary Optimization*' (AUREO).

Uncertainty handling and robustness analysis are nowadays salient research fields in decision-making. Many contributions have been made during the last years regarding different topics of this concept [194, 190, 80, 91]. This thesis aims at the same direction. In particular we strive to encompass in AUREO the previous efforts [194, 91, 93, 92, 34, 200] in evolutionary computation bringing them out and getting them together into the same analytic tool. Moreover, we extend the scope of the efforts of our predecessors giving a glimpse into the incorporation of the more representative theoretical frameworks to represent uncertainty into the evolutionary algorithms.

The concept of robustness plays a central role in the AUREO. In consequence, one of the contribution of this thesis is to provide a deeper insight into distinct formulations and nuances of this idea. Afterwards a global vision of robustness is presented based on an information perspective. This view makes possible to analyze the different ways of modelling robustness in the context of the available information of the problem, as well as some issues that come along afterwards when the efficient set is sought with evolutionary algorithms.

The second part of this work, composed of two chapters, is devoted to the practical application of the AUREO in dependability problems, namely reliability and planning problems and vulnerability analysis respectively.

1.1 Outline

The contents of this thesis are organized in two parts, the first one (chapters 2 to 4) covers the state of the art in uncertainty representation and Evolutionary Multiple-objective Optimization and afterwards introduces the AUREO which is the main contribution of this thesis. The second part (chapters 5 and 6) is devoted to the study of practical problems linked to the different robustness formulations by applying them to real-life technical decision-making problems in the fields of reliability, planning and vulnerability analysis. The distribution of matters is as follows:

Chapter 2 presents an opening discussion about the different facets of uncertainty, then a a brief review of the mathematical concepts associated with uncertainty and a survey of the most representative theoretical frameworks for representing uncertainty, namely fuzzy logic, probability theory, Demspter-Shafer evidence theory, imprecise probability theories and Info-gap models, is presented. A reader familiarized with these concepts can skip this chapter.

Chapter 3 brings the basics of Multiple-Criteria Decision Making. Then it gives an introduction to Evolutionary Algorithms and to Multiple-Objective Evolutionary Algorithms (MOEA). Finally a reviews of the state of the art in Evolutionary Multiple-objective Optimization in the presence of uncertainty is given. The exposition about MOEA with and without the uncertainty topic is recommended even to experts since it comprises key concepts necessary to understand the subsequent work.

Chapter 4 is the core of this thesis and comprises the main contributions of this work. It analyzes the many options to cope with uncertainty in terms of its sources and the available information about the decision-making problem, and how these options affect the MOEA design. Concurrently, the notion of robustness is thoroughly studied regarding theoretical and algorithmic issues. Finally, the whole elements are organized into the *Analysis of Uncertainty and Robustness in Evolutionary Optimization* (AUREO) methodology.

Chapter 5 reports some applications of robust design in real life reliability and planning problems. This chapter exemplifies the whole use of AUREO from the mathematical formulation of the decision-making problem to some algorithmic efficiency issues tackled by the introduction of a special encoding named 'percentage representation'.

Chapter 6 applies AUREO to a salient topic named multiple-objective vulnerability analysis, giving new insights into this subject, and finally

Chapter 7 presents the concluding remarks.

1.2 List of publications

Most of the contributions of this thesis have been or are about to be published in international journals and conferences proceedings, viz.:

- Daniel E. Salazar A., Claudio M. Rocco S., Enrico Zio (Forthcoming): Optimal Protection of Complex Networks Exposed to Terrorist Hazard: A Multi-Objective Evolutionary Approach. *Proceedings of the Institution of Mechanical Engineers, Part O, Journal of Risk and Reliability.*

- Claudio M. Rocco S., Daniel E. Salazar A., Enrico Zio (Forthcoming): Application of Advanced Computational Techniques to the Vulnerability Assessment of Network Systems Exposed to Uncertain Harmful Events. Special Issue on "System Survivability and Defense against External Impacts". *International Journal of Performability Engineering.*

- Daniel E. Salazar A., Claudio M. Rocco S , Enrico Zio (2007): Robust Reliability Design of a Nuclear System by Multiple Objective Evolutionary Optimization. Special Issue on "Soft Computing Methods in Nuclear Engineering". *International Journal of Nuclear Knowledge Management* 2(3):333–345.

- Daniel E. Salazar A. and Claudio M. Rocco S. (2007): Solving Advanced Multi-Objective Robust Designs By Means Of Multiple Objective Evolutionary Algorithms (MOEA): A Reliability Application. *Reliab Eng Sys Saf* 92(6):697–706.

- Daniel E. Salazar A., Claudio M. Rocco S. and Blas J. Galván G. (2006): Robustness Analysis: An Information-Based Perspective. *Newsletter of the European Working Group «Multiple Criteria Decision Aiding»* (Robustness Analysis Forum). Series 3, No 14, Fall 2006:17–22.

- Claudio Rocco and Daniel Salazar (2007): A Hybrid Approach Based on Evolutionary Strategies and Interval Arithmetic to Perform Robust Designs. In: S. Yang, Y. Ong, and Y. Jin (editors), *Evolutionary Computation in Dynamic and Uncertain Environments.* Studies in Computational Intelligence, Vol. 52, Springer, ISBN-10: 3-540-49772-2, ISBN-13: 978-3-540-49772-1 : Chapter 22.

- Daniel Salazar Aponte, Xavier Gandibleux, Julien Jorge and Marc Sevaux (2006): A robust-solution-based methodology to solve multiple-objective problems with uncertainty. To appear in the MOPGP'06 post-conference volume - LNEMS, Springer.

- Enrico Zio, Claudio M. Rocco S., Daniel E. Salazar A. and Gonzalo Müller (2007): Complex Networks Vulnerability: A Multiple-Objective Optimization Approach. *The Annual Reliability and Maintainability Symposium RAMS 2007.* January, 22-24, 2007, Florida, USA.

- Claudio M. Rocco S., Enrico Zio and Daniel E. Salazar A. (2007): Multi-objective evolutionary optimisation of the protection of complex networks exposed to terrorist hazard. In: Terje Aven and Jan Erik Vinnem (Eds.) *Risk, Reliability and Societal Safety*, volume 1: Specialisation Topics. Taylor & Francis, ISBN: 978-0-415-44783-6:899–905.

- Daniel E. Salazar A., Claudio M. Rocco S. and Enrico Zio (2006): Multiple Objective Evolutionary Optimisation for Robust Design. In: Da

Ruan, Pierre D'hondt, Paolo Fantoni, Martine De Cock, Mike Nachtegael, Etienne Kerre. *Applied Artificial Intelligence. Proceedings of The 7th Interntional FLINS Conference.* Genova, Italy, 29-31 August, 2006. World Scientific. ISBN 981-256-690-2 : 930–937.

- Philipp Limbourg and Daniel E. Salazar Aponte (2005): An Optimization Algorithm for Imprecise Multi-Objective Problem Functions. *IEEE Congress On Evolutionary Computation (CEC 2005).* Edinburgh, UK, September 2-5, 2005.

Part I

Background and fundamentals for the AUREO

Chapter 2

Uncertainty

> *"When one admits that nothing is certain one must, I think, also admit that some things are much more nearly certain than others. [...] Certainly there are degrees of certainty, and one should be very careful to emphasize that fact, because otherwise one is landed in an utter skepticism, and complete skepticism would, of course, be totally barren and completely useless".*
>
> From *Am I An Atheist Or An Agnostic?*
> by Bertrand Russell (1872 - 1970).

Information is the fundamental good in decision-making problems. Information is needed to describe problems and building models; to assess and forecast the performance of systems and their interaction with the environment; to analyze the quality of each alternative to solve a problem. In other words, the whole apparatus of a decision-making process is made of information. In that sense, information can be viewed as a constraint (cf. [227]) to intellectual operations and thus to decision-making.

By constraint we allude here to the intuitive meaning of some good or commodity that cannot be exploited in a perfect and unlimited way. Therefore, information is or acts as a constraint because it is normally limited, usually sparse, or even absent in some cases. Moreover, even having enough information, it cannot be exploited perfectly, due to the inherent cognitive and linguistic limitations of human beings to express, to understand and to handle information, enforcing decision-makers (DM) and analysts to recourse to different heuristics to make decisions in diverse information availability scenarios [198].

With the aforementioned ideas in mind, in the next sections we bring a survey of the theory of uncertainty and some techniques conceived to cope with it, i.e. we study the state of the art in dealing with lack of information. Afterwards we focus on a particular concept, related to systems fed by uncertain inputs, namely robustness, which constitutes the main concern of this thesis.

2.1 Uncertainty: an opening discussion

An elegant definition offered by Rowe states that: *"uncertainty is essentially the absence of information, information that may or not be obtainable"* [166, pg. 743]. However, it is easy to realize that the lack of information is a constant

Keywords	Question	Goal
What? Information	*What* do we need?	*Information*, or conversely, uncertainty reduction
Why? Purpose	*Why* do we need it?	*Purpose* of information or uncertainty reduction
Obtainable? Nature	Is that information *obtainable*?	*Nature* of uncertainty
Reducible? Price	Is uncertainty *reducible*?	*Price* of information
Where? Sources	*Where* does uncertainty *come from*?	*Sources* of uncertainty
Where? Influence	*Where* does uncertainty *affect* the system?	Systemic relations, *input–system–output*

Table 2.1: Keywords to analyze uncertainty.

for mankind: we ignore more than we know and we do not have full knowledge about anything (recall Socrates' famous quote *All I know is I know nothing*).

Even when uncertainty is ubiquitous, it does not imply that people feel surrounded by it. Actually, people's attitude towards uncertainty is more or less utilitarian: it is perceived as negative as long as it might carry undesirable consequences, otherwise it is called chance or it is simply neglected. This last point is due to an important fact that must be kept in mind in order to understand and to analyze uncertainty: information has a purpose and in consequence the way we handle its absence strongly relies on such purpose. However, purpose is not the only relevant matter regarding uncertainty, but the nature and the form of uncertainty itself must be investigated as well.

As a rule of thumb to analyze uncertainty, let us propose some key words that can lead to a better understanding of uncertainty in practice. These keywords and the associated concepts are summarized in Table 2.1. As we go along with this task, we shall survey the relevant facets of the current theories of uncertainty.

2.1.1 *What?–Information*

Our society is deemed to be an 'information society' concerned not only to account for the world but for knowledge itself [105]. This last follows the fact that knowledge is a finite recourse and therefore it must be handled in an efficient way. Our knowledge is *'the sum total of a system's learnings, memories, and experiences'* [72, pg. 1007], but since our ability to understand, to remember and to perceive is limited by nature, so is our knowledge.

Knowledge acquisition entails performing a group of operations on information, such as gathering, interpretation and transmission. These operations interconnect with each other and rarely follow a linear sequence. Failures or deficiencies on any of these operations induce uncertainty. Table 2.2 offers an insight into the classes of uncertainty derived from the imperfections on each class of operation. Part of the classification is adapted from [166].

Although paradoxical, the term uncertainty is a good example of translational uncertainty. It is broadly accepted that the word uncertainty means an imperfection in information, but authors disagree in what amount. According to

Operation Class	Activity Method	Uncertainty Class	Crucial Information
Information Gathering (Data collecting)	Measuring Research Observation	Metrical	New Measurements Stored Measurements Descriptions
Information Gathering (Data mining)	Prediction Retrodiction	Temporal	Future Past
Information Linking (Data composition)	Modelling	Structural	Complexity
Information Transmission (by Natural Language)	Communication and Descriptions	Translational	Concepts, Goals and Values

Table 2.2: Operations on Information (partially adapted from [166]).

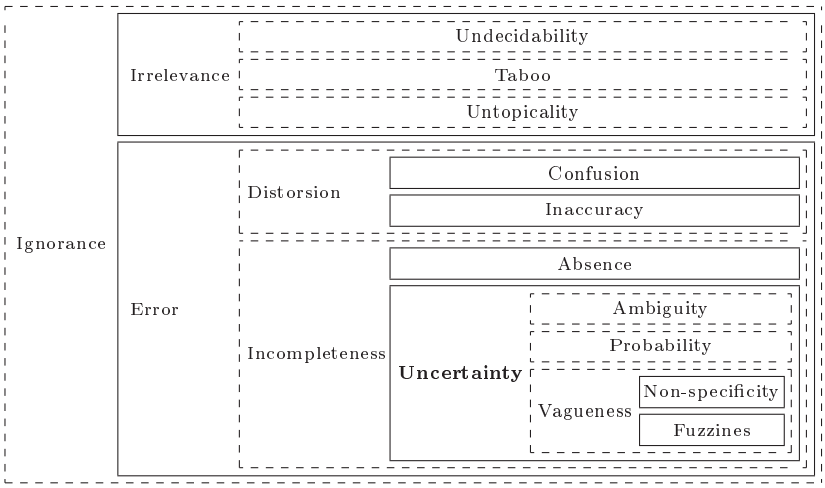


Figure 2.1: Smithson's taxonomy of ignorance (taken from [95, 188]).

Smithson, (cf. [95]) the higher level of imperfection is ignorance, from which we can identify two branches according with the willingness to be in such state: the deliberate act of ignoring (due to *irrelevance* of information) or the unwanted state of ignorance (*error*). In this view (see fig. 2.1), uncertainty refers to the degree of incomplete knowledge that results from the imperfections in the actual amount of true information. Ayyub and Klir offer a similar but comprehensive and more refined categorization of types of ignorance in [7, 6].

P. Smeth (cf. [95]) places uncertainty in higher level in his taxonomy of imperfection of information (see fig. 2.2); however in this view the term uncertainty is not yet at the highest tier. For Smeth, uncertainty is induced by either objective or subjective imperfections of information. In the former case it is a property of information whereas in the latter is a property of the observer [95].

Uncertainty can be fully identified with imperfection of information. This is the view of Krause and Clark, Bouchon-Meunier and Nguyen (see [95] for insight into these taxonomies) and Klir. This last author makes correspondence between information and uncertainty reduction, thus any piece of evidence is

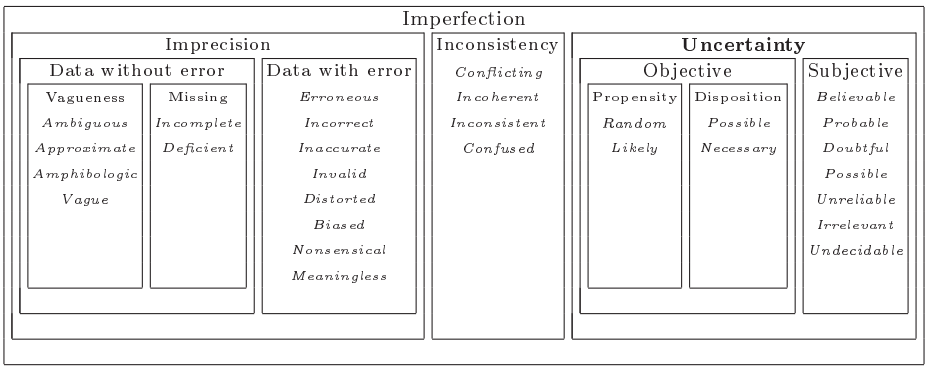


Figure 2.2: Adaptation of Smeth's thesaurus on imperfection (taken from [95]). Concepts are in plain text, synonyms in italic.

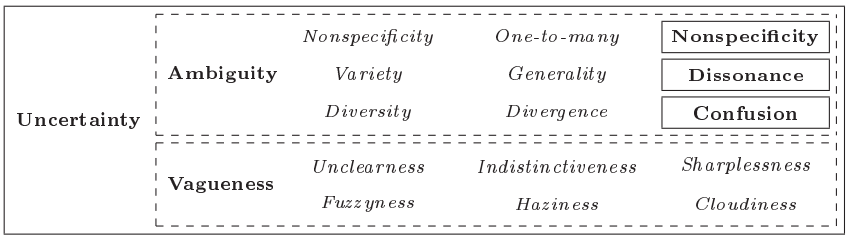


Figure 2.3: Klir's taxonomy of uncertainty [105] (Concepts are in bold text, synonyms in italic).

measurable and strictly equivalent to the reduction of uncertainty that results from the gain in information. This vision is deeply connected with the seminal work of Shannon [187] broadly known as *Classical Information Theory* and has been extensively developed in the *Generalized Information Theory* [104].

Klir distinguishes two types of uncertainty, namely vagueness and ambiguity (see fig. 2.3). The former concept refers to the difficulty of making sharp and precise observations of the world, evoking the notion of fuzzy sets. On the other hand, Klir detaches the idea of ambiguity into three possible defects in the available information, viz. nonspecificity, dissonance and confusion. Roughly speaking, given a piece of evidence and the subsets of possibilities related to the information in question, the concept of nonspecificity points out the size of the subset, thus the larger the subset, the less specific the information. Dissonance, on the other hand, designates the conflicts between the prospective disjoint subsets connected with the evidence. Finally, confusion is associated with the number of subsets in partial or total clash. These categories and their mathematical formulae are built on the notion of fuzzy measures [105].

The variety of conceptions and classification of information and uncertainty, as those reported above, evidences the difficulty of coping with the unknown, difficulty that any good analyst must be aware. As far as concerns to this work, from now on we stand on emphasizing objective uncertainty and to consider subjective uncertainty as a result of the former. In other words, we assert

that any rational DM has no reason to show mistrustful of information if it is perfect and complete and, conversely, any sign of suspicion is related somehow to imperfection of information. Our reading of knowledge is a combination of information and experience (cf. [72] and Davenport's view in [75, pg. 175]); likewise, we opt for considering information as data or evidence. Finally, we adopt the full identification of uncertainty with imperfection of information herein.

In practice, the challenge of any analyst in decision-making is to define the suitable conceptual framework within which the operations on information (cf. table 2.2) are going to be applied, and to fulfil as good as possible the following aims:

- To collect the larger amount of data,
- To identify the classes of imperfections associated with such data,
- To find a suitable way for representing any kind of uncertainty, and
- To minimize the effect of such uncertainty on the whole decision process and, if it is the case, on the chosen alternative itself.

2.1.2 *Why?–Purpose*

It is helpful to consider two temporal frameworks in order to appreciate the central role played by information in decision-making: present and future. The present does not refer here to the very moment when an action takes place, but a period of time when all the variables and parameters affecting the decision-making process remain steady. In other words, during the present, records are not to be updated, DM's preferences do not change, variables and parameters are not expected to exhibit some states different from the prior ones.

As a first abstraction, let us say that decision-making takes place in the present. The present, therefore, designates the time elapsed form the very moment the DM realize they have a problem (an unsatisfied need or an unfulfilled aspiration) to the moment when the solving alternative is definitively identified. Within this time window, the analyst assists the DM in modelling their desires and preferences in connection with the physical, economical and technological constraints imposed by the environment where the decision is to be made. This phase corresponds to the 'modelling stage' in fig. 2.4. Here information is relevant as long as it contributes to build or to improve the model of the problem.

Uncertainty comes along in this framework as the consequence of the imperfections of the operations performed on information by the analyst (cf. table 2.2): translational uncertainty results from the difficulty of capturing the whole meaning of the DM preferences when they are enquired to express them. Metrical uncertainty appears due to the errors and imprecision in past and current measures of the environment. Structural uncertainty derives from incorrectness or inadequacy in the model of the problem and its amount of complexity reduction. The global effect of uncertainty is that the model may mislead the analyst into identifying the optimal alternatives, leading to a final wrong choice.

Another characteristic concern during 'the present' is the uncertainty inherent to the resolution stage. Should the model be flawless, the resolution method may still introduce some uncertainty due to, e.g., its intrinsic inability to produce exact results (as is the case of Monte Carlo simulations) or the lack of

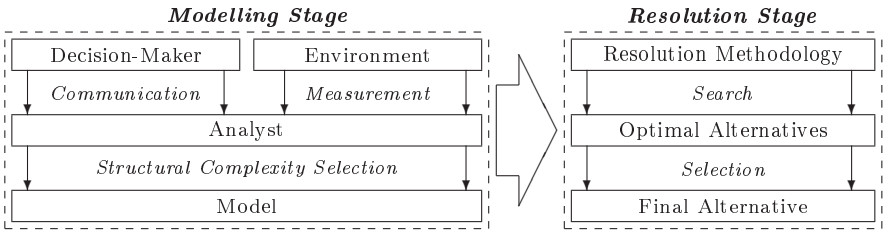


Figure 2.4: Model building and resolution stages in decision-making.

optimality proof (like in Genetic Algorithms applications). The analyst must be aware of the different sources of uncertainty and their implications in the process. In chapter 4 this view of a two-stage process is revisited to face not only the present but the hurdles that brings the other temporal framework: the future.

The validity of an alternative as final solution relies upon the circumstances that surrounded its selection. If the state of affairs changes such an alternative might be no longer appropriate. Time is the matter of change -and vice versa-, that explains the convenience of detaching temporal uncertainty from the rest of classes. The worry that characterizes the encounter with the future is the degree in which the current decisions will carry on fulfilling the decision-maker's needs and aspirations as times goes by: preferences may vary, variables and parameters definitively do.

Given a system, the typical strategy to hedge against temporal uncertainty is to anticipate eventualities and to identify their consequences, trying to minimize the regrettable ones: in other words, we seek to minimize risk. Risk is defined as a combination of possible consequences and the uncertainty associated with such consequences [5]. In particular, it is of interest the amount of damage or loss that any possible event can produce and its likelihood. In the context of risk analysis, such an insight is obtained answering to the following questions [97, 96]:

1. What can go wrong?

2. How likely is it that that will happen?

3. If it does happen, what are the consequences?

The scope of risk analysis is broader than that mentioned above, but the concept of risk as presented so far is quite enough to clarify the fundamental issues of temporal uncertainty. Information is necessary to attend to the internal and external features whose influence on a generic system may prevent it from accomplishing its mission. Once the unwanted scenarios are identified and their consequence properly quantified, policies of prevention, mitigation, transference and acceptance of risk may be delineated. This procedure is applicable to a wide range of systems and problems, albeit most of the people tend to associate it only with security and prevention of accidents.

The earlier questions of risk analysis formulated in terms of an optimal alternative in decision-making become:

1. What can change? (variable, parameter, preference)
2. How likely is that?
3. If it does happen, how much does it affect the optimality condition of the chosen alternative?

A worthwhile strategy to cope with risk is to promote the intrinsic ability of a system to resist to undesirable stimuli, lessen the consequences caused by its reactions. This approach recalls notions like resilience and robustness. By attending to the robustness of a system, whatever the system is, the analyst could contribute to decrease risk and consequently to increase the system's quality in the events to come. This way of thinking underpins some strategies to cope with uncertainty like Info-gaps models [13] and constitutes one of the basis of this thesis.

2.1.3 *Obtainable?–Nature*

There is a long tradition in some disciplines like engineering and risk management of discriminating uncertainty in terms of its nature. The underlying (practical) issue is that if uncertainty is reducible with more information or not. The answer leads to a two-class classification of uncertainty according to whether it is affirmative or negative.

The first class of uncertainty, known as Type I, systemic or *aleatory* uncertainty, distinguishes by a variability that cannot be reduced by further empirical effort. In other words, such an uncertainty does not rely upon the information available but on the inherent randomness and stochasticity of the system under study and its environment. It is therefore irreducible although it is still susceptible of being characterized [145, 146, 47, 50, 49].

On the other hand, the second class called Type II or *epistemic* uncertainty is the complement to the former. This kind of uncertainty is assumed to be reducible by additional information of the system or its environment [145, 146, 47, 50, 49].

Practitioners recognize the challenge of suitably represent epistemic uncertainty in opposition to aleatory uncertainty [145, 47]. As an example, consider the case of a unique unknown value bounded in [3,5] and a variable value uniformly distributed in the same interval. It is widely assumed that the former (epistemic) uncertainty can be represented with a uniform distribution U[3,5], i.e., assigning the same density of probability to all the possible values that the variable can take within this interval. However, this recourse makes only an artificial equivalence of our ignorance to a (uniform) random variable, for the true value is unique; in other words, we resort to a probabilistic reasoning to represent something that is not probabilistic in nature. Thus, with more information the assumption can change e.g. by pruning the interval. In contrast, if the variability associated with the latter variable is well characterized, no further improvements in the description of its behaviour are expected with more data.

2.1.4 *Reducible?–Price*

In practice, analysts and decision-makers should balance the price of information against its worth in terms of the benefits of uncertainty reduction. Investments

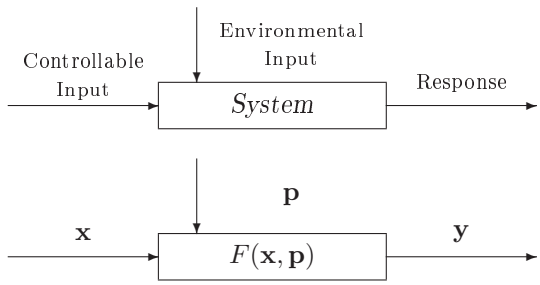


Figure 2.5: Instances and sources of uncertainty in a systemic view: the system (top) and its functional representation (bottom).

in information could be no longer justifiable beyond certain point at which the acceptance of the residual uncertainty is cheaper than its further reduction; especially when alternatives like searching for robust systems are possible. Probability distribution (see sec. 2.2.4) are a good example of what we meant about the price of information. Sometimes it is possible but not practical to collect all the data about some feature of a population, rather it is cheaper to sample it appropriately in order to characterize the population by means of a distribution, assuming the uncertainty introduced by giving up part of the available information.

The trade-off between information and price is noticeable not only in data gathering but in modelling. According to the *principle of minimum uncertainty* we should accept those models (and in general those solutions) that minimize the loss of information [103]. In practice, it corresponds to favouring more complex models that are, in most cases, more difficult to be solved. Therefore, analysts are paradoxical compelled to resort to solving heuristics that introduce uncertainty at resolution stage.

2.1.5 *Where?–Sources and Influence*

To clarify this point, consider that the interaction of a system with its surroundings can be modelled with a function of the type $F(\mathbf{x}, \mathbf{p})$ where $\mathbf{x}$ denotes the variables and $\mathbf{p}$ the environmental parameters (see fig. 2.5). Vectors $\mathbf{x}$ and $\mathbf{p}$ represent the sum total of elements entering the system that the modeller is aware of. The response of the system is denoted as a vector of attributes $\mathbf{y}$.

Uncertainty arises from many sources, being one of the most important the descriptive power of the model. Its complexity in terms of number of variables and parameters considered as well as the functional relationships between them is a key issue in uncertainty analysis. Another point that must be kept in mind is the instance of the universe that is subject and/or source of uncertainty. Uncertainty may come from the variables ($\mathbf{x}$), the environment ($\mathbf{p}$), the system itself or the way we assess it. If the system is suitably modelled, at least all the controllable inputs that significantly affect the system are included in vector $\mathbf{x}$. Some variables could be conveniently neglected for practical reasons or not considered for incorrectness of the model. On the other hand, there is always a set of environmental parameters that must be considered during the modelling in order to capture the relevant uncontrollable interactions that affect the system's

responses.

The assessment of $\mathbf{y}$ is subject to two instances of uncertainty, viz. the adequacy of the model $F(\mathbf{x})$ and the suitability of the calculations. Errors in computer codes and/or simulation models are sources of epistemic uncertainty (cf. [146]). Likewise, measures of $\mathbf{x}$ and $\mathbf{p}$ are usually subjected to epistemic uncertainty due to inherent limitations and incorrectness of the measuring procedures and devices. Besides, aleatory uncertainty must be considered and rightfully modelled, as resulting from environmental stochasticity, inhomogeneity of materials, fluctuations in time, variations in space [50, pg. 11], etc.

2.2 Main Theories of Uncertainty

Theories of uncertainty are meant to provide a coherent mathematical apparatus to express what we know and what we ignore about reality. However, since the sources and classes of imperfection in information and thus in our knowledge are varied, so they are the theories of uncertainty and the difficulties they try to overcome.

In the following subsections we survey the fundamental features of the most relevant systems for representing information, or conversely, uncertainty. The exposition begin with a short review of set theory, since this body of knowledge underpins the discussions hereafter. Afterwards the very systems are presented, highlighting the static and dynamic components of each system, i.e. the way we represent our knowledge or belief and the mechanism to update it given a new piece of information (as suggested in [192]).

2.2.1 Mathematics of Uncertainty

This section surveys the mathematic foundations that support the current theories of uncertainty. Most of the exposition follows [7, 105]. See these references for further details.

Sets

A *set* is any collection of distinct *elements* (also termed *individuals* or *members*) defined from a universe. Such a universe comprises the sum total of elements that pertain to a particular context, thus is named universe of interest, universe of discourse or simply universal set (usually denoted $\mathcal{S}$).

Relevant facts concerning a set are the way it is defined, its relation with elements, with other sets and the operations on them.

Sets and elements

A generic element x can belong ($x \in \mathcal{X}$) or not ($x \notin \mathcal{X}$) to a generic set $\mathcal{X}$. If the elements belonging $\mathcal{X}$ can be labelled by the positive integers the set is *countable*, otherwise it is *uncountable*. A set without any element is an empty set ($\emptyset$).

The total number of elements that make up a set is its *cardinality*. Countable sets can be *finite* or *countably infinite*; uncountable sets are *infinite*. The cardinality of a finite set of n elements is n ($|\mathcal{X}| = n$). An infinite set has a

cardinality equal to the size of its extension (e.g. $|\{x|a \leq x \leq b\}| = b - a$). Empty sets have cardinality zero.

Sets and operations on other sets

Operators on sets are to express the existing relations amongst them or to modify them, open the way for new sets. Giving sets $\mathcal{A}$ and $\mathcal{B}$, the following relations are possible:

- $\mathcal{A}$ is *equal* to $\mathcal{B}$ (denoted $\mathcal{A} = \mathcal{B}$): $\mathcal{A}$ and $\mathcal{B}$ comprise uniquely the very same elements,
- $\mathcal{A}$ is a *subset* of $\mathcal{B}$ (denoted $\mathcal{A} \subseteq \mathcal{B}$): All the elements of $\mathcal{A}$ are elements of $\mathcal{B}$,
- $\mathcal{A}$ is a *proper subset* of $\mathcal{B}$ (denoted $\mathcal{A} \subset \mathcal{B}$): All the elements of $\mathcal{A}$ are only a part of the total elements of $\mathcal{B}$.

Likewise, negations are still possible:

- $\mathcal{A}$ is *not equal* to $\mathcal{B}$ (denoted $\mathcal{A} \neq \mathcal{B}$): $\mathcal{A}$ and $\mathcal{B}$ differ in at least one element,
- $\mathcal{A}$ is *not a proper subset* of $\mathcal{B}$ (denoted $\mathcal{A} \not\subset \mathcal{B}$): $\mathcal{A}$ contains at least one element not contained in $\mathcal{B}$,
- $\mathcal{A}$ is *not a subset* of $\mathcal{B}$ (denoted $\mathcal{A} \not\subseteq \mathcal{B}$): $\mathcal{A}$ is either not equal or not a proper subset of $\mathcal{B}$ or both.

Other operators serve not only to express the current relations amongst sets but also to define news sets[1]:

- $\mathcal{A}$ *union* $\mathcal{B}$ (denoted $\mathcal{A} \cup \mathcal{B}$): The composite comprises all the elements of $\mathcal{A}$ and $\mathcal{B}$, i.e. $\mathcal{A} \cup \mathcal{B} = \{x|x \in \mathcal{A} \vee x \in \mathcal{B}\}$,
- $\mathcal{A}$ *intersection* $\mathcal{B}$ (denoted $\mathcal{A} \cap \mathcal{B}$): The composite contains only the elements common to $\mathcal{A}$ and $\mathcal{B}$, i.e. $\mathcal{A} \cap \mathcal{B} = \{x|x \in \mathcal{A} \wedge x \in \mathcal{B}\}$. If $\mathcal{A} \cap \mathcal{B} = \emptyset$, then $\mathcal{A}$ and $\mathcal{B}$ are *disjoint* sets.
- $\mathcal{A}$ *difference* $\mathcal{B}$ (denoted $\mathcal{A} - \mathcal{B}$): The composite is comprised only of those elements of $\mathcal{A}$ not comprised by $\mathcal{B}$, , i.e. $\mathcal{A} - \mathcal{B} = \{x|x \in \mathcal{A} \wedge x \notin \mathcal{B}\}$,
- $\mathcal{A}$ *complement*(denoted $\bar{\mathcal{A}}$): The complement indicates all the elements of the universal set not contained in $\mathcal{A}$, i.e. let $\mathcal{S}$ the universal set, thus $\bar{\mathcal{A}} = \{x|x \in \mathcal{S} \wedge x \notin \mathcal{A}\}$.

Additional operations on sets are reported in table 2.3.

Power Set

A *power set*, denoted $P_{\mathcal{X}}$, is the set of all possible different subsets that can be defined from a given set $\mathcal{X}$, included the empty set $\emptyset$ and the very set $\mathcal{X}$. For finite and discrete sets, the cardinality of the power set is given by $|P_{\mathcal{X}}| = 2^{|\mathcal{X}|}$.

[1] For the sake of convenience, in algorithmic operations on sets union and complement could be also denoted with + and ! respectively. Assignment is denoted with := operator.

Rule Type	Operations
Identity	$\mathcal{A} \cup \emptyset = \mathcal{A}$ $\mathcal{A} \cap \emptyset = \emptyset$ $\mathcal{A} \cup \mathcal{S} = \mathcal{S}$ $\mathcal{A} \cap \mathcal{S} = \mathcal{A}$
Idempotence	$\mathcal{A} \cup \mathcal{A} = \mathcal{A}$ $\mathcal{A} \cap \mathcal{A} = \mathcal{A}$
Complement	$\mathcal{A} \cup \bar{\mathcal{A}} = \mathcal{S}$ $A \cap \bar{A} = \emptyset$ $\bar{\bar{\mathcal{A}}} = \mathcal{A}$ $\bar{\mathcal{S}} = \emptyset$ $\bar{\emptyset} = \mathcal{S}$
Commutation	$\mathcal{A} \cup \mathcal{B} = \mathcal{B} \cup \mathcal{A}$ $\mathcal{A} \cap \mathcal{B} = \mathcal{B} \cap \mathcal{A}$
Association	$(\mathcal{A} \cup \mathcal{B}) \cup \mathcal{C} = \mathcal{A} \cup (\mathcal{B} \cup \mathcal{C})$ $(\mathcal{A} \cap \mathcal{B}) \cap \mathcal{C} = \mathcal{A} \cap (\mathcal{B} \cap \mathcal{C})$
Distribution	$(\mathcal{A} \cup \mathcal{B}) \cap \mathcal{C} = (\mathcal{A} \cap \mathcal{C}) \cup (\mathcal{B} \cap \mathcal{C})$ $(\mathcal{A} \cap \mathcal{B}) \cup \mathcal{C} = (\mathcal{A} \cup \mathcal{C}) \cap (\mathcal{B} \cup \mathcal{C})$
Absorption	$\mathcal{A} \cup (\mathcal{A} \cap \mathcal{B}) = \mathcal{A}$ $\mathcal{A} \cap (\mathcal{A} \cup \mathcal{B}) = \mathcal{A}$ $\mathcal{A} \cup \left(\bar{\mathcal{A}} \cap \mathcal{B}\right) = \mathcal{A} \cup \mathcal{B}$ $\mathcal{A} \cap \left(\bar{\mathcal{A}} \cup \mathcal{B}\right) = \mathcal{A} \cap \mathcal{B}$
de Morgan's laws	$\overline{\mathcal{A} \cup \mathcal{B}} = \bar{\mathcal{A}} \cap \bar{\mathcal{B}}$ $\overline{\mathcal{A} \cap \mathcal{B}} = \bar{\mathcal{A}} \cup \bar{\mathcal{B}}$

Table 2.3: Operations on classical sets [7, 105].

Characteristic Function and Types of Sets

The relation between a generic element x and a given set $\mathcal{X}$ can be expressed by means of a *characteristic function* $\mu : \mathcal{S} \rightarrow \mathcal{T}$, that discriminates whether the element belongs to $\mathcal{X}$ or not. If $\mathcal{T} = \{0, 1\}$ the set is a (classical) *crisp set* and μ strictly indicates if x is contained (1) or not (0) in $\mathcal{X}$:

$$\mu(x) = \begin{cases} 1 & : \quad \forall x \in \mathcal{X} \\ 0 & : \quad \text{otherwise} \end{cases} \tag{2.1}$$

On the other hand, if the qualification of membership is not binary but continuous, i.e. if the membership can be ranged from 0 to 1, where 0 means not being a member and 1 being a member, then $\mathcal{X}$ is a *fuzzy set*, and $\mu : \mathcal{S} \rightarrow [0, 1]$ a *membership function*. Fuzzy sets shall be treated with more detail later on in section 2.2.2.

Both crisp and fuzzy sets can be described in terms of *rough sets* approximations. A rough set is a set defined regarding another set under the requirement of being a *lower approximation* or a *upper approximation*. Therefore, a generic set $\mathcal{X}$ has lower $\underline{R}(\mathcal{X})$ and upper $\overline{R}(\mathcal{X})$ approximation sets. Fig. 2.6 shows an

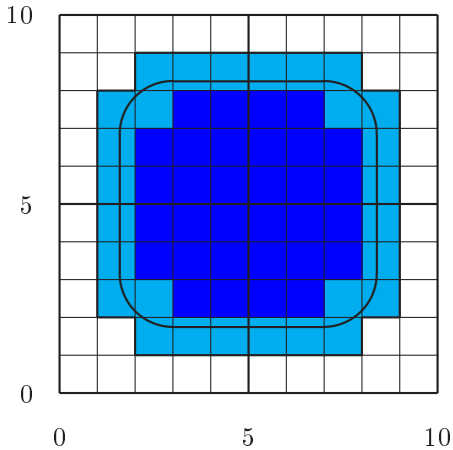


Figure 2.6: Example of rough sets: lower (dark shaded) and upper (light shaded) approximations to an oval shaped set.

example of rough sets. Notice that both approximation sets are defined from crisp partitions.

From the definition above it follows that $\underline{R}(\mathcal{X}) \subseteq \mathcal{X} \subseteq \overline{R}(\mathcal{X})$. For example, given $\mathcal{X} = \{x|x \in [1.5, 6.3]\}$ and a crisp partition equal one, the rough sets are $\underline{R}(\mathcal{X}) = \{x|x \in [2, 6]\}$ and $\overline{R}(\mathcal{X}) = \{x|x \in [1, 7]\}$. Larger upper sets and smaller lower sets are feasible (e.g. [0,8] and [3,5] respectively) but with a lesser quality of approximation. The accuracy of an approximation can be measured as

$$\delta = \frac{|\underline{R}(\mathcal{X})|}{|\overline{R}(\mathcal{X})|} \tag{2.2}$$

where δ ranges within [0,1], thereby the better the approximation the greater δ.

Finally, note that only partial assertions of the membership of elements are possible from rough sets, preventing any characteristic function from being an universal classifier. At most a function like

$$\mu(x) = \begin{cases} 1 & : \quad \forall x \in \underline{R}(\mathcal{X}) \\ 0 & : \quad \forall x \notin \overline{R}(\mathcal{X}) \\ ? & : \quad \text{otherwise} \end{cases} \tag{2.3}$$

–where '?' means '*don't know*'– is possible.

Crisp, fuzzy and rough set lay the foundations for a mathematical formalization of the most relevant theories of uncertainty employed in the realms of engineering, risk analysis and decision science in general.

Interval Arithmetic

To conclude this section, let us take a look to interval arithmetic (IA). IA is a mathematical framework to cope with imprecisions in computation. This framework is also valid for mathematical structures that cannot be identified with a single scalar, such as continuous closed sets in $\mathbb{R}$.

R. E. Moore [141] is recognized as the most influential researcher in this field, although the prior attempts in IA go back forty years previous his PhD Thesis [99, 74].

For two intervals of the form $x = [\underline{x}, \overline{x}]$ and $y = [\underline{y}, \overline{y}]$, such that $\underline{x} \leq \overline{x} \wedge \underline{y} \leq \overline{y} : x, y \in \mathbb{R}$, the basic operations are:

$$x + y = [\underline{x} + \underline{y}, \overline{x} + \overline{y}] \tag{2.4}$$

$$x - y = [\underline{x} - \overline{y}, \overline{x} - \underline{y}] \tag{2.5}$$

$$x \times y = [min(\underline{x}\underline{y}, \underline{x}\overline{y}, \overline{x}\underline{y}, \overline{x}\overline{y}), max(\underline{x}\underline{y}, \underline{x}\overline{y}, \overline{x}\underline{y}, \overline{x}\overline{y})] \tag{2.6}$$

$$\frac{1}{x} = [1/\overline{x}, 1/\underline{x}] \quad (\text{if } \underline{x} > 0 \text{ or } \overline{x} < 0) \tag{2.7}$$

$$x \div y = x \times \frac{1}{y} \tag{2.8}$$

$$kx = [k\underline{x}, k\overline{x}] \quad (k \geq 0 \text{ is a scalar}) \tag{2.9}$$

Powers are also possible in IA:

$$\begin{aligned} x^n &= [1, 1] && (\text{if } n = 0) \\ &= [\underline{x}^n, \overline{x}^n] && (\text{if } \underline{x} \geq 0 \vee \underline{x} \leq 0 \leq \overline{x} \wedge n \text{ is odd}) \\ &= [\overline{x}^n, \underline{x}^n] && (\text{if } \overline{x} \leq 0) \\ &= [0, max\,(\underline{x}^n, \overline{x}^n)] && (\text{if } \underline{x} \geq 0 \vee \underline{x} \leq 0 \leq \overline{x} \wedge n \text{ is even}) \end{aligned} \tag{2.10}$$

Interval Function

An extension of the preceding concepts stands for functions. An interval function is an interval-valued function of one or more interval arguments. The connection between this type of function and the crisp ones is given by the *interval extension* concept [141]. An interval function $F(\cdot)$ with argument $\mathcal{X}_1, \mathcal{X}_2, \ldots, \mathcal{X}_n$ is an extension of a real valued function $f(\cdot)$ with argument $x_1, x_2, \ldots, x_n$ if $F(x_1, x_2, \ldots, x_n) = f(x_1, x_2, \ldots, x_n)$. In general the *inclusion property* establishes that $f(x_1, x_2, \ldots, x_n) \subseteq F(\mathcal{X}_1, \mathcal{X}_2, \ldots, \mathcal{X}_n)$ for all element $x_i \in \mathcal{X}_i,\ i \in \{1, \ldots, n\}$. From here, it follows that the range of $F(\cdot)$ can be calculated using interval arithmetic [151].

Further developments in interval arithmetic and interval functions exist, like derivative and integral operations, for more details see [141, 70].

2.2.2 Fuzzy Logic

Fuzzy logic was first advocated by L. Zadeh in 1965 [226]. Since then, it has been thoroughly studied through a large amount of papers, books, seminars, congresses and scientific journals especially dedicated to this realm.

The main claim of fuzzy logic is that bivalent logic is unable to represent many situations that involve uncertainty in a suitable way. Instead, fuzzy logic proposes smooth transitions from acceptance to denial upholding the existence of graded membership. This statement is particular appealing to dealing with natural language uncertainty as such induced by linguistic qualifiers like '*cold*' or '*big*' in '*it's cold outside*' or '*I want a big car*'. Fig. 2.7 exemplifies 2 membership functions to qualify 'cold' and 'very cold'. Notice that the same temperature has different membership values according with the meaning given to these

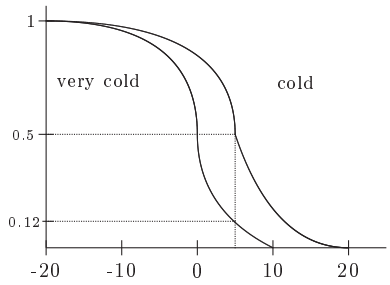


Figure 2.7: Fuzzy sets representing 'cold' and 'very cold' (in Celsius).

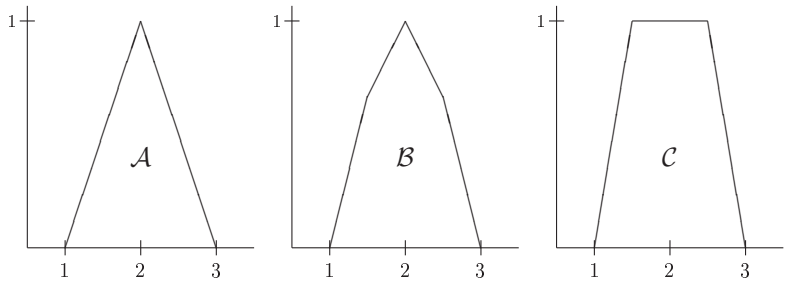


Figure 2.8: Example of membership functions of three fuzzy sets 'around 2': sets $\mathcal{A}$ (left), $\mathcal{B}$ (centre) and $\mathcal{C}$ (right).

qualifiers. In other words, 5 ° C is *more or less* 'cold' but it is definitely far from being 'very cold'.

It is of particular relevance to this work the notions of fuzzy numbers, fuzzy intervals and fuzzy arithmetic, since they are the base of input-to-output uncertainty propagation. For a comprehensive insight into fuzzy logic the readers are referred to [105].

Static components

Fuzzy logic is based on the assertion that the membership of an element x of a set $\mathcal{A}$ can be graded by ranging it from 0 (perfect exclusion) to 1 (perfect inclusion). This is represented by means of a membership function $\mu : \mathcal{S} \rightarrow [0, 1]$ that has the following canonical form:

$$\mu_{\mathcal{A}}(x) = \begin{cases} f_{\mathcal{A}}(x) & : \quad \forall x \in [a, b) \\ 1 & : \quad \forall x \in [b, c] \\ g_{\mathcal{A}}(x) & : \quad \forall x \in (c, d] \\ 0 & : \quad \text{otherwise} \end{cases} \tag{2.11}$$

where $a \leq b \leq c \leq d \in \mathcal{A}$ and $f_{\mathcal{A}}(x) : [a, b) \rightarrow [0, 1]$ and $g_{\mathcal{A}}(x) : (c, d] \rightarrow [0, 1]$ are real-valued increasing and decreasing functions respectively [7]. Fig. 2.8 shows an example of three possible membership functions 'around 2'.

Each fuzzy set can be characterized in terms of the distribution of its elements. This is done by means of subsets of $\mathcal{A}$ called α-cuts, which are defined

as $\mathcal{A}_\alpha = \{x \in \mathcal{A} | \mu_\mathcal{A}(x) \geq \alpha\}$. Notice that every α-cut is a crisp set. Two noticeable α-cuts are the *core set* or $\mathcal{A}_1$ and the *support set* or $\mathcal{A}_0$. In eq. 2.11, the core and support sets are $\mathcal{A}_1 = \{x | x \in [b, c]\}$ and $\mathcal{A}_0 = \{x | x \in [a, d]\}$ respectively.

The cardinality of a fuzzy set is defined in several ways: the *scalar cardinality* of a (discrete) fuzzy set is defined as

$$|\mathcal{A}| = \sum_{x \in \mathcal{A}} \mu_\mathcal{A}(x) \tag{2.12}$$

or in a relative way, with respect to an universal set $\mathcal{S}$ as

$$||\mathcal{A}|| = \frac{|\mathcal{A}|}{|\mathcal{S}|} \tag{2.13}$$

on the other hand, the *fuzzy cardinality* $C_\mathcal{A}(|\mathcal{A}_\alpha|)$ is a set of pairs $|\mathcal{A}_\alpha|$ (cardinality of the α-cut) and α, e.g. if $C_\mathcal{A}(|\mathcal{A}_\alpha|) = \{1/1, 5/0.5, 10/0\}$, it means that $|\mathcal{A}_1| = 1$, $|\mathcal{A}_{0.5}| = 5$ and $|\mathcal{A}_0| = 10$.

Let us consider now the concepts of fuzzy number, and fuzzy interval with the examples shown in fig. 2.8. In a crisp world, the number 2 can be represented exactly or approximately, as 2 or e.g. [1.5,2.5] respectively. When these numbers become fuzzy, their bounds blur yielding set $\mathcal{A}$ (a fuzzy number) and set $\mathcal{C}$ (a fuzzy interval).

Fuzzy numbers and intervals are combined arithmetically in terms of their α-cuts and interval-arithmetic rules, according to the following expressions [101]:

$$\mathcal{A}_\alpha + \mathcal{B}_\alpha = [\underline{a}, \overline{a}]_\alpha + [\underline{b}, \overline{b}]_\alpha = [\underline{a} + \underline{b}, \overline{a} + \overline{b}]_\alpha \tag{2.14}$$

$$\mathcal{A}_\alpha - \mathcal{B}_\alpha = [\underline{a}, \overline{a}]_\alpha - [\underline{b}, \overline{b}]_\alpha = [\underline{a} - \overline{b}, \overline{a} - \underline{b}]_\alpha \tag{2.15}$$

$$\mathcal{A}_\alpha \times \mathcal{B}_\alpha = [\underline{a}, \overline{a}]_\alpha \times [\underline{b}, \overline{b}]_\alpha = [l(\mathcal{A}_\alpha, \mathcal{B}_\alpha), u(\mathcal{A}_\alpha, \mathcal{B}_\alpha)]_\alpha \tag{2.16}$$

$$\mathcal{A}_\alpha \div \mathcal{B}_\alpha = [\underline{a}, \overline{a}]_\alpha \div [\underline{b}, \overline{b}]_\alpha = [\underline{a}, \overline{a}]_\alpha \times [1/\underline{b}, 1/\overline{b}]_\alpha \tag{2.17}$$

where

$$\begin{aligned} l(\mathcal{A}_\alpha, \mathcal{B}_\alpha) &= min(\underline{ab}, \underline{a}\overline{b}, \overline{a}\underline{b}, \overline{a}\overline{b}) \\ u(\mathcal{A}_\alpha, \mathcal{B}_\alpha) &= max(\underline{ab}, \underline{a}\overline{b}, \overline{a}\underline{b}, \overline{a}\overline{b}) \end{aligned}$$

and $\mathcal{A}_\alpha \div \mathcal{B}_\alpha$ holds if $0 \notin [\underline{b}, \overline{b}]_\alpha$. Some restrictions apply when operating one number by itself (see [101] for further details). These relations make possible the practical implementation [98][40] of *extension principle* (see [63] and references therein) and thus allow to propagate uncertainty.

Dynamic components

In fuzzy logic, the update of assignments is not carried out in the traditional way established by Bayes' theorem (eq. 2.25) due to the logic and semantic differences between fuzzy sets and probabilities. Nevertheless, some rules are necessary in order to generate new sets given the existence ones.

In fuzzy sets the basic operations of union $\mathcal{A} \cup \mathcal{B}$, intersection $\mathcal{A} \cap \mathcal{B}$ and complement $\bar{\mathcal{A}}$ are not unique as their counterparts in crisp sets (table 2.3). Instead, they are defined by classes of operations which generalization leads

to *t-conorms* and *t-norms* (see [223, 102, 105, 7] for details). The simpler realizations of such classes are:

$$\begin{aligned}
\mu_{\overline{\mathcal{A}}}(x) &= 1 - \mu_{\mathcal{A}}(x) \\
\mu_{\mathcal{A}\cup\mathcal{B}}(x) &= max(\mu_{\mathcal{A}}(x), \mu_{\mathcal{B}}(x)) \\
\mu_{\mathcal{A}\cap\mathcal{B}}(x) &= min(\mu_{\mathcal{A}}(x), \mu_{\mathcal{B}}(x))
\end{aligned}$$

2.2.3 Possibility Theory

The notion of a smooth transition that underpins fuzzy logic can be used beyond the context of quantifying the membership of an element to a set to qualify and moreover to quantify the possibility of an element regarding a set by means of a *possibility measure* or *possibility distribution*.

Static components

The simpler example of a possibility measure is the characteristic function of a set. Given a set $\mathcal{E}$ and its characteristic function $\mu_{\mathcal{E}}(x)$, the assertion $y \in \mathcal{E}$ has possibility $\mu_{\mathcal{E}}(y)$, thus a possibility equal to zero means that $y \notin \mathcal{E}$ while a possibility equal to one entails $y \in \mathcal{E}$. This binary-valued view of possibility is contemplated in the *Classical Possibility Theory*. For instance, if x is a unique value such that $x \in \mathcal{E}$, then [41]:

$$\Pi_{\mathcal{E}}(\mathcal{A}) = \begin{cases} 1 & \text{if} \quad \mathcal{A} \cap \mathcal{E} \neq \emptyset \\ 0 & \text{otherwise} \end{cases} \tag{2.18}$$

which means that $\mathcal{A}$ has the possibility of containing x if $\Pi_{\mathcal{E}}(\mathcal{A}) = 1$.

On the other hand, if a fuzzy instead of a binary logic is adopted, the characteristic function becomes a membership function, opening the door to a *Graded Possibility Theory*. Following the previous example, the possibility of $\mathcal{A}$ is assessed in terms of the supremum degree of membership of the intersection between $\mathcal{A}$ and the core set of $\mathcal{X}$:

$$\Pi_{\mathcal{E}}(\mathcal{A}) = \sup_{x \in \mathcal{A}} \mu_{\mathcal{E}}(x) \tag{2.19}$$

A second functional called *necessity* is defined in possibility theory as a dual measure built with reference to the complementary set $\bar{\mathcal{A}}$ of that under study. Thereby, the necessity $N_{\mathcal{E}}(\mathcal{A})$ is:

$$N_{\mathcal{E}}(\mathcal{A}) = 1 - \Pi_{\mathcal{E}}(\bar{\mathcal{A}}) = \begin{cases} 1 & \text{if} \quad \mathcal{E} \subseteq \mathcal{A} \\ 0 & \text{otherwise} \end{cases} \tag{2.20}$$

Two axioms of 'maxitivity' and 'minitivity' [41] complete the core of definitions of possibility theory:

$$\Pi_{\mathcal{E}}(\mathcal{A} \cup \mathcal{B}) = \max\{\Pi_{\mathcal{E}}(\mathcal{A}), \Pi_{\mathcal{E}}(\mathcal{B})\} \tag{2.21}$$

$$N_{\mathcal{E}}(\mathcal{A} \cap \mathcal{B}) = \min\{N_{\mathcal{E}}(\mathcal{A}), N_{\mathcal{E}}(\mathcal{B})\} \tag{2.22}$$

Clearly the possibility measure resemble probability as both quantify the evidence about the occurrence of an event, however possibility allows weaker

assertions, since one can state that something is possible without knowing nothing about its likelihood. In fact, to say that something is 'possible' is open to several interpretations that can find echo in some version of probability theory [41] (like those surveyed later on in this chapter)[2]. This fact makes possible the use of possibility theory to handle aleatory and epistemic uncertainty as well.

Dynamic components

The updating of information is possible via conditional possibility, an index that expresses the possibility of one set given a previous knowledge about the actual possibility of another set ($\Pi(\mathcal{B}) \neq 0$), thus:

$$\Pi(\mathcal{A}|\mathcal{B}) = \frac{\Pi(\mathcal{A} \cap \mathcal{B})}{\Pi(\mathcal{B})} \tag{2.23}$$

The rule for updating the possibility is then given by:

$$\Pi(\mathcal{H}|\mathcal{E}) = \min\left\{1, \frac{\Pi(\mathcal{E}|\mathcal{H}) \cdot \Pi(\mathcal{H})}{\Pi(\mathcal{E}|\bar{\mathcal{H}}) \cdot \Pi(\bar{\mathcal{H}})}\right\} \tag{2.24}$$

More details about the these dynamic components can be found in [41] and references therein.

2.2.4 Probability Theory

This section brings the basic of probability theory. Most of the contents follow [142, 7, 165].

Static components

Probability is an index in $[0, 1]$ that denotes our degree of belief in the likelihood of an event. Such a belief can rely on some prior knowledge about its observed frequency of occurrence (*objective probability*) or might express the inner state of confidence of a subject with respect to its likelihood (*subjective probability*). In other words, probability is a projection of what we think may happen regarding the different sources that feed our knowledge, i.e., the observation of the past and our interpretation of it.

The set of all possible outcomes or events is the sampling universe. Thus, the probability of an event not only expresses its likelihood in an isolated way, but also a characteristic of the whole population. In probability theory, all the assignments of probability must conform to the Kolmogorov axioms:

Axiom 1: The probability $p(A)$ of an event A is a non negative number such that $0 \leq P(A) \leq 1$.

Axiom 2: The probability of the sampling universe S is the unit, i.e. $P(S) = 1$.

Axiom 3: For any sequence of mutually exclusive or disjoint events $A_1, A_2, \ldots$ it holds that $P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i), \quad n = 1, 2, \ldots, \infty$.

[2] About the possible bridges between the two theories see e.g. [43, 41] and references therein.

The connections amongst events is expressed in terms of *dependence* and *independence*. The former holds when the occurrence of one event influence the likelihood of the other, while the latter indicates no influence between each other. The properties that characterize the possible connections of the elements of the power set P_S of the sampling universe S are:

Property 1: The probability of the union verifies $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

Property 2: If A and B are mutually exclusive or disjoint events property 1 yields $P(A \cup B) = P(A) + P(B)$. [3]

Property 3: The connection between complementary events is given by $P(\bar{A}) = 1 - P(A)$.

Property 4: The conditional probability of A given the prior occurrence of B is given by $P(A|B) = \frac{P(A \cap B)}{P(B)}$.

Property 5: If two events are disjoint $P(A \cap B) = P(A) \cdot P(B)$ holds.

Dynamic components

The aforementioned axioms and properties allow to construct a coherent framework for probabilistic reasoning in static situations, i.e. when the amount of evidence is complete. Nevertheless, when a new piece of evidence comes along, it is necessary to update the assignments of probability according with the new information. In probability theory this is done using Bayes' theorem:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \tag{2.25}$$

or more generally,

$$P(A_i|B) = \frac{P(B|A_i)P(A_i)}{\sum_j P(B|A_j)P(A_j)} \tag{2.26}$$

This rule establishes that the probability of A given that B happened must be updated proportionally to the prior value of A and the standardized likelihood of B, or in Bayesian jargon:

$$\text{posterior} = \frac{\text{likelihood} \times \text{prior}}{\text{normalizing constant}}$$

Probability distributions

The data about a phenomenon can be de used to predict its future outcomes, likewise the partial information about a population can be used to infer current features of it or to predict behaviours to come. The fundamental abstraction to do this is the notion of *random variable*, i.e. a variable associated to some feature of interest (e.g. the height or age of children in a classroom, the number of customer that arrive in a period of time, the number of accidents in industrial facilities) that takes random values according to some law.

[3] This property is called *additivity* and it is characteristic of classical probability theory.

Random variables and distribution functions are used to express the probability that a phenomenon hits a particular state. The relation between these two elements can be expressed by mean of a *cumulative distribution function* (CFD) of the type $\mathtt{F}(x) \equiv P(\mathtt{X} \leq x)$ which gives the probability of the random variable $\mathtt{X}$ being less than or equal x, or by its derivative. For *discrete random variables* the derivative is called *probability mass function* (PMF) and gives the probability of a $\mathtt{X}$ being equal x: $\mathtt{P}(x) = P(\mathtt{X} = x)$. Likewise, for *continuous random variables* the *probability density function* (PDF) represents the density of probability of x, which is zero for each single value, and has a non-null value for any range of values given that $\mathtt{f}(x)\Delta x$ holds due to PDF being a derivative.

Characterizing sets statistically

In general, any set can be seen as the feasible domain of a random variable. In consequence, statistical analysis can be applied to extract relevant information about aleatory and non-aleatory sets. As to the former, the underlaying law that governs the aleatory phenomenon is approximated using distribution functions, and both the set and the distribution are characterized by *moments* and other relevant parameters associated with the latter. In contrast, to non-aleatory sets the assignment of a distribution is simply an artificial recourse. This difference has repercussions in the way of analyzing and propagating both kind of sets (cf. [47]). Nevertheless, distributions apart, moments are commonly used to characterize non-aleatory sets.

The first group of parameters are those that characterize central tendencies by identifying some axis of symmetry:

Mean or expected value This is the first central moment, defined as

$$\mu = \int_{\mathtt{X}} x\mathtt{f}(x)dx \tag{2.27}$$

$$\mu = \sum_{i=1}^{n} x_i \mathtt{P}(x_i) \tag{2.28}$$

for PDF and PMF respectively. This moment is estimated with m samples of x doing

$$\bar{x} = \frac{1}{m}\sum_{i=1}^{m} x_i \tag{2.29}$$

Median The median x_m of m samples is the value that divides the data into two groups of the same cardinality, one of which has its elements below x_m and the other above x_m.

The second group characterizes variability:

Variance This is the second central moment, defined as

$$\sigma^2 = \int_{\mathtt{X}} (x-\mu)^2 \mathtt{f}(x)dx \tag{2.30}$$

$$\sigma^2 = \sum_{\mathtt{X}} (x_i-\mu)^2 \mathtt{P}(x_i) \tag{2.31}$$

for PDF and PMF respectively. In occasions variability is also expressed in terms of the *standard deviation* σ, which is the square root of the variance. This moment is estimated with m samples of x as

$$S^2 = \frac{1}{m-1}\sum_{i=1}^{m}(x_i - \bar{x})^2 \tag{2.32}$$

Coefficient of variation (COV) This measure defined as $\text{COV} = \frac{S}{\bar{x}}$ is a dimensionless quantity that can be interpreted as how much is the amount of the actual deviation with respect to $\bar{x}$ in percentage terms. Therefore, in terms of uncertainty reduction, the lesser the amount the better.

Table 2.4 presents the CDF, and PDF as well as the 1st and 2nd moments of some common distribution of probability used in engineering.

To close this section, consider also another kind of parameters commonly used parameters to characterize distributions, named fractiles or quantiles. A x_p fractile of a sample is the value such that $p\%$ of the data is lesser or equal x_p. In terms of probability: $P(\mathtt{X} \leq x_p) = p\%/100$.

2.2.5 Dempster-Shafer Evidence Theory

Dempster-Shafer Evidence Theory (DST) was introduced by Shafer [186] based on the mathematics developed by Dempster [39]. It is considered a generalization of classical probability theory to a multi-valued space, i.e. the mass resides on sets instead of precise points, as in PMF; concomitantly an assignment can be made to a group of events (called *focal points*) without discriminating the precise probability of the elements within it. This characteristic distinguishes DST from classical probability theory, making the former a suitable alternative to overcome the limitations encountered in the latter when modelling epistemic uncertainty.

Static components

The three fundamental elements in DST are:

The *Basic Probability Assignment*: Denoted as bpa or m, is the amount of evidence that support the claim that a particular element of the universal set $\mathcal{S}$ belongs to a subset of the powerset $P_{\mathcal{S}}$ [183, pg. 13]. Mathematically we say that bpa are mappings $m : P_{\mathcal{S}} \rightarrow [0, 1]$ such that

$$m(\emptyset) = 0 \tag{2.33}$$

$$\sum_{\mathcal{A}\in P_{\mathcal{S}}} m(\mathcal{A}) = 1 \tag{2.34}$$

Likewise, bpa can be assessed from belief measures via Möbius transformation:

$$m(A) = \sum_{\mathcal{B}|\mathcal{B}\subseteq\mathcal{A}} (-1)^{|\mathcal{A}-\mathcal{B}|}\text{Bel}(\mathcal{B}) \tag{2.35}$$

Normal distribution

Functional representation

PDF: $\mathtt{f}(x) = \frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \quad -\infty \leq x \leq \infty$

CDF: There is no analytical form. Tables and numerical methods exists.

Moments

1st *moment:* μ

2nd *moment:* σ^2

Lognormal distribution

Functional representation

PDF: $\mathtt{f}(x) = \frac{1}{\phi x\sqrt{2\pi}}\exp\left(-\frac{(\ln(x)-\xi)^2}{2\phi^2}\right) \quad -\infty \leq x \leq \infty$

where ξ and ϕ are parameters of the lognormal distribution.

CDF: Computed from normal distribution of $\ln(x)$.

Moments

1st *moment:* $\mu = exp(\xi + \phi^2/2)$

2nd *moment:* $\sigma^2 = exp(\phi^2)\left(exp(\phi^2) - 1\right) exp(2\xi)$

Exponential distribution

Functional representation

PDF: $\mathtt{f}(x) = \lambda\exp(-\lambda x) \quad 0 \leq x \leq \infty$

where λ is a parameter of the exponential distribution.

CDF: $\mathtt{F}(x) = 1 - \exp(-\lambda x)$

Moments

1st *moment:* $\mu = 1/\lambda$

2nd *moment:* $\sigma^2 = 1/\lambda^2$

Weibull distribution (two parameters)

Functional representation

PDF: $\mathtt{f}(x) = (k/c)(x/c)^{k-1}\exp\left(-(x/c)^k\right) \quad 0 \leq x \leq \infty$

where c and k are parameters of the Weibull distribution.

CDF: $\mathtt{F}(x) = 1 - \exp\left(-(x/c)^k\right)$

Moments

1st *moment:* $\mu = c\Gamma\left(1 + \frac{1}{k}\right)$

2nd *moment:* $\sigma^2 = c^2\left(\Gamma\left(1 + \frac{2}{k}\right) + \Gamma^2\left(1 + \frac{1}{k}\right)\right)$

Uniform distribution

Functional representation

PDF: $\mathtt{f}(x) = \frac{1}{b-a} \quad a \leq x \leq b$

where a and b are parameters of the Uniform distribution.

CDF: $\mathtt{F}(x) = \frac{x-a}{b-a}$

Moments

1st *moment:* $\mu = \frac{a+b}{2}$

2nd *moment:* $\sigma^2 = \frac{(b-a)^2}{12}$

Table 2.4: Common Probability Distribution Functions for continuous random variables employed in engineering (adapted from [142]).

The *Belief measure*: Also called *support* and denoted as Bel, is the sum of all the masses of the subsets of the set of interest:

$$\mathrm{Bel}(\mathcal{A}) = \sum_{\mathcal{B}|\mathcal{B}\subseteq\mathcal{A}} m(\mathcal{B}) \tag{2.36}$$

Belief measures verify

$$\mathrm{Bel}(\emptyset) = 0 \tag{2.37}$$

$$\mathrm{Bel}(\mathcal{S}) = 1 \tag{2.38}$$

Finally, belief measures are supperadditive, i.e. $\mathrm{Bel}(\mathcal{A}\cup\mathcal{B}) \geq \mathrm{Bel}(\mathcal{A}) + \mathrm{Bel}(\mathcal{B})$ for $\mathcal{A}\cap\mathcal{B} = \emptyset$.

The *Plausibility measure* Denoted as Pls, is the sum of the masses of all sets intersecting the set of interest:

$$\mathrm{Pls}(\mathcal{A}) = \sum_{\mathcal{B}|\mathcal{B}\cap\mathcal{A}\neq\emptyset} m(\mathcal{B}) \tag{2.39}$$

Plausibility measures also verify

$$\mathrm{Pls}(\emptyset) = 0 \tag{2.40}$$

$$\mathrm{Pls}(\mathcal{S}) = 1 \tag{2.41}$$

$$\mathrm{Pls}(\mathcal{A}\cup\mathcal{B}) \geq \mathrm{Pls}(\mathcal{A}) + \mathrm{Pls}(\mathcal{B}) \tag{2.42}$$

The relations between Pls and Bel are given by:

$$\mathrm{Pls}(\mathcal{A}) \geq \mathrm{Bel}(\mathcal{B}) \tag{2.43}$$

$$\mathrm{Pls}(\mathcal{A}) = 1 - \mathrm{Bel}(\overline{\mathcal{A}}) \tag{2.44}$$

$$\mathrm{Pls}(\overline{\mathcal{A}}) = 1 - \mathrm{Bel}(\mathcal{A}) \tag{2.45}$$

$$\mathrm{Bel}(\mathcal{A}) = 1 - \mathrm{Pls}(\overline{\mathcal{A}}) \tag{2.46}$$

$$\mathrm{Bel}(\overline{\mathcal{A}}) = 1 - \mathrm{Pls}(\mathcal{A}) \tag{2.47}$$

Fig. 2.9 presents a typical example of how the basic probability assignments are made. In this case the analyst has identified seven subsets within $\mathcal{S}$. Notice that the likelihood of an element x belonging to set $\mathcal{A}$ is bounded within $[\mathrm{Bel}(A), \mathrm{Pls}(A)]$, i.e. the minimal amount of evidence that support the assertion $x \in \mathcal{A}$ which is $\mathrm{Bel}(A) = \sum_{i=\{2,5\}} m(\mathcal{B}_i)$ and the maximal amount of evidence that is consonant with such an assertion is $\mathrm{Pls}(A) = \sum_{i=\{1,2,3,4,5\}} m(\mathcal{B}_i)$.

Dynamic components

In DST the updating of the assignments can be made in different forms. Dempster's rule of combination was the first proposal to combine evidence. This rule aggregate the evidence of different sources (expert 1 and 2) as

$$m_{1,2}(\mathcal{A}) = \frac{\sum\limits_{\mathcal{B}\cap\mathcal{C}=\mathcal{A}\neq\emptyset} m_1(\mathcal{B})m_2(\mathcal{C})}{1 - \sum\limits_{\mathcal{B}\cap\mathcal{C}=\emptyset} m_1(\mathcal{B})m_2(\mathcal{C})} \tag{2.48}$$

but since it has received some criticisms, authors have formulated alternatives rules. Amongst others, one can mention Yager's rule, Inagaki's rule and Mixed Averaging Rule. See [183] for a comprehensive review of combination rules in DST.

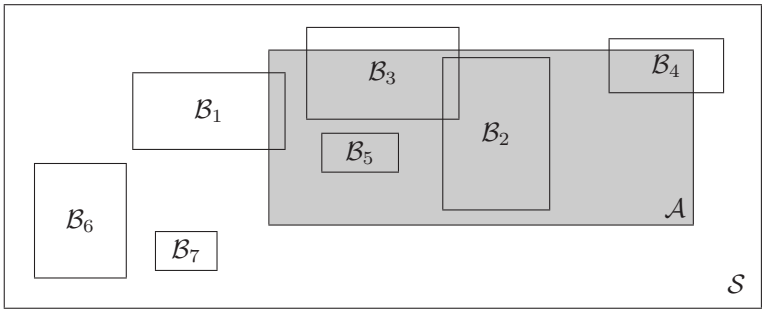


Figure 2.9: Example of bpa, Bel and Pls in Dempster-Shafer Theory.

2.2.6 Imprecise Probabilities

Imprecise probability theory was developed by P. Walley [207] based on earlier ideas of Keynes, Smith and Good [211, pg. 462]. It is nowadays a well developed theoretical framework with an important research community, a biannual conference and its own site on the internet (`http://www.sipta.org`).

Static components

In classical probability theory, the probability of an event x is mapped from the universe of concern to a precise value in [0,1] by means of a probability law. In imprecise probability theory, such a mapping is treated as impossible in practice. Instead, a class $\mathfrak{M}$ of plausible probability distributions to map event x is considered, thus the *lower probability* $\underline{P}(x)$ and the *upper probability* $\overline{P}(x)$ are defined as the the infimum and supremum of probabilities $P(x)$ over all probability distribution P belonging $\mathfrak{M}$, respectively [211, pg. 462], i.e.:

$$\underline{P}(x) = \inf\{P(x)|P \in \mathfrak{M}\} \tag{2.49}$$

$$\overline{P}(x) = \sup\{P(x)|P \in \mathfrak{M}\} \tag{2.50}$$

Upper and lower probabilities are seen as a lower and upper envelopes of probability. These measures satisfy the following properties [210]:

Property 1: The lower probability is a non-negative number in [0,1] such that $\underline{P}(\emptyset) = 0$ and $\underline{P}(\mathcal{S}) = 1$.

Property 2: Given the lower probability, the upper probability is defined by $\overline{P}(A) = 1 - \underline{P}(\bar{A})$.

Property 3: For disjoint sets the super-additivity of lower probabilities holds, i.e. $\underline{P}(A \cup B) \geq \underline{P}(A) + \underline{P}(B)$.

Property 4: Likewise, disjoint sets verify the sub-additivity of upper probabilities, i.e. $\overline{P}(A \cup B) \leq \overline{P}(A) + \overline{P}(B)$.

Property 5: $\overline{P}(A) \geq \underline{P}(B)$.

Property 6: $A \supset B \Longrightarrow \underline{P}(A) \geq \underline{P}(B) \wedge \overline{P}(A) \geq \overline{P}(B)$.

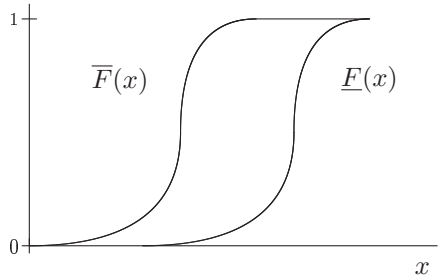


Figure 2.10: An example of a p-box: the CDF is circumscribed by $\overline{F}(x)$ and $\underline{F}(x)$.

Property 7: $A \cap B = \emptyset \Longrightarrow \overline{P}(A \cup B) \geq \overline{P}(A) + \underline{P}(B) \geq \underline{P}(A \cup B)$.

Property 8: $\underline{P}(A) + \underline{P}(B) \leq \overline{P}(A \cup B) + \underline{P}(A \cap B) \leq 1 + \underline{P}(A \cap B))$.

Dynamic components

Imprecise probability is generalized in upper and lower previsions that provide the procedures to update information (*natural extension*) and to make decisions based on the concept of gambles. The interested reader is referred to [209, 208] to see the details of such procedures.

2.2.7 Probability Boxes

Static components

A probability box (p-box) is a type of imprecise probability structure for distribution functions. It is composed of two binding non-decreasing functions $\overline{F} : \mathbb{R} \to [0, 1]$ and $\underline{F} : \mathbb{R} \to [0, 1]$, such that the p-box $[\overline{F}, \underline{F}]$ denotes or contain every unknown CDF $F(x)$ for the random variable $\mathtt{X}$ (see fig. 2.10). The connections between p-boxes, imprecise probability and Dempster-Shafer are given by [50]:

with Imprecise probability theory

$$\overline{F}(x) = 1 - \underline{P}(\mathtt{X} > x) \tag{2.51}$$

$$\underline{F}(x) = \underline{P}(\mathtt{X} \leq x) \tag{2.52}$$

with Dempster-Shafer theory For subsets expressed as an interval plus its mass $([\underline{x}_i, \overline{x}_i], m_i)$, the p-box can be determined as

$$\overline{F}(x) = \sum_{\underline{x}_i \leq x} m_i \tag{2.53}$$

$$\underline{F}(x) = \sum_{\overline{x}_i < x} m_i \tag{2.54}$$

p-boxes allow to express different levels of uncertainty; for instance if the shape of the distribution is known, say a normal function, but its parameters are

uncertain, then the p-box adopt the form $N([\underline{\mu}, \overline{\mu}], [\underline{\sigma^2}, \overline{\sigma^2}])$. Thus every normal distribution having its mean in $[\underline{\mu}, \overline{\mu}]$ and its variance in $[\underline{\sigma^2}, \overline{\sigma^2}]$ conforms the definition of the p-box.

Dynamic components

Due to the connection between DS theory and p-boxes, the combination of the evidence is possible via DS' rules of combination. The reader is referred to [183] where a comprehensive study of the pros and cons of each rule is presented.

2.2.8 Info-gap Models

An info-gap model is a non-probabilistic methodology proposed by Y. Ben-Haim that aims at quantifying the disparity between what the DM knows and what could be known [13]. An info-gap model defines a family of nested sets $\mathcal{U}(\alpha, \tilde{\Phi})$, $\alpha \geq 0$ of vectors or functions $\tilde{\Phi}(\mathbf{x})$ that approximate or estimate the true value $\Phi(\mathbf{x})$ of the quantity of concern. Thus, as the *uncertainty parameter* α grows the sets become more inclusive [14]: $\alpha \leq \alpha' \Rightarrow \mathcal{U}(\alpha, \tilde{\Phi}) \subseteq \mathcal{U}(\alpha', \tilde{\Phi})$.

From the preceding model, Ben-Haim prescribes two functions that measure the *robustness* $\hat{\alpha}(x, r_c)$ and the *opportunity* $\hat{\beta}(x, r_c)$ given the current level of uncertainty:

$$\hat{\alpha}(x, r_c) = \max\left\{\alpha : \left(\min_{\tilde{\Phi} \in \mathcal{U}(\alpha, \tilde{\Phi})} R(\mathbf{x}, \Phi)\right) \geq r_c\right\} \tag{2.55}$$

$$\hat{\beta}(x, r_w) = \min\left\{\alpha : \left(\max_{\tilde{\Phi} \in \mathcal{U}(\alpha, \tilde{\Phi})} R(\mathbf{x}, \Phi)\right) \geq r_w\right\} \tag{2.56}$$

The robustness of the choice $\mathbf{x}$ reflects the higher horizon of uncertainty α at which reward R no less than r_c is guaranteed. In other words, $\hat{\alpha}(x, r_c)$ points out the higher discrepancy between the real value $\Phi(\mathbf{x})$ and the estimate $\tilde{\Phi}(\mathbf{x})$ that can be allowed without incurring in a loss of quality below r_c. Likewise, $\hat{\beta}(x, r_w)$ indicates the least level of knowledge deficiency at which the performance of reward of the system can reach the windfall r_w [13]. Some examples of this methodology can be found in [14, 13].

2.3 Generalized Theoretical Frameworks

The preceding subsections took a glance at some of the most relevant theories of uncertainty. Further efforts strive to establish the connection between such theories offering a unified vision of the different aspect related to information and its absence. In this context L. Zadeh [227] proposed the Generalized Theory of Uncertainty, P. Walley [209] its Unified Theory of Imprecise Probabilities and G. Klir [104] the Generalized Information Theory. The extension and complexity of these frameworks, as well as the scope of this thesis advise against going beyond a simple mention. However, the reader interested in a deep insight into these matters is encouraged to study the cited references.

2.4 Summary of the chapter

In this chapter we have presented the basics of uncertainty handling, from an opening discussion to the mathematics and theoretical foundations that underpin the current techniques of uncertainty handling.

Amongst the different facets of uncertainty, five key questions have been proposed to help any decision-making analysts and/or algorithms designer to tackle an uncertainty analysis:

1. *What? - Information:* what is known? what is ignored? what is worthwhile to be investigated?

2. *Why? - Purpose:* why the presence of uncertainty constitutes a problem in decision-making? Does it hamper the identification of optimal solutions now or will it do it in the future?

3. *Obtainable? - Nature:* what are the characteristic of such uncertainties? Is it possible to reduce them with more information (epistemic uncertainty) or not (aleatory uncertainty).

4. *Reducible? - Price:* Should the uncertainty be reducible, is it worthwhile to reduce it? What is the cost and the payoff of additional information?

5. *Where? - Sources and Influence:* where do the uncertainties come from (measures, environment, lack of records, etc.)? What do they impact (models, forecasts, comparisons, etc.)?

Afterwards, the main concepts of set theory and interval arithmetic have been presented. Finally, the most important theoretical frameworks to treat uncertainty used nowadays, namely probability theory, fuzzy logic, possibility theory, Dempster-Shafer evidence theory, imprecise probabilities, probability boxes and info-gap models, have been surveyed.

In subsequent chapters these concepts will serve as a base to perform in-depth analyses of multiple-criteria decision-making problems based on evolutionary optimization techniques in the presence of uncertainty. To do so, the next chapter shall present the basics of multiple-criteria decision-making and evolutionary optimization, with a survey of the state of the art in uncertainty handling in this realm.

Chapter 3

Evolutionary Optimization-based Multi-Criterion Decision Making

Evolutionary Algorithms (EA) have shown outstanding capabilities to solve a wide range of optimization and artificial learning problems in different domains of applied sciences. This success bred the incorporation of EA techniques into the classical world of multiple-criteria decision science, opening the way for a brand new discipline: Evolutionary Multiple-objective Optimization (EMO).

In the course of barely two decades EMO has endowed practitioners with powerful heuristics that significantly enhanced their ability to identify the set of optimal solutions in an efficient way. Moreover, such heuristics known as Multiple-Objective Evolutionary Algorithms (MOEA) are able to deal with non-continuous, non-convex and non-linear spaces, as well as with problems whose objective functions are not explicitly known (as is the case, e.g., of Monte Carlo simulations), thereby overcoming the typical limitations of classical approaches to cope with the aforementioned types of domains.

The tremendous progress in EMO is partially due to the boost in efficiency introduced by the successful integration of the advances achieved in general EA into MOEA, (viz. elitism, complexity analysis, parallelization) which makes MOEA an attractive *all-purpose* tool for practitioners committed to Multiple-Criteria Decision-Making (MCDM). Nonetheless, the significant achievements in efficiency have not been reproduced yet by the efforts in other other relevant issues like preference incorporation or uncertainty handling in EMO-based decision-making.

This chapter brings the basics of multiple-criteria decision-making and EA. Then it offers an overview of MOEA structure, representative algorithms, performance metrics and some other issues. Afterwards the state of the art in uncertainty handling involving EMO is surveyed.

3.1 Multiple-Criteria Decision-Making

3.1.1 Basics of Decision-Making

Essentially, decision-making is about making choices to satisfy needs and aspirations. Such choices have a subject, namely decision-makers, and an object, the available alternatives of decision. Rational DM aim at maximizing their level of fulfilment by selecting that alternative that produces a higher return. This principle of decision is expressed in the axiom of choice of Zeleny [230, 229].

When decision-making is assisted by formal decision-aid methods, DM's needs and aspirations -hereafter called *preferences*- are to be mathematically modelled. For that matter the analyst, i.e. the actor that helps DM during the decision-making process, must translate preferences into attributes, objectives, goals and criteria entailed in the problem.

Attributes refers to the valuable characteristics of a system that allow DM to distinguish between alternatives. Objectives point the directions of improvement on the attributes. Goals suggest minimal levels of achievement of such improvements. Finally, criteria are principles of judgment constructed from the combination of attributes, objectives and goals [163]. These elements are integrated into a mathematical program of the type:

$$\begin{array}{ll} & Opt\ F(\mathbf{x}) \\ \text{s.t.:} & \\ & G(\mathbf{x}) \leq 0 \\ & H(\mathbf{x}) = 0 \end{array} \tag{3.1}$$

where Opt indicates the direction of improvement (*maximization* or *minimization*) of the function $F(\mathbf{x})$. The feasible domain $\mathbf{X}$ is defined by $G(\mathbf{x})$ and $H(\mathbf{x})$ which are vectors of inequality and equality constraints respectively.

If the problem can be properly modelled by a unique criterion, it means that only one aspect of the reality seems pertinent to measure the level of satisfaction, to guide the search towards the best solution. According to Roy [167] such a problem or system spontaneously tends to an extreme value of its single criterion. In this scenario, $F(\mathbf{x})$ becomes a single function that entirely expresses the DM's preferences whereas optimization amounts to the main activity to solve this *single-criterion paradigm* program.

Nevertheless, the richness of possibilities and perceptions that human mind is able to conceive cannot be always reduced to nor reflected upon a single criterion. Instead, a more complex conception of reality characterized by multiple -*and conflicting*- criteria is an alternative way for modelling problems according to the *multiple-criteria paradigm* [167]. In this view different aspects of reality are considered relevant to assess the level of satisfaction that any possible alternative may provide. As a result, $F(\mathbf{x})$ becomes a vector of objective functions.

For some authors single-criterion problems are just a particular case of the multiple-criteria ones (e.g. [31]), while others claim that the former paradigm is not a collapse of the latter but a different way to view reality [167]. Perhaps a good way to account for this last reasoning is to realize that a single criterion allows a complete order of the alternatives in terms of their quality, thus the problem is reduced to single (*technological* in decision-making parlance) search [230, 229, 163] for the optimal alternative. In other words, there is no decision to make in the strict sense since the optimum is unquestionably wished. In

contrast, only a partial order is possible in multiple-criteria problems due to the inherent situation of conflict between criteria. The search of optimal solutions converges to a set of optimal alternatives (also called *efficient*, *non-dominated* or *Pareto* solutions) such that further improvements in one criterion lead to decrease quality in at least another one. The DM is compelled therefore to accept a trade-off between attributes.

Types of Problems

The existence of multiple criteria gives room for different classes of decisional problems that rely upon distinct features like the purpose and the temporal framework of the model and the characteristics of the alternatives. Regarding the first issue, authors recognize different kinds of *problematics*, being the most important ones [81, 80]:

Choice or determining the subset of best alternatives,

Sorting or partitioning the set of alternatives into subsets with respect of pre-established norms, and

Ranking or ranking the set of alternatives.

Sorting and ranking problematics are typically of problems whose alternatives are measured as discrete variables. Choice, however, may include alternatives with either discrete or continuous attributes or even both. Classical methods like Analytical Hierarchical Process and outranking methods like ELECTRE and PROMETHEE tackle problematics with discrete variables. In contrast, Multiattribute Utility Theory (MAUT) focuses on choice problems with both discrete and continuous dimensions aggregating the different criteria in one single function. Moreover, Multiple-Criteria Decision Making (MCDM) provides a framework for other classical approaches that disregard aggregation, within the same class of problematic. Such methods have had a major impact on EMO.

On the one hand, the DM may demand at once the optimality condition, forcing the search for strictly efficient alternatives. This amounts to a multiple-objective problem (see section 3.1.2) that can be solved by a complete determination of the set of efficient alternatives (*multiple-objective optimization*) or directing the search towards a reference ideal point (*compromise programming*) [228, 229]. Perfect information and *present* as temporal framework (cf. sec. 2.1.2) are the two fundamental assumptions that underpin this approach. On the other hand, the DM can formulate a scenario where partial instead of full satisfaction of preferences is the guiding principle. The underlying idea is that optimality cannot be immediately achieved under the current (unstable) state of affairs; instead a series of incremental steps (goals) are prescribed under the *goal programming* approach.

Table 3.1 summarizes the fundamental points of this section. See [44] and references therein for a comprehensive review the history of MCDM and MAUT. The main contributions in goal programming are surveyed in [3]. In [168] B. Roy describes the nexus between decision-making and decision-aid. For some classical treaties see [230, 171, 100, 163]. For a survey of MCDM methods in engineering see [129].

Actors
Decision-Maker (DM)
Entity (person or group) with the authority to make the final decision (choice) on the current problem.
Analyst
Entity that helps the DM in modelling the problem, solving the mathematical program and identifying the final alternative.
Elements
Attributes
Valuable characteristics of a system that allow DM to distinguish between alternatives.
Objectives
Directions of improvement on the attributes (*maximization* or *minimization*).
Goals
Minimal levels of achievement of attributes improvements.
Criteria
Principles of judgment constructed from the combination of attributes, objectives and goals.
Type of problem
Single-Criterion
Ordinary optimization.
Multiple-Criteria
Multiple-objective programming (search for the set of efficient alternatives).
Compromise programming (search for the closer alternative to an ideal reference point).
Goal programming (search for an incremental improvement of satisfaction).

Table 3.1: Basics of Decision-Making.

3.1.2 Multiple-Objetive Optimization

A multiple-objective optimization problem consists of solving a mathematical program of the form:

$$\begin{aligned} & Opt\left(F(\mathbf{x}) = (f_1(\mathbf{x}), f_2(\mathbf{x}), \ldots, f_k(\mathbf{x}))^t\right) \\ \text{s.t.:} \quad & G(\mathbf{x}) \leq 0 \\ & H(\mathbf{x}) = 0 \end{aligned} \tag{3.2}$$

where *Opt* stands for optimization (maximization or minimization). F is a vector of functions $f_i : \mathbb{R}^n \rightarrow \mathbb{R}$ such that $F : \mathbf{X} \rightarrow \mathbf{Y}$ maps a vector of n decision variables $\mathbf{x} = (x_1, x_2, \ldots, x_n)^t$ (called a *decision vector*, a *solution vector* or simply an *alternative*) from the *decision space* $\mathbf{X}$, defined by vectors $G(\mathbf{x})$ of g inequality and $H(\mathbf{x})$ of h equality constraints, into a k-dimensional *objective vector* $\mathbf{y} = (y_1, y_2, \ldots, y_k)^t$ in the *objective space* $\mathbf{Y} \subseteq \mathbb{R}^k, k \in \mathbb{N}$ [238].

As pointed out earlier on, the concept of optimality in single objective cannot be directly extrapolated to multiple-objective problems. Instead, a different notion of optimality that captures the multiplicity of attributes and criteria is necessary. As mentioned before, some theoretical approaches like MAUT aggregate the criteria into one single-criterion function called utility. Other approaches like most of MCDM, classify the solutions in terms of Pareto optimality.

Without a loss of generality, in terms of minimization[1]:

Definition 1 (Pareto Dominance) $\mathbf{x}_1$ *dominates* $\mathbf{x}_2$*, denoted* $\mathbf{x}_1 \succ \mathbf{x}_2$ *, iff* $f_i(\mathbf{x}_1) \leq f_i(\mathbf{x}_2) \wedge F(\mathbf{x}_1) \neq F(\mathbf{x}_2)$; $i \in \{1, 2, \ldots, k\}$*. If there are no solutions which dominates* $\mathbf{x}_1$*, then* $\mathbf{x}_1$ *is non-dominated.*

Definition 2 (Weak Pareto Dominance) $\mathbf{x}_1$ *weakly dominates* $\mathbf{x}_2$*, denoted* $\mathbf{x}_1 \succeq \mathbf{x}_2$ *, iff* $f_i(\mathbf{x}_1) \leq f_i(\mathbf{x}_2)$; $i \in \{1, 2, \ldots, k\}$*. This weak relation includes the previous one* ($\succeq \subseteq \succ$)*.*

Definition 3 (Pareto Optimal) *A solution vector* $\mathbf{x}^* \in \mathbf{X}$ *is a Pareto optimal solution iff* $\neg\exists x \in X : \mathbf{x} \succeq \mathbf{x}^*$*. These solutions are also called true Pareto solutions.*

Definition 4 (Pareto Approximation Set) *A set of non-dominated solutions* $\{\mathbf{x}^* | \neg\exists \mathbf{x} : \mathbf{x} \succ \mathbf{x}^*;\ \mathbf{x}^*, \mathbf{x} \in \mathbf{D} \subseteq \mathbf{X}\}$ *is said to be a Pareto approximation set.*

Definition 5 (Pareto Front) *the image of a Pareto Set, i.e.* $\{F(\mathbf{x}^*) | \neg\exists \mathbf{x} : \mathbf{x} \succ \mathbf{x}^*;\ \mathbf{x}^*, \mathbf{x} \in \mathbf{D} \subseteq \mathbf{X}\}$

Definition 6 (True Pareto Set and Front) *An approximation set such that* $\mathbf{D} = \mathbf{X}$ *is a true Pareto Set, likewise its image is a True Pareto Front.*

3.2 Evolutionary Algorithms

3.2.1 Fundamentals

In Artificial Intelligence, Evolutionary Algorithms (EA) belongs to a school of though, characterized by non-symbolic, non-statistical and often trial-and-error approaches, known as Soft Computing [214]. In particular, EA denotes a class of stochastic learning procedures that *evolve* towards a state of higher knowledge about a subject of concern, by the iterative application of learning rules (*heuristics*) inspired in some natural mechanism, especially those of living beings' dynamic. One can exploit this learning capacity to solve optimization problems, thereby evolving towards local optima. This facet of EA application shall constitute the kernel of our discussion about EA herein.

Iterative search is not at all exclusive of EA, but in this paradigm this feature acquire a different nuance. In classic optimization methods like derivative methods or conjugate directions methods, the iterative process is designed to exploit some mathematical properties associated with the shape or landscape of the objective function. Another characteristic of this paradigm is the use of a single vector at each iteration. In EA, however, the dynamic of exploration does not rely, in general, upon the objective function but upon the current sample of (many) potential optimal solutions to such function. These solutions are combined according to some rules that guarantee a global improvement of the sample through successive iterations.

To account for the philosophy behind EA, consider the principles of natural evolution postulated by Charles Darwin: *survival* is the central problem to be

[1] The same relations can be redefined for maximization just changing $\leq$ for $\geq$. Operators $\succ$ and $\succeq$ stand for both minimization and maximization in this work.

Nature		EA
Natural selection: how to survive?	⇒	**Artificial selection:** how to approach the optimum?
Individual: living being.	⇒	**Individual:** decision vector.
Population's origin: unknown.	⇒	**Population's origin:** initial sample.
Natural encoding: DNA.	⇒	**Artificial encoding:** representation scheme.
Natural mechanism of reproduction: crossover, mutation, cloning.	⇒	**Recombination mechanisms:** emulation of crossover and mutation.
Temporal step: lifetimes.	⇒	**Temporal step:** iterations.

Table 3.2: The abstraction of Darwinian postulates in EA: how life is emulated to solve optimization problems through EA (adapted from [135]).

solved by living beings. This is tackled in nature by individuals and by populations, and the interaction between these two instances is the base of *natural selection*. Thereby populations learn how to survive (*adaptation*) through an iterative process where the best individuals have higher odds to breed (*natural selection*), hence giving their successful features to the next generations. Therefore, if survival is now viewed as optimization and thus individuals become decision vectors, the iterative process will tend to find local optima if the mechanisms of *selection* and *recombination* are suitable emulated [135]. Some of the *metaheuristics*[2] inspired in Darwinian postulates are:

Genetic Algorithms (GA): Evolve populations of potential solutions to a particular problem, that born, breed and die according to their distance to the optima (*fitness*),

Evolution Strategies (ES): Similar to GA, but based on real vector optimization with self-adaptative mutation rate,

Genetic Programming (GP): Evolves computer programs structures.

To gain a deeper insight into these topics see the classical treaties [84, 64, 138, 8, 9]. Table 3.2 summarizes how these algorithms emulates life.

Darwinian algorithms are broadly known and have been successfully applied to a wide variety of problems. Yet, life also offers to EA many non-genetical sources of inspirations to explore domains in the search of optima. Some of the metaheuristics behind these alternative approaches are [214]:

Differential evolution: Search based on vector differences [196],

Particle swarm optimization: Search based on the observation of animal flocking behaviour,

Ant colony optimization: Search based on the ideas of ant foraging by pheromone communication to form path.

[2] In EA's jargon, *metaheuristics* denotes a class of heuristics. Thereby, GA is a metaheuristic and any implementation of GA is just an heuristic.

Algorithm 1 Evolutionary Algorithm: $\mathbf{EA}(F_a(I), M, t^{max})$

```
    /* Initialization */
1:  Initialize(P^(0))
2:  t ← 0
    /* Exploration */
3:  while not Terminate(P^(t), t, t^max) do
4:    M_p := Select(P^(t), F_a(I), M)          /* Select individuals to recombinate */
5:    P^(t+1) := Generate(M_p, M)              /* Generate a new sample (offspring) */
6:    P^(t+1) := Update(P^(t+1) + P^(t), F_a(I), M)          /* Update population */
7:    t ← t + 1
8:  end while
9:  Output(P^(t))
```

Figure 3.1: Algorithm 1: Evolutionary Algorithm.

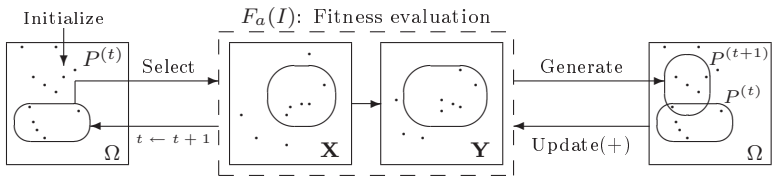


Figure 3.2: Flowchart of an EA.

3.2.2 Structure of EA

The algorithmic structure of a EA is presented in fig. 3.1. This pseudocode is, to our opinion, an improved version of that presented in [115, pg. 16]. It is complemented graphically by fig. 3.2, which sketches the flow of information when performing an EA search. Others versions of general EA pseudocodes can be found in [181, pg. 21][110, pgs. 28-29][203, pg. 2-19][235, pgs. 21-22].

As sketched in fig. 3.2, EA work with three spaces, viz. the search space Ω composed of *individuals* I, the decisional space $\mathbf{X}$ and the objective space $\mathbf{Y}$. The mapping $\rho(I) : \Omega \rightarrow \mathbf{X}' \subseteq \mathbf{X}$ is called the representation or encoding of individuals, and emulates the genotypic codes found in DNA[3]. In practice every representation is based on some (discrete and finite) numeric or symbolic alphabet, and every dimension defined from such alphabet as well. Thus, if we want to codify n-dimensional vectors $\mathbf{x}$, the number of states that any of their component x_i can adopt is restricted by the encoding. The result is that the use of mapping $\rho(I)$, whatever may be its functional expression, obliges to restrict the exploration of $\mathbf{X}$ to a discrete subset $\mathbf{X}' \subseteq \mathbf{X}$. Hence the representation is a critical issue regarding the efficacy and efficiency of an EA. Some typical system of encoding are:

Binary representation where individuals are represented by means of binary numbers,

[3]Space Ω is traditionally called as the genotypic spaces and $\mathbf{X}$ as the phenotypic space. However, we find this terminology confusing and unnecessary, rather we prefer to use the terminology mentioned in the text.

Gray code where binary numbers are assigned is such a way that to an unitary change in x_i correspond a change in the binary string of I with Hamming distance equals one,

Real representation real numbers are represented as themselves, so $\rho(I) = \mathbf{x}$ is an identity function and $\Omega = \mathbf{X}'$,

Integer representation integer numbers are used to represent a discrete and finite list of states.

Further details about these schemata of codification and others not explained as the *messy Genetic Algorithm*, can be found in [110, 64, 138].

The first operation performed by an EA is the initialization ($\texttt{Initialize}(P^{(0)})$) which consists in taking a random sample from Ω to fill up the initial population $P^{(0)}$ with M individuals. Then the three main procedures $\texttt{Select}(P^{(t)}, F_a(I), M)$, $\texttt{Generate}(M_p, M)$, and $\texttt{Update}(P^{(t+1)} + P^{(t)}, F_a(I), M)$ are repeated until the termination condition is verified. Notice that `Select`, that fills up the *matting pool* M_p, and `Update` depend on the fitness or adaptation function $F_a(I)$ which is an expression that scores an individual I in terms of its quality (e.g. its distance to an optimum). This expression is somehow related to the objective function $F : \mathbf{X} \rightarrow \mathbf{Y}$ and serves to rank the population in order to determine which one are to be recombined by `Generate` and to be kept or removed by `Update`. `Generate` invokes some recombination operators traditionally called *crossover* and *mutation* which are nothing more than multiple-individuals (usually binary) and unary sampling operators respectively. On the other hand, `Update` ranks the union $P^{(t+1)} + P^{(t)}$ and reduces it to M individuals. Even when both procedures rely upon the fitness function, the way they handle the fitness scores are not necessarily the same.

3.2.3 Representation and Sampling Operators in EA

Representation

In EA field, individuals are usually referred to as *chromosomes*. Every chromosome is a vector of *genes*, where each *gene* is the encoding of some feature (normally a vector component) regarding the decision vectors $\mathbf{x}$. Each gene has a limited number of allowed states, so the cardinality of the search space is the cartesian product of all genes' domain cardinality.

As an example consider the binary string '1001100100' composed of two genes of five bits; hence the first gene is '00100' while the second is '10011'. Both can allow any state between '00000' and '11111' so each gene has cardinality 32 (2^5), and the search space has cardinality 32*32=1024.

In general, every binary individual $I \in \Omega \subseteq \{0,1\}^L$, where L is the length or number of bits of I. Vector I has 2^L possible states that can be related to $\mathbf{x}$ making $\rho(I)$ a linear relation between the co-ordinates pairs $(min\{I : I \in \Omega\}, min\{\mathbf{x} : \mathbf{x} \in \mathbf{X}\})$ and $(max\{I : I \in \Omega\}, max\{\mathbf{x} : \mathbf{x} \in \mathbf{X}\})$ [181, pg. 16]. In contrast, in real and integer encoding $\rho(I)$ is usually the identity function.

Genetic-based Sampling

As we stated before, EA search strategy emulates some natural mechanisms, whether inspired in the behavioural dynamics of populations or in genetics. To

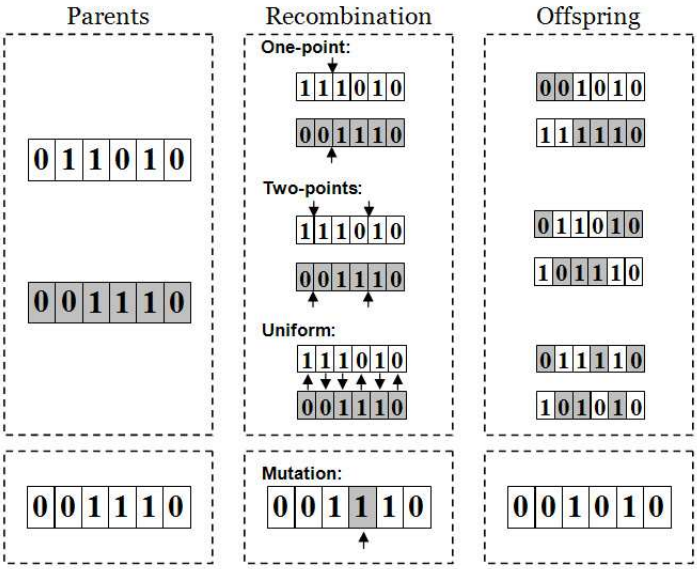


Figure 3.3: Example of the crossover mechanism for binary chromosomes [181, pg. 19].

the latter belong some widespread and well known operators called *mutation* and *crossover*. These two mechanisms are sampling operators of the space Ω that operate over one or multiple-individuals I respectively.

Fig. 3.3 exemplifies how they work with binary encoding. Notice that crossover takes two individuals (also called *parents*) to compose new individuals (also called *offspring*) by relating the bits' states of the offspring with those of the parents at the same position of the bit string. This way crossover generates new samples as the combination of some (encoded) characteristics of the previous samples. In contrast, mutation perform a local change in the bit string over only one individual, thereby generating a single new sample.

The intensity of the sampling is regulated by some control parameters called crossover (ρ_c) and mutation rate (ρ_m), which are the probabilities of applying such mechanisms each time `Generate` is invoked. Therefore in average each new population of size m has $m * \rho_c$ individuals produced through crossover. Mutation, however, allows different interpretation that rely on the EA designer. An extended practice to fix ρ_m is the rule of $1/L$, which established that, in average, the state of only one bit will be swaped for each time an individual mute.

By contrast, real representation requires a very different strategy of sampling. A common alternative for crossover is to perform an affine combination of parents, while mutation can be carried out by shifting the individuals by a gaussian noise.

3.2.4 Representation and Recombination in our EA-based Research

All the EA-based research directly or indirectly related to this thesis was performed relying partially [134, 180, 60, 58, 158, 234, 161] or totally [175, 174, 178, 176, 179] upon an *ad-hoc* code developed in C/C++ by the author. This code allows the use of mixed chromosomes and different types of crossover and mutation. The following subsections give some implementation details about this code.

Mixed Chromosomes

The remarkable feature about our implementation of chromosomes is the use of the type `union`, that allows handling the same string of bits from memory as different basic types simultaneously. Thus when a chromosome is programmed as a string of `union`, it is possible to have an heterogeneous vector, composed of integer, real and/or binary components.

To each `union` there is a structure attached that store the bounds of the associated component of **x** and other internal parameters of the code. Likewise, there are several mechanisms of crossover and mutation to be applied depending on the way the union is decoded (i.e. like binary, real or integer). Such mechanisms are explained next.

Such an implementation not only makes the code handy and appropriate to be used with different types of problems, but allows a better control of the search space size. This is especially noticeable when handling integer variables. With binary encoding, due to the restriction to 8-bit multiples, integer variables can be over represented, i.e. at least 7 bits may left over for each component. In order to assure efficiency, practitioners should take care to avoid this situation, in especial if the application is time-consuming. On the other hand, if a string-of-bits is adopted such that each gene (that takes values 0 or 1) corresponds to a basic variable (a byte), there would be an unnecessary waste of memory. These points are, naturally, open to discussion, but they are worthy to be mentioned since they affect the efficiency of academical and industrial application as well.

Crossover Operator

Fig. 3.4 exemplifies crossover for mixed chromosomes. The recombination mechanism is one-point crossover, adapted for a mixed chromosome. The crossover point is randomly selected between n possibilities corresponding to the number of components of the chromosome. In order to keep the possible values (states) of the integer or binary variables within their limits, if the crossover point hit on components of such types, components are only permuted but in any case a new value is calculated. The strategy changes when the crossover point lays on a real variable. In this case a convex combination is performed.

An alternative to the aforementioned mechanism is to allow crossing binary and integer components. For the binary case the procedure was explained earlier on. On the one hand, for integer encoded components, an integer mutation is applied over it.

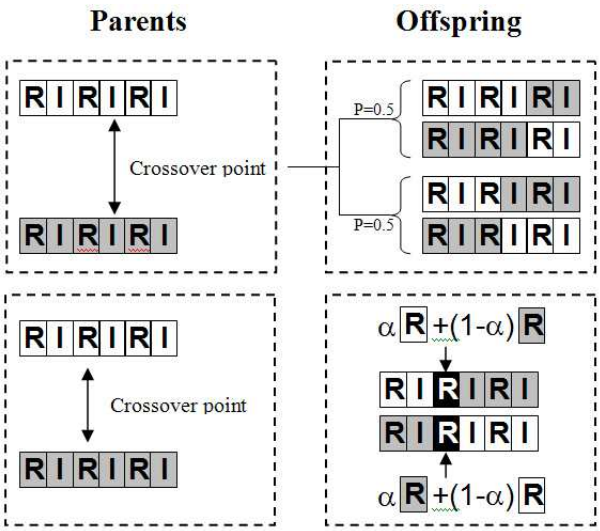


Figure 3.4: Example of the crossover mechanism for mixed chromosomes: real (R) and integer (I) variables [174, pg. 1062].

Mutation Operator

Real encoded are mutated with a gaussian noise of the type $x^{new} = x^{old} + N(0, \sigma^2)$ where $3\sigma = min\{|x - \underline{x}|, |x - \overline{x}|\}$ and $\underline{x}$ and $\overline{x}$ are the lower and upper bounds of x.

For integer variables two mechanisms are implemented. In the first one, with probability 0.4, the current value is replaced by a new value randomly chosen with equal probability from the list of possible values for that particular variable. The second mechanism selects, with probability 0.6, a new value through a *triangular mutation operator* that attempts to emulate the Gaussian mutation. Fig. 3.5 shows an example of how this mutation operator works.

First, note that the length of the triangle base is equal to the number of possible states the integer variable can assume. With the current state a triangle with total area 1 is built. Then a uniformly distributed random number $U(0, 1)$, equivalent to a cumulative probability or area A_r is generated. If A_r is less than the area up to the current value, the new state that replaces the current one is the state that makes the maximal triangle with area less than or equal to A_r (see fig. 3.5, dotted area). Likewise, if A_r is greater than the area up to the current state, the new state is that which makes the minimal triangle area larger than or equal to A_r (fig. 3.5, shaded area).

3.2.5 A Proposal: Percentage Representation

The percentage representation was first introduced in [175] as a response to an efficiency issue generated in a floating bound scheduling problem [134]. The problem is characterized by a dependency between the current value of some variables and the bounds of others.

As an example, assume that a chromosome is composed by two variables x_1 and x_2, where $x_1 \in [\underline{x_1}, \overline{x_1}]$ and $x_2 \in [\underline{x_2}, x_1]$ for some fixed bounds $\underline{x_1}$, $\overline{x_1}$ and

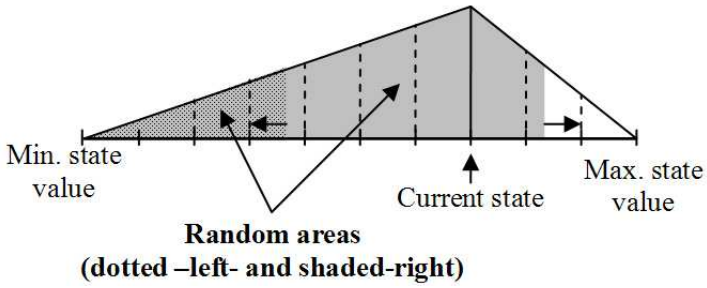


Figure 3.5: Triangular mutation operator [174, pg. 1063].

$\underline{x_2}$. Now suppose we have two individual $A = (x_1^A, x_2^A)$ such that $x_1^A \in [\underline{x_1}, \overline{x_1}]$ and $x_2^A \in [\underline{x_2}, x_1^A]$; and $B = (x_1^B, x_2^B)$ such that $x_1^B \in [\underline{x_1}, \overline{x_1}]$ and $x_2^B \in [\underline{x_2}, x_1^B]$. Thus, if x_1 is not previously fixed for all chromosomes; in the valid case when $x_2^A < x_1^A < x_2^B < x_1^B$ one potential offspring of a crossover of A and B are $A' = (x_1^B, x_2^A)$ which is feasible, and $B' = (x_1^A, x_2^B)$ which is unfeasible since it violates $x_2 \in [\underline{x_2}, x_1]$.

Clearly, to be able to explore the whole domain of the problem without generating unfeasible individuals, x_1 must be fixed before assigning values to x_2. This leads to a '*double-loop*' configuration, such that the value of x_1 is to be controlled by the external loop while the value of x_2 is handled by the nested loop. Consequently, the search space is composed of a set of disjoint subspaces.

In the percentage representation, the values of dependent variables are considered as percentage of the variables they depend on. In this way, the search space is defined as an augmented domain where only feasible individuals are generated and the recombination operators can be applied without restrictions. The next subsections offers an insight into the the search space issue regarding the percentage representation, through an industrial application.

3.2.6 Search Space Analysis and Percentage Representation: An Industrial Application Example

Problem Setting

The following case of study was presented in [175] as an alternative to solve the double-loop algorithm proposed in [134] through percentage representation.

As a part of the technical specifications of nuclear power plants, the surveillance requirements establish the Test Intervals (TI) and the Test Strategy or Planning (TP) of the safety systems of such nuclear plants. Given a safety system, its TI consists in a period between tests, whereas the TP consists in the schedule for testing each component of the system. Fig. 3.6 depicts a simplified high pressure injection system, which is used to remove heat from the reactor in some accidents. The system consists of 3 pumps and several valves. For this system, it is necessary to establish the TI, which comprises 3 intervals $\{T1, T2, T3\}$, and the TP composed by 6 times to first test $\{TA, TB, TC, TD, TE, TF\}$. E.g., the decision variables associated to pump PA correspond to the set $\{T1, TA\}$. After some assumptions, the search space is reduced to four variables so that the decision vector $\mathbf{x}$ is $(T1, TA, TB, TC)$ [134].

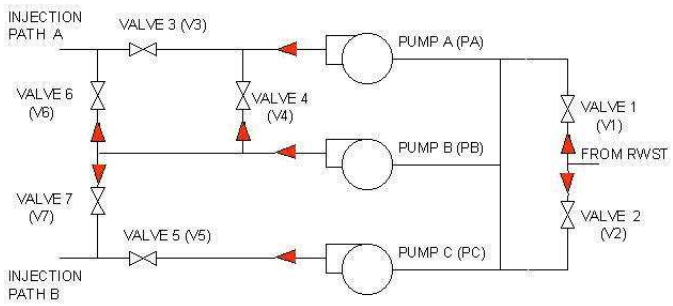


Figure 3.6: High Pressure Injection System [175][134].

The minimum and maximum values for the decision variables can be seen below. Notice that the lower bound are fixed for all the variables, but the maximum bound is fixed for only $T1$, whereas for TA, TB, and TC such bounds are variable, i.e. depends on the current value of $T1$. In [175] it offers more details about this setting:

x	**Minimum (days)**	**Maximum (days)**
$T1$	30	58
TA	0	Variable (equal to current T1-1)
TB	0	Variable (equal to current T1-1)
TC	0	Variable (equal to current T1-1)

Search Space Analysis

Given that the maximal period between consecutive tests is $T1$, the times for first test encoded in $\mathbf{x}$, cannot be higher than $T1$. For this reason, once the $T1$ value is chosen, the search space for this particular $T1$, denoted by $S(T1)$, is defined as the combinatorial space of (TA, TB, TC) which has size $T1^3$. Notice that all subspaces $S(T1)$ $\forall$ $T1$ are disjoint. Hence, in general, the complete search space S is a collection of disjoint sets, which has size:

$$|S| = |S(T1_{min})| + \ldots + |S(T1_{max} - 1)| = \sum_{n=T1_{min}}^{T1_{max}-1} n^3$$

According to the precedent equation, the $|S|$ for the problem considered ($T1 \in \{30, \ldots, 58\}$) is 2738296. In some cases a search space of such dimension is not affordable[4], resorting to the employment of alternative heuristics to speed up the search.

The first alternative studied in this example, namely a double-loop strategy, is derived in a natural way from the structure of the problem. Notice that in a double-loop scenario the search space is composed by disjoint subspaces, each of which is to be explored. In our example it means that an external loop fixes the values of $T1$, whereas an internal nested loop performs the EA over each subspace. The complete enumeration of subspaces is affordable since its number of subspaces is discrete, finite and small; otherwise such a strategy is

[4] In [175] a complete enumeration represented 760 hours of CPU for a complete enumeration.

not possible, as shall be seen later on in chapter 5. On the other hand, observe that, due to the double-loop works on disjoint subspaces, any solution found in one subspace is not necessarily defined in the others; consequently the good solutions cannot be extrapolated or mixed directly between subspaces. This fact seems to us to contradict the foundations of metaheuristics, for they try to take advantage of any good feature discovered so far during the process.

Hence, we introduced a second alternative, based on the observation that the dependency between the TI and the TP can be expressed in relative terms (percentages) instead of absolute terms, like in the double-loop configuration. In fact, a representation of the type $(T1, TA_{\%} = \%T1, TB_{\%} = \%T1, TC_{\%} = \%T1)$ unifies the search space making possible to explore the whole domain of $T1$ at the same time. In general, the percentage representation embeds the disjoint subspaces into an augmented unified space. Although the resulting search space is larger, it does not make this encoding less appealing, since this unification makes possible to perform genetic-based sampling without restrictions. However, in some cases analysts can find a trade-off to make the new space as smaller as possible.

Consider, e.g. the case when the percentages are one-byte encoded, the search space amounts to (58-30+1)*255^3 = 480859875, which is 175.6 times bigger than the double-loop's search space. Additionally, if $T1$ is larger than say 300, one byte is not enough to represent all possible states of TA, TB and TC. Of course, the use of bigger binary variables is possible, as well as the use of real representation for the percentages, but with the consequent augment of the search space.

The minimal unified search space that contains any point of the original one can be defined making $TA_{\%}$, $TB_{\%}$, and $TC_{\%}$ integer variables between 0 and $(T1_{max} - 1)$ and then decoding each individual by calculating the percentages in terms of its current $T1$ value as:

$$T = T_{\%} \frac{T1}{T1_{max}}$$

For example, consider two individuals A = (58, 45, 57, 30) and B= (40, 10, 50, 57) of type $(T1, TA_{\%}, TB_{\%}, TC_{\%})$. According to the above equation for $T1_{max} = 58$, the final (decoded) value of A is (58, 45, 57, 30), whereas for B the decoded value is (40, 6, 34, 39). In this way, one can apply crossover and mutation operators over any chromosome without restrictions, overcoming the limitations imposed by the disjoint structure of the original search space. Obviously, the expansion of the search space remains as a drawback, but the size of the space is smaller than in any other type of percentage representation. For the problem considered the search space has size[5] (58-30+1)*58^3 = 5658248.

3.3 Multiple-Objective Evolutionary Algorithms

Multiple-Objective Evolutionary Algorithms (MOEA) is a class of evolutionary algorithms especially tailored to deal with multiple-criteria problems. They merge the potentiality of metaheuristics with the some principles of MCDM thus yielding algorithms of outstanding capabilities.

[5]The trick consists in realizing that $T1$ cannot be divided into more than 58 parts, since (TA, TB, TC) are integers. Percentage of real variables must be treated in a different way.

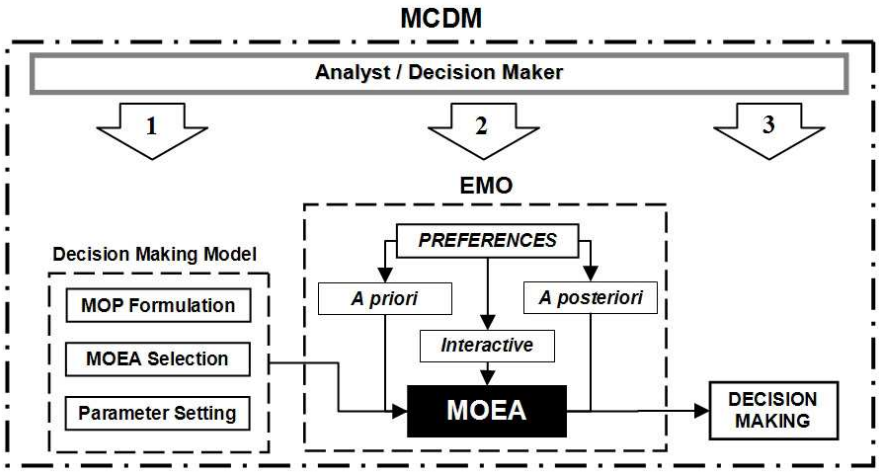


Figure 3.7: A model for MCDM and EMO: MOEA as a black box optimizer embedded in a decision-making process (taken from [172]).

In practice, every MOEA is embedded in a decision-making process that entails three stages as is shown in fig. 3.7. To our opinion, the right understanding of the strengths and weaknesses of EMO, in special concerning salient issues as uncertainty handling, is only possible within this broader scope.

Even when EMO is a hybrid branch relatively new in comparison with its roots, namely MCDM and EA, there is an extensive amount of pages devoted to EMO[6] in the applied sciences and engineering. Hence, this section will focus upon those topics that we consider salient issues in EMO, in particular uncertainty handling, while reserving to the remainder topics only a brief mention.

3.3.1 Historic Outline

EMO has a short but exciting history that can be traced back two decades, within which the research efforts have met two fundamental phases (fig. 3.8). According to Coello [28] the initial stage or first generation is characterized by an emphasis in simplicity. To this group correspond the first attempts in EMO, a heterogeneous collection of EA more concerned with efficacy than efficiency. The approaches were rather simple and there were no consensus about the techniques employed to classify the global quality. To this generation belong some methods like the aggregation method, the ϵ-constrained method, NSGA [195], NPGA [90] and MOGA [53].

First generation MOEA allowed Pareto and non-Pareto ranking. A typical example of the latter is the aggregation method. This is an *a priori* method, i.e. the DM is expected to express any kind of preferences before performing the optimization. Then the objectives are aggregated according with the weights

[6]This field is nowadays very popular between researchers of computational sciences and engineering. EMO has a biannual conference entirely devoted to it that counts already four editions (EMO 01, 03, 05, 07), as well as special track sessions in the most important local and international scope conferences all over the world (e.g. CEC, GECCO, etc.). Likewise there is an extense public repository of publications in EMO in http://www.lania.mx/~ccoello/EMOO/

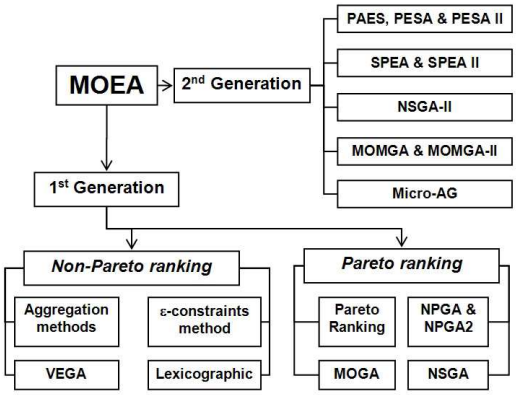


Figure 3.8: Historic development of EMO (taken from [181, pg. 24]).

assigned in a convex combination of the form

$$\sum_{i=1}^{k} w_i f_i(\mathbf{x}) \tag{3.3}$$

such that $w \geq 0$ and $\sum_{i=1}^{k} w_i = 1$. This way the multiple-objective problem collapses into a single-objective problem easily solvable, at least in principle, with either heuristics or classic optimization procedures. If the DM is not able to weight the objectives properly, the analyst might repeat the procedure for a representative number of weights, thereby getting a good sample of the True Pareto front. However, aggregation procedures are usually subject of critics due to its inherent inefficacy facing non-convex spaces [27].

The introduction of Pareto ranking methods during this phase were strongly influenced by David E. Goldberg, who suggested a Pareto ranking procedure in his book [64]. The procedure was incorporated in NSGA and later on in NSGA-II (see fig. 3.13a) and constitutes a key-stone in MOEA, since it demonstrated the potential of Pareto ranking. First generation methods also glanced other important issues as constraint handling, preference incorporation and optimality/diversity trade-off.

Second generation methods, on the other hand, comprises more elaborated methods that incorporates successful features that formerly enhanced single objective EA, namely elitism, complexity analysis, parallelization, hybridization, amongst others, as well as a focus on Pareto-based ranking methods. The standardization of elitism, the introduction of the ϵ-dominance [117] and the development of a genuine concern on efficiency and quality of MOEA are considered keystones [52] of the current state-of-the-art in EMO. In that sense, some of the main contributions during this period are the popularization of MOEA by means of an extensive number of benchmark studies [240] along with the introduction of high-efficiency algorithms (like SPEA2 [239], NSGA-II [36], PAES [107], PESA [108]) and the improvement on the methodological issues related with quality assessment, viz. the introduction of test functions frameworks for benchmark studies [33, 37] and a considerable effort in quality metrics [241, 54].

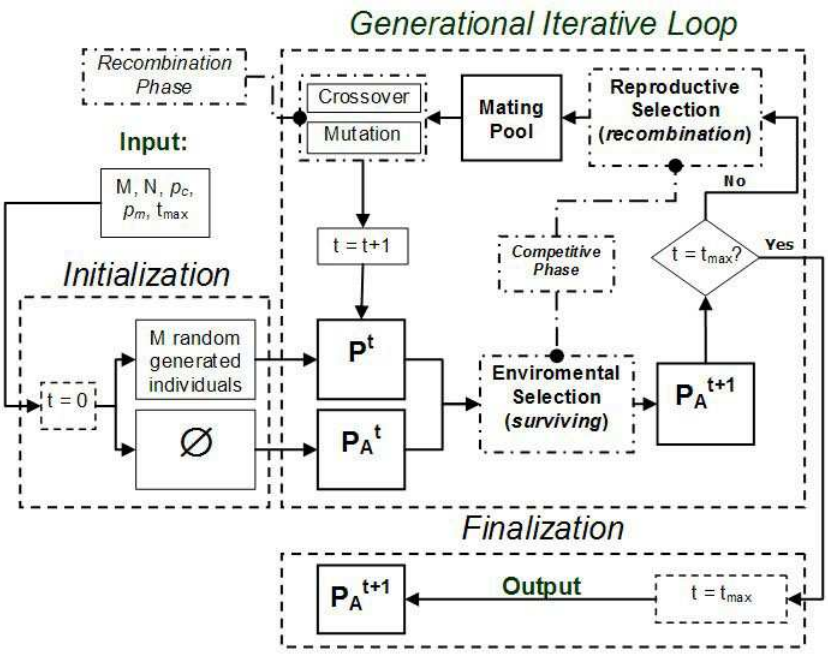


Figure 3.9: Flowchart of a second-generation MOEA based on GA (taken from [174, pg. 1060]).

3.3.2 Algorithmic Structure

Every MOEA attempts to solve a multiple-objective problem of the form shown in eq. 3.2 by tackling the following questions:

1. How to sample the decisional space?
2. How to rank the alternatives found?
3. How to keep the good solutions?

The way these questions are answered distinguishes each realization of MOEA found in the literature. The first question concerns the use of the current information to generate new information, i.e. the selection schema and the search engine. The former feature defines the seed while the latter the rules to sample the decisional space. The second question aims at classifying the solutions in terms of their global quality (Pareto optimality) and local quality (density). Finally the third question refers to how many and what alternatives are to be presented to the DM amongst the all optimal solutions discovered during the evolutionary search.

The three constituent elements connected to the preceding questions are: search engine, fitness assignment, archiving and selection strategies. These elements can be designed and arranged in different fashion, but most of them follows a structure similar to that depicted in fig. 3.9. The other crucial factor is the parameter setting, which is the subject of many benchmarks and studies of the like (e.g. [181]), since it has a direct impact on the efficiency of MOEA.

Search Engines

Search engines are based on the different metaheuristics (see sec. 3.2.1) that can be applied to explore the search space. In other words, one can recourse to different search strategies to sample the decisional space with MOEA. Genetic Algorithms are perhaps the most popular search engine, but some others have gained importance like Differential Evolution [154, 219, 78], Particle Swarm Optimization [153] and Ant Colony Optimization for combinatorial problems (e.g. [24]). Non-evolutionary metaheuristics have been successfully used as well like Multiple-Objective Simulated Annealing (MOSA). The selection of one or another can impact the speed of convergence as well as the spread of the final Pareto approximation front. This is noticeable when switching from continuous to discrete decisional spaces.

Fitness Assignment

In MOEA, candidate alternatives must compete for not being discarded and for seeding the next samples. This is comparable to the environmental and reproductive selection in nature. Both process are ruled by figures of merit that reflect the global and local quality of each alternative amongst the population.

Global quality is assessed in modern MOEA in terms of Pareto dominance (see sec. 3.1.2). Some optimizers, like PAES [107] and PESA [108], discard every solution that come to be dominated during the evolutionary search. In contrast, some other popular methods make use of non-dominated and dominated solutions as well, but giving the former group much more chances to breed and to survive. Some typical schemes are [238]:

Dominance Rank: the smaller the number of vectors that dominate an individual the better, e.g. MOGA [53], NPGA [90], SPEA [235, 240] and SPEA2 [239] (see fig. 3.15),

Dominance Count: the larger the number of vectors an individual dominates the better, e.g. SPEA [235, 240] and SPEA2 [239] (see fig. 3.15), and,

Dominance Depth: the closer to the optimal the front an individual is located the better, e.g. NSGA [195] and NSGA-II [36] (see fig. 3.13).

Local quality, on the other hand, aims at promoting well spread and even distributions amongst the approximation fronts. For that matter, different density-measuring strategies have been proposed, being the most representative in state-of-the-art MOEA [238]:

Nearest Neighbour: the larger the distance to the kth neighbour the better. This scheme, adopted e.g. by SPEA2 [239], measures the density in terms of Euclidean distances, thereby promoting those alternatives with larger hyper-spherical empty volumes around them.

Crowding Distance: each individual is characterized by the Manhattan distance between its two nearest neighbours, so the larger the distance the better. This approach was introduced in NSGA-II [195] (see fig. 3.13b).

A recent trend in MOEA constitutes the indicator-based optimizers. This class of algorithms aims at optimizing a global quality indicator based on performance metrics (see sec. 3.3.5). Such metrics attempt to express in a single

Algorithm 2 Infinite Size Archiving Algorithm: $\mathbf{Archive_0}(P_A, I)$

1: If I is not dominated by the members of P_A then $P_A := P_A + I$
2: If I dominates any member I' of P_A then $P_A := P_A - I'$
3: `Output`(P_A)

Figure 3.10: Algorithm 2: Infinite Size Archiving Algorithm.

scalar both the global and local quality of an approximation front. Concomitantly, the subset of the current population that maximizes the indicator is considered the best amongst the whole population and thus the heuristic efforts focus on increasing such indicator. Some examples of this strategy can be found in [45, 144, 237].

Archiving and Selection Strategies

Promising individuals are selected to survive and to seed the next sampling. In modern MOEA the criteria to enter the archive are almost the same to perform the mating selection, i.e. fitness expresses global and local quality (giving prevalence to the former over the latter) in a scalar form. In consequence, the mating selection is possible for instance, by simply taking two candidates from the population and keeping that with better fitness (binary selection). Since elitism in MOEA is possible through a external memory or archive, mating selection is performed on this component.

Archiving is one of the fundamentals of MOEA. Since multiple-objective problems does not have a single optimum but a set of efficient solutions, these have to be stored as they are found. Therefore the archiving strategy directly affects the quality of the information provided to the decision maker. On the other hand, archiving provides a repository for elitism in MOEA that has a crucial effect on the performance of the algorithm [119].

Archives can be of finite or infinite size. The latter case simply indicates that every efficient solution is to be incorporated into the archive, whereas the dominated ones are wiped out (see fig. 3.10). The former case, however, is more complex since it entails a selection of elements not only in terms of their efficiency but its distribution along the Pareto approximation front as well.

Finite size archiving implies a limited storage capacity of non-dominated solutions. Some strategies accept only non-dominated solutions into the archive, which implies an upper-limit bounded size. In contrast, other approaches accept dominated solutions when the number of non-dominated ones is not sufficient to fill up the archive. Currently this scheme seems to be the most popular amongst practitioners.

A generic finite size archiving procedure is described in fig. 3.11. The inclusion of a new non-dominated solution may generate a reduction of the archive if the incoming individual I dominates some members of the archive P_A. Otherwise the size of the archive must be checked and if it exceeds the limit n, then the archive have to be truncated.

The procedure `Truncation`(P_A) distinguishes one archiving approach from another. Modern MOEA always keep the extreme solutions and begin pruning the more crowded zones. As we stated in the previous section, optimizers like SPEA2 and NSGA-II carrying it out by deleting those solutions whose distances

Algorithm 3 Finite Size Archiving Algorithm: $\mathbf{Archive_1}(P_A, I, N)$

1: If I is not dominated by the members of P_A then:
2: $P_A := P_A + I$ /* Adds the new individual */
3: If I dominates any member I' of P_A then $P_A := P_A - I'$
4: If `Size`$(P_A) > N$ then:
5: `Truncation`(P_A) /* The archive is reduced to N individuals */
6: Else go to 8
7: Else `AddsDominated`(P_A, I, N) /* Decides the addition of the new individual */
8: `Output`(P_A)

Figure 3.11: Algorithm 3: Finite Size Archiving Algorithm.

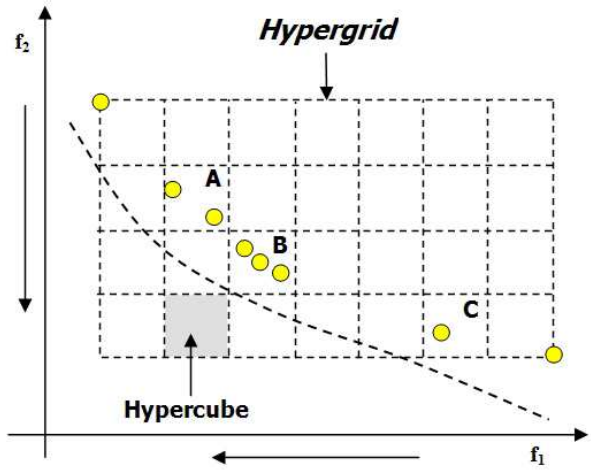


Figure 3.12: Hypergrid archiving strategy introduced by Knowles et al. [107](taken from [181, pg. 35]).

to their nearest neighbours are smallest amongst the population. The difference relies upon the way the distances are measured, resorting to Manhattan or Euclidean norms respectively. In contrast, other approaches like the Hypergrid (see fig. 3.12) divides the objective space evenly in boxes or hypercubes, and estimates the density of an individual as the number of solutions that share the same box. Therefore the most crowded box (hypercube B in the figure) is reduced by deleting at random one of its contained solutions. This approach was introduced as a feature of PAES [107] but has been implemented in other optimizers [26].

The other distinctive feature in archiving is the `AddsDominated`(P_A, I, N) procedure. The hypergrid does not accept dominated solutions which means that this procedure is not included in the hypergrid. Indicator-based procedures neglect dominated solutions as well. When dominated vectors are allowed, they enter the archive when either the maximal size has not been reached or the incoming vector is better than other dominated vectors in the archive. 'Better' means higher global quality (the individual dominates solutions contained in the archive) or same global quality and better local quality (the individual is

placed in a less crowded area with respect to its counterparts in the archive).

Despite the of the way `Truncation`(P_A) and `AddsDominated`(P_A, I, N) are implemented, every finite archive strategy is expected to meet two desirable properties:

True Pareto convergence: When the size of the approximation front reached surpasses the size of the archive, it is necessary to give up the exceeding. Even when the archive is an approximation set, i.e. all its content is made up of non-dominated solutions, only a portion of such content -or perhaps nothing- belongs to the True Pareto frontier. If it does happen the pruning of the exceeding should favour true Pareto solutions over non-true Pareto solutions. This is known as True Pareto convergence property.

Promotion of well distributed fronts: Since most general-purpose MOEA approaches assume *a posteriori* exercise of preferences, approximation fronts are expected to be even in order to offer the DM an unbiased front. The extension of the approximation set is another crucial parameter of quality. Even distributed samples of the true Pareto frontier along its whole extensions are the ideal.

Different archiving schema have been proposed so far, each one showing with different levels of success in achieving the aforementioned properties. The current trend is to make the better use of ϵ-dominance to assure True Pareto convergence. The so called ϵ-archiving has been studied both theoretically and practically [117, 118, 116], showing a tremendous potential to improve significantly the performance of MOEA.

3.3.3 ϵ-Dominance

In this subsection we bring an additional definition that have a considerable weight in modern EMO is the ϵ-dominance notion. In terms of minimization [117, 116]:

Definition 7 (ϵ-Dominance) $\mathbf{x}^1$ *dominates* $\mathbf{x}^2$*, denoted* $\mathbf{x}^1 \succ_\epsilon \mathbf{x}^2$*, iff for* $i \in \{1, 2, \ldots, k\}$ *and* $\epsilon > 0$ *it verifies:*

$$\begin{array}{rcll} f_i(\mathbf{x}^1) & \leq & (1+\epsilon_i) f_i(\mathbf{x}^2) & \textit{(multiplicative form)} \\ f_i(\mathbf{x}^1) & \leq & f_i(\mathbf{x}^2) + \epsilon_i & \textit{(additive form)} \end{array} \tag{3.4}$$

In essence the ϵ-dominance means that given two vectors $\mathbf{x}^1$ and $\mathbf{x}^2$ such that $\mathbf{x}^2 \succ \mathbf{x}^1$, if the amount of such domination is not greater than ϵ, i.e. if $\mathbf{x}^1 \succ_\epsilon \mathbf{x}^2$, then the real dominance is not relevant in terms of classification. In other words, during the evolutionary search at least one member of the population is expected to improve in a proportion greater than ϵ to consider the sample successful.

3.3.4 Some Representative MOEA

Amongst the different heuristics that have been proposed as general-purpose MOEA, only a few have become popular between practitioners. In this section we bring some details of relevant MOEA related directly or indirectly with the research presented in this work.

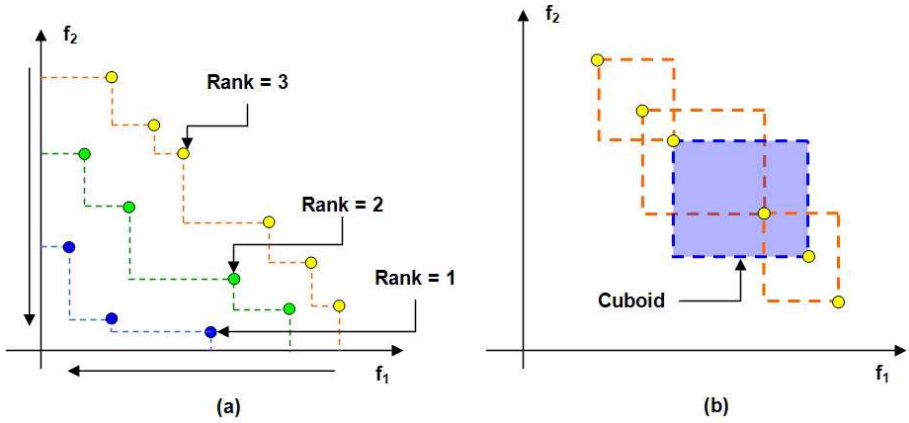


Figure 3.13: Quality assessment in NSGA-II: (a) Dominance depth ranking and (b) Cuboid-based density control in minimization (taken from [181, pg. 30][174, pg. 1061]).

Nondominated Sorting Genetic Algorithm II (NSGA-II)

The NSGA-II [36] is a fast and efficient algorithm that has attracted the attention not only from the entire EMO community but from many applications fields as well. The main feature of NSGA-II is the crowding operator, that performs two operations of classification. The first operation reflects the global quality of the solutions in terms of a depth-ranking, i.e, the non-dominated set of the population is determined and labeled with the best fitness, then it is removed from the population and the new non-dominated set of the resulting population is determined. This process is repeated until the whole population is labeled in terms of depth, as suggested in [64] (see fig. 3.13a). The second operation is the assessment of the relative density or local quality of the solutions regarding the front they belong to. Each solution fitness is added with inverse of its cuboid, or the Manhattan distance between the two nearest neighbours. Extreme solutions has infinite cuboids. This way each front is subclassified giving better fitness to those solutions placed in less crowded areas (see fig. 3.13b).

Our implementation of NSGA-II is reported in fig. 3.16. Notice that the fitness function $F_a(I)$ corresponds to the crowding operator. The subroutine $\texttt{Rank}(P^{(t)}, F_a(I))$ sorts the argument according to $F_a(I)$. $\texttt{Update}(P^{(t)}, F_a(I), N)$ assigns the N best individuals of the argument to the new archive, whereas $\texttt{Select}(P_A^{(t+1)}, F_a(I), M)$ fills the mating pool with $M/2$ pairs, each one producing two new individuals. This implementation was employed in [175, 174, 178, 176, 179, 58, 158, 234].

Strength Pareto Evolutionary Algorithm 2 (SPEA2)

The SPEA2 is a powerful alternative for solving MOP, based on a mixed mechanism of ranking, where count and rank dominance are combined by means of the fitness function $F_a(I) = R(I) + D(I)$. The contribution of $R(I)$ was modified from SPEA [240] to SPEA2 [239] as is depicted in figures 3.15 (a) and (b) respectively. Fitness is to be minimized; this way SPEA2 assigns zero fitness to

Algorithm 4 Nondominated Sorting GA II: **NSGAII**$(F_a(I), M, N, t^{max})$

```
    /* Initialization */
 1: Initialize(P_A^(0))                          /* Initialize population randomly */
 2: P^(0) := ∅
 3: t ← 0
    /* Exploration */
 4: while (t < t^max) do
 5:   P^(t) := P^(t) + P_A^(t)
 6:   Rank(P^(t), F_a(I))                        /* Apply Crowding Operator */
 7:   P_A^(t+1) := Update(P^(t), F_a(I), N)      /* Store n best individuals */
 8:   M_p := Select(P_A^(t+1), F_a(I), M)        /* Select m individuals to recombinate */
 9:   P^(t+1) := Generate(M_p, M)                /* Generate a new sample (offspring) */
10:   t ← t + 1
11: end while
    /* Finalization */
12: Output(P_A^(t))
```

Figure 3.14: Algorithm 4: Nondominated Sorting Genetic Algorithm II (NSGA-II).

all non-dominated vectors, whereas the fitness of the remainder is calculated as the sum of the number of vectors dominated by those vectors that dominate a particular solution plus a density factor D(I) (see fig. 3.17).

The truncation operator $\texttt{Truncate}(P_A^{(t+1)}, N)$ in fig. 3.17 is a complex mechanism that eliminates those solutions placed in more crowded areas according to their Euclidean distance to the kth nearest neighbour. This is done by means of the $\leq_d$ operator which works in such a way that $I \leq_d J$ entails $\sigma_I^k = \sigma_J^k$ for all k or $\sigma_I^l = \sigma_J^l$ for all $l < k$ and $\sigma_I^k < \sigma_J^k$ for a particular k [238]. Thus $\leq_d$ is applied recursively to delete that solution that verifies $I \leq_d J : \forall J \in R^{(t)}$ until the top size N is reached.

Multiobjective Simulated Annealing (MOSA)

General purpose MOEA like NSGA-II can be employed to solve discrete problems with more or less success; however some heuristics have been especially developed to cope with combinatorial optimization problems, like the Multiobjective Simulated Annealing (MOSA) method. This algorithm is based on single-objective simulated annealing but introduce an archiving procedure for non-dominated vectors. Fig. 3.17 shows the MOSA algorithm proposed by Ulungu et al. [201].

Some remarks about this algorithm: first of all, notice that it is possible to promote some search directions varying the weights λ_i in Δ_S. This fact might be very useful if there is some knowledge about the regions of interest *a priori*; however, if the archiving routine is applied according to what is shown in line 6 to 9, some non-dominated solutions may be disregarded due to the bias imposed by Δ_S. Some variations can be introduced in order to handle such an effect, like invoking the archiving routine before updating the reference individual $I_{ref}^{(t)}$. Likewise, the type of archiving is to be selected. Infinite (fig. 3.10) or finite archiving (fig. 3.11) are indeed possible for MOSA.

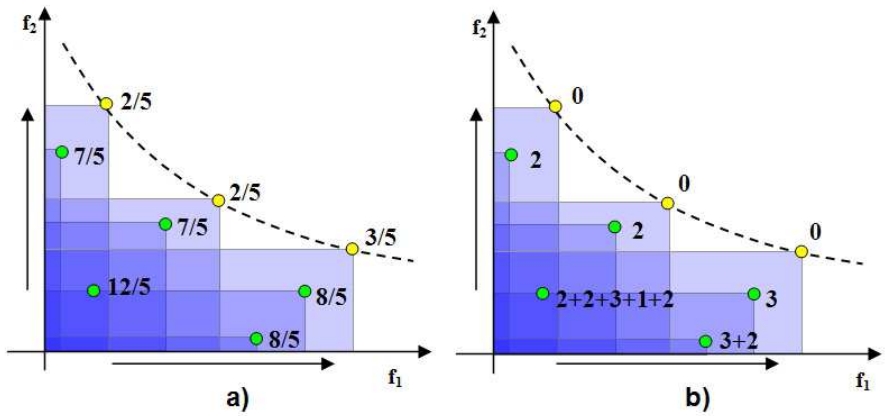


Figure 3.15: Dominance rank + count strategy in (a) SPEA and (b) SPEA2 for maximization (taken from [181, pg. 33]).

Algorithm 5 Strength Pareto EA 2: **SPEA2**$(F_a(I), M, N, t^{max})$

/* Initialization */
`Initialize`$(P^{(0)})$ /* Initialize population randomly */
$P_A^{(0)} := \emptyset$
$t \leftarrow 0$
/* Exploration */
while $(t < t^{max})$ **do**
  $R^{(t)} := P^{(t)} + P_A^{(t)}$
  `Rank`$(R^{(t)}, F_a(I))$ /* Assign adaptation $F_a(I)$ */
  `Sort`$(R^{(t)})$ /* Sort population in increasing order of $F_a(I)$ */
  $P_A^{(t+1)} := \{I \in R^{(t)} | I \text{ is not dominated}\}$ /* Update the archive */
  If $|P_A^{(t+1)}| < N$ then
    $P_A^{(t+1)} := P_A^{(t+1)} + \{\text{First } N - |P_A^{(t+1)}| \text{ individuals } I \in R^{(t)}\}$
  Else
    `Truncate`$(P_A^{(t+1)}, N)$ /* Reduce archive to N individuals*/
  $M_p :=$ `Select`$(P_A^{(t+1)}, F_a(I), M)$ /* Select m individuals to recombinate */
  $P^{(t+1)} :=$ `Generate`(M_p, M) /* Generate a new sample (offspring) */
  $t \leftarrow t + 1$
end while
/* Finalization */
`Output`$(P_A^{(t)})$

Where:

$F_a(I) = R(I) + D(I)$
$S(I) = |\{I' : I' \in R^{(t)} \wedge I \succ I'\}|$ /* Number of individuals I dominates */
$R(I) = \sum_{I' \in R^{(t)}, I' \succ I} S(I')$
$D(I) = \frac{1}{\sigma_I^k + 2}$
σ_I^k is the distance to the kth nearest neighbour, with $k = \sqrt{M+N}$

Figure 3.16: Algorithm 5: Strength Pareto Evolutionary Algorithm 2 (SPEA2).

Algorithm 6 MO Simulated Annealing: **MOSA**$(F_a(I), \alpha, T^{(0)}, T^{max}, t^{step}, t^{max})$

/* Initialization */
1: $I_{ref}^{(0)} \leftarrow$ `Initialize`$(I^{(0)})$ /* Initialize reference individual at random */
2: $P_A^{(0)} := I^{(0)}$
3: $t \leftarrow 0$
/* Exploration */
4: **while** ($t^{count} < t^{max}$ AND $T^{(t)} \geq T^{max}$) **do**
5: $I^{(t)} \leftarrow$ `GenerateNeighbour`$(I_{ref}^{(t)})$ /* Draw randomly a neighbour of $I_{ref}^{(t)}$ */
6: If `Decide`$(I_{ref}^{(t)}, I^{(t)})$ then
7: $I_{ref}^{(t)} \leftarrow I^{(t)}$
8: $P_A^{(t+1)} :=$ `Archive`$(P_A^{(t)}, I^{(t)})$ /* Update archive with $I^{(t)}$ */
9: If $P_A^{(t+1)} \neq P_A^{(t)}$ then $t^{count} \leftarrow 0$
10: Else
11: $I_{ref}^{(t+1)} \leftarrow I_{ref}^{(t)}$
12: $t^{count} \leftarrow t^{count} + 1$
13: If (t `mod` t^{step})=0 then $T^{(n+1)} = \alpha * T^{(n)}$ else $T^{(n+1)} = T^{(n)}$
14: $t \leftarrow t + 1$
15: **end while**
/* Finalization */
15: `Output`$(P_A^{(t)})$

Decide$(I_{ref}^{(t)}, I^{(t)})$

1: If $\Delta_S \leq 0$
2: return `true`
3: Else
4: with probability p return `true`
5: Otherwise return `false`

Where:

$p = exp\left(-\frac{\Delta_S}{T^{(n)}}\right)$

$\Delta_S = \sum_{i=0}^{k} \lambda_i \left(f_i(I_{ref}^{(t)}) - f_i(I^{(t)})\right)$

Figure 3.17: Algorithm 6: Multiobjective Simulated Annealing (MOSA).

3.3.5 Performance Measures

The quality assessment of an approximation front has shown to be one of the most challenging tasks and obviously an open issue, in EMO. The first attempts were based on 'unary' metrics, i.e. on measures of quality that involve only one approximation front. While these metrics are useful when an utility function on the members of the approximation front can be defined [71], both empirical and theoretical results [241] shown the limitations of such metrics to draw coherent conclusions when no additional information is available.

In response, researchers proposed the use of 'binary' metrics, i.e. quality indicators based on pair comparisons. The theoretical framework introduced in [241] allows to analyze the power of inference of unary and binary metrics as well, and has served as reference for a methodological improvement in quality assessment.

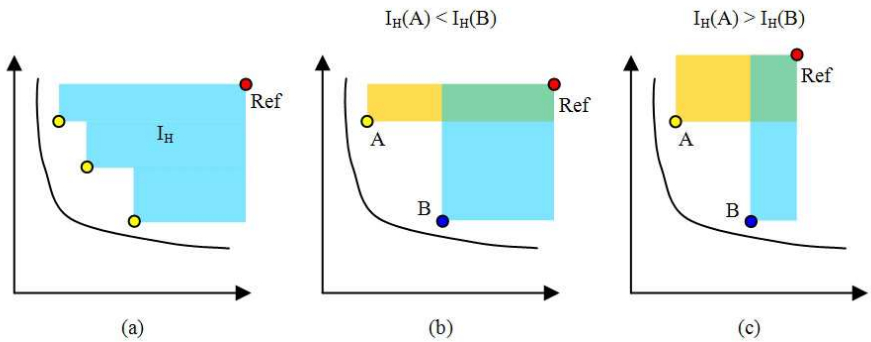


Figure 3.18: Example of the hypervolume or S metric in minimization: the conclusions drawn from the hypervolume metric (a) depends on the reference point: $I_H(A) < I_H(B)$ (b) and $I_H(A) > I_H(B)$ (c) for the same outcomes A and B.

Two trends proposed in [55] to practitioners conform with what is expected after the horizon set by [241]. The first trend consists in a statistical assessment of the level of attainment of the approximation fronts (e.g. [54]), i.e. after several experiments the probability of attaining some regions of the objective space is empirically determined. In contrast, the second trend is to use dominance-compliant indicators that assess the quality with help of some extra information, like a reference point or set or an utility function. The recommended indicators are the S-metric or hypervolume indicator [235], the unary ϵ-indicator [241] and Hansen and Jaszkiewicz's [71] based R_R indicators [241].

S-metric

The hypervolume metric I_H of an approximation front corresponds to the Lebesgue measure of the polytope defined by such a front and a reference point (see fig. 3.18). When comparing two fronts, the one with higher hypervolume is preferred since -as is the intuitive meaning of this metric- a larger coverage of the objective space is seemed as an indication of both good convergence and spread. However, the conclusions drawn from this metric rely upon the selection of the reference point, as is show in fig. 3.18 (b) and (c). Nevertheless an anti-ideal point can be identified in many problems, which seems as a good option to fix such a reference. On the other hand, the computational complexity of the S-metric is clearly a drawback of this metric, although some efforts for finding optimal algorithms of calculus have been proposed in [56, 213, 212].

The S-metric has been used not only as quality measure but as the base for archiving procedures [106] and indicator-based MOEA [45, 144].

ϵ-Indicators

The concept of ϵ-dominance exposed in sec. 3.3.3 is the base of the binary and unary ϵ-indicators. The former, that operates on two approximation fronts A and B, is defined as

$$I_\epsilon(A, B) = \inf_{\epsilon \in \mathbb{R}} \{\forall \mathbf{x}^2 \in B \exists \mathbf{x}^1 \in A : \mathbf{x}^1 \succ_\epsilon \mathbf{x}^2\} \tag{3.5}$$

while the latter is based on a reference set R such that $I_\epsilon(A, R)$ defines the unary indicator of A [241].

3.4 Uncertainty Handling in EMO

Although the great importance of uncertainty handling in decision-making, the efforts devoted so far to this topic in EMO are still sparse.

The presence of uncertainty hinders the ability of algorithms -and humans- to distinguish what is *certainly* best amidst two or more options. In consequence the core of EA, which is to give more chances to the most fitted alternatives, might be misled due to the difficulty -or impossibility- to know the real quality of the alternatives.

The strategies to cope with uncertainty in EA are oriented, according to [93], toward four categories of problems, viz.:

Noise: it is characterized by a noisy fitness function that can be modeled as $F(\mathbf{x}) + \delta$, where the first addend is the invariant part and the second the noisy factor, which is often assumed to be a gaussian noise. In this type of problem the objective function returns a different value every time the function is called with the same argument.

Robustness: corresponds to arguments affected by uncertainty in such a way that the actual value of a nominal argument $\mathbf{x}$ could vary. As happens with noisy fitness functions, this effect can be modeled as a one-to-many function, but where the disturbances belong to the decisional instead of the objective space, i.e. $F(\mathbf{x} + \delta)$. We will return to robustness in the subsequent chapters.

Fitness approximation: comprises all the cases where the fitness is approximated by surrogate models. In this case neither the original function nor the argument is necessarily subject to uncertainty, but the cost of evaluating $F(\mathbf{x})$ compels to approximate $F(\mathbf{x})$ in order to reduce the computational burden. As a result, the analyst may have at hand two functions, making use of the surrogate one for economy sake or of the original one for precision sake.

Time-varying fitness functions: this category enclose those fitness functions that varies in time, i.e. $F(\mathbf{x}, t)$. This is the case, for example, of the payoffs of stocks and shares in the market.

One can guess that the aforementioned categories are not comprehensive, since they do not distinguish between aleatory and epistemic uncertainty, for instance. On the other hand, in real problems these types of problems are often mixed. This does not mean that such a classification is usefulness but there has to be a considerable additional effort to tackle the uncertainty handle in EA and particularly in EMO, from both the theoretical and the algorithmic viewpoints.

The efforts in MOEA design for optimization under uncertainty are shown in table 3.3 and shall be discussed next.

Realm of study	Authors and works
Probabilistic Dominance:	
Optimization with interval fitness value	J. Teich [197]
Optimization with noisy fitness function	E. Hughes [89, 88], J.E. Fieldsen & R.M. Everson [51]
Quality Indicator-Based Procedures:	
Indicator-based optimization	M. Basseur & E. Zitzler [11]
Robust Optimization:	
Optimization with uncertainty propagation	K. Sörensen [194], T. Ray [152], K. Deb & H. Gupta [34], C. Barrico and C.H. Antunes [10]
Info-gap based robust design	D. Lim et al. [124]
Reliability-based optimization	K. Deb et al. [35]
Probability bounding approaches:	
Tolerance interval	S. Martorell et al. [136], J.F. Villanueva et al. [204]
Non-Probabilistic Procedures:	
Optimization with epistemic uncertainty	P. Limbourg [125], P. Limbourg & D.E. Salazar A. [126, 173]

Table 3.3: State of the art in MOEA design for optimization under uncertainty.

3.4.1 Probabilistic Dominance-based Procedures

The first attempts to design MOEA that cope with uncertainty belong precisely to this category, and correspond to the work of E. Hughes [89, 88] and J. Teich [197], published in 2001. Later on the probabilistic dominance was revisited in the work of J.E. Fieldsen & R.M. Everson [51] in 2005.

The studies conducted by E. Hughes [89, 88] assumes the existence of a noisy fitness function, such that each solution has a PDF associated in the objective space. Additionally, the mathematical treatment assumes gaussian distribution because, as the author argues, in real engineering problem the noise is often gaussian. Then, given two solutions to compare, the kernel of the approach is the calculation of the probability that one of them dominates the other, or conversely, the probability of making a wrong decision. Hughes provides formulae to determine such a probability when both the mean and the variance of each distributions are known, and when only the variances are available. The author remarks that a similar mathematical treatment is possible for other types of PDF, but clearly this approach is only applicable if the distribution shape is known and its moments are available.

The approach proposed by J. Teich [197], on the other hand, considers each solution is mapped into a random variable bounded in a *property interval*. The further mathematical development assumes the random variables are uniformly distributed. When two multiple-objective intervals overlap, the probability of one of them dominating the other is calculated and if it is larger than an acceptance threshold, the dominance is accepted.

J.E. Fieldsen & R.M. Everson [51] expands the studies of Hughes considering unknown variance and how to estimate parameters making use of Monte Carlo

procedures.

3.4.2 Quality Indicator-based Procedures

M. Basseur & E. Zitzler proposed in [11] an indicator-based algorithm that handles uncertainty by ranking the individuals in terms of the expected value of its indicator I_ϵ with respect to the rest of the population. More precisely, the ranking strategy assumes that each solution has PDF in the objective space from which a set of samples is available. Then, the indicator is averaged over such samples, reflecting the mean ϵ value of the solution regarding the population. Afterwards, the archiving and the recombination is ruled by the expected indicator-based fitness as usual. In addition, M. Basseur & E. Zitzler tested their approach on both noise and robustness problems with uncertainty propagation, using several selection procedures.

One interest aspect of this study is the difficulty to assess the quality of the approximation fronts in the presence of uncertainty. We will return to this point in sec. 3.4.6.

3.4.3 Robust Optimization Procedures

The search for robust solutions in EMO follows the line proposed in [200] by S. Tsuitsui and A. Ghosh, which consists in assuming the uncertainty associated to the alternatives $\mathbf{x}$ follows a PDF. Thus the objectives values, which are uncertain by the nature of the problem, are represented by its expected value calculated via Monte Carlo as $\sum_i F(\mathbf{x}+\delta_i)$. This approximation, called *effective function* in [200], is the base of the contribution made by T. Ray [152], K. Sörensen [194, 193] and K. Deb & H. Gupta [34].

K. Sörensen assumes the PDF is somehow known while K. Deb & H. Gupta intrinsically assumes an uniform distribution that is sampled by the Latin Hypercube method. Both cases tackle the search for robust solutions by propagating the uncertainty, but while the latter approach constraints the maximal deviation from the expected value attempting to control dispersion, the former does not establish a variability control, although further works propose optimizing the worst case scenario or weighting the importance of the disturbances as $\sum_i w_i F(\mathbf{x}+\delta_i)$ [184, 185]. In contrast, T. Ray employs a three-objective approach with individuals' performance, the mean performance of its neighbours and the standard deviation of its neighbours' performance.

Later, K. Deb et al. [35] tested a robustness related concept to handle uncertainty: the reliability-based optimization. Concurrently D. Lim et al. [124] proposed a bi-objective scheme for robust design, where the performance of the alternatives constitutes one objective and a function of robustness constitutes the other. This approach is inspired in the info-gap models and aims at maximizing the allowed degradation in the performance for a fixed horizon of uncertainty in the decisional space.

A propagating approach that also takes into account the size of the domain propagated is that introduced by C. Barrico and C.H. Antunes [10]. This approach bears resemblance to the view adopted by K. Deb & H. Gupta [34] in the formulation of robustness, but extends it through the notion of degree of robustness, which make reference to the maximal integer coefficient by which the

original domain can be multiplied -enlarging it- without incurring in a deviation greater than a pre-specified value in the objective space.

3.4.4 Probability bounding approaches

Recently, S. Martorell et al. [136] and J.F. Villanueva et al. [204] have proposed the use tolerance intervals to estimate the upper bounds of binding intervals for the uncertainty quantities, within wich some probability mass is enclosed. The multiple-objective optimization is thereby performed based on such upper bounds comparisons. This approach is brand new regarding EMO and have been applied to dependability problems. Nevertheless, the idea of binding a mass of probability for making-decisions in not new, but have been applied in non-evolutionary approaches e.g. using fractiles [143].

3.4.5 Non-Probabilistic Procedures

The main drawback of probabilistic procedures is that, since the rely upon the assumption of a prior knowledge of the PDF associated with the involved uncertainty, if this information is not available or is not sufficient, the applicability of the methods is therefore compromised. Another aspect that must be considered is the nature of the uncertainty. Epistemic and aleatory uncertainty require a different treatment, and the principle of insufficient reason or maximal entropy [103], that amounts to assume a uniform distribution when nothing points out to a different PDF, may be misleading as underly many authors (cf. e.g. [47, 13]).

Non-probabilistic MOEA that handle uncertainty were proposed by P. Limbourg in [125] and by P. Limbourg & D.E. Salazar A. in [126, 173]. The approaches aims at handling epistemic uncertainty that arises from imprecision in measurements or fitness approximation and that is expressed in property intervals. However, unlike Teich's approach, these approaches do not assume a PDF inside the intervals. The algorithm proposed by P. Limbourg & D.E. Salazar A. in 2005 [126, 173] is summarized in sec. 4.2.

3.4.6 Metrics and Quality Assessment under Uncertainty

Representing an assessing uncertain approximation fronts is an open problem that has been tackled resorting to different approaches in [51, 11, 126, 173]. J.E. Fieldsen & R.M. Everson [51] represent the front using the expected values and then calculate the probability that the obtained front dominates a solution by drawing samples. M. Basseur & E. Zitzler [11] proposed to classify the quality of different approximation fronts (A) by calculating the front with lower expected value of the ϵ-indicator regarding a reference front R, $E(I_\epsilon(A, R))$ that is statistically different (using a Mann-Whitney test with 5% significance level). In addition these authors proposed an adaptation of the empirical attainment function (cf. section 3.3.5) to the uncertain case, which allows to build a map of the objective space and the frequency by which each subregion of such a space is dominated.

As we shall see later on, a complementary approach was introduced by P. Limbourg & D.E. Salazar A. [126, 173] based on the hypervolume metric (see sec. 4.2), thereby characterizing the quality of the approximations by $[\underline{I_H}(A), \overline{I_H}(A)]$. It must be noticed that the MOEA proposed by P. Limbourg & D.E. Salazar

A. and by M. Basseur & E. Zitzler maximize a metric; therefore if the quality is measured by the same metric they maximize it is logical to expect they beat other alternative methods. Nevertheless, the former approach does it only partially, since it recourses to the hypervolume only for density reasons. Anyhow it would be interesting to test the quality of each approach using different metrics.

3.5 Summary of the chapter

This chapter has treated the different aspects of MCDM and EMO that should be known to cope with MCDM problems under uncertainty, where the EA play a central role as solving procedures.

As the ability to solve MCDM problems using EMO depends on an adequate understanding of the decision-making problem as well as a suitable implementation of MOEA, we have highlighted the following issues along this chapter:

- *Formulation of the decision-making problem:* is the problem one of single-criterion or belongs to a multiple-criteria view of the reality. If so, what are its elements and what type of solutions optimize it?
- *General characteristic of the solving MOEA:* what is the selected search engine? Why? What are its parameters and sampling operators? How are the solutions represented?
- *Particular characteristic of the solving MOEA:* is the search engine suitable to sample the decisional space? How are the alternatives raked? How is the information stored (archiving)? How is the MOEA quality to be assessed (metrics)?

Afterwards, the state of the art in uncertainty handling with MOEA has been surveyed. Specifically, the following points have been treated:

- *Types of problems:* what are the different classes of uncertainty handling that exist in the EA world.
- *State of the art:* what are the existing proposals in uncertainty handling in EMO and their characteristics.
- *Metrics and uncertainty:* how researchers have been tackled the problem of assessing the quality of MOEA that handle uncertainty. Current limitations.

The body of knowledge brought so far represents the theoretical foundation of the analytic perspective introduced in the next chapter for the Analysis of Uncertainty and Robustness in Evolutionary Optimization (AUREO).

Chapter 4

Analysis of Uncertainty and Robustness in Evolutionary Optimization (AUREO)

> *"...c'est une vérité très certaine que, lorsqu'il n'est pas en notre pouvoir de discerner les plus vraies opinions, nous devons suivre les plus probables; et même qu'encore que nous ne remarquions point davantage de probabilité aux unes qu'aux autres, nous devons néanmoins nous déterminer à quelques unes, et les considérer après, non plus comme douteuses en tant qu'elles se rapportent à la pratique, mais comme très vraies et très certaines, à cause que la raison qui nous y a fait déterminer se trouve telle."*
>
> *Discours de la méthode*
> by René Descartes (1596 - 1650).

In previews chapters we have discussed relevant facets of uncertainty, decision making and evolutionary optimization almost separately. In this chapter, we strive to give a complete picture of how these topics relate to each other and what can be done in order to cope with uncertainty in EMO-based MCDM. Within this view, the idea of robustness is investigated in terms of its conceptual meaning and its relevancy as criterion. Afterwards, these reflections are wholly embraced in an information-based perspective for the *Analysis of Uncertainty and Robustness in Evolutionary Optimization* (AUREO) for EMO-based MCDM under uncertainty.

4.1 An insight into uncertainty in decision-making

4.1.1 Uncertainty induces comparisons of sets

To account for the effect of uncertainty in decision-making, let us introduce a new space $\mathbf{K}$ (*kappa*) into the conceptualization of the whole decisional process. This space $\mathbf{K}$ denotes the cosmos ($\mathit{Κοσμος}$) and by extension it covers all the real states that the variables of interest may adopt. The universe of discourse $\mathbf{X}$ of any particular problem is a subset of $\mathbf{K}$. Let Γ (from $\mathit{Γνωσις}$) be a function

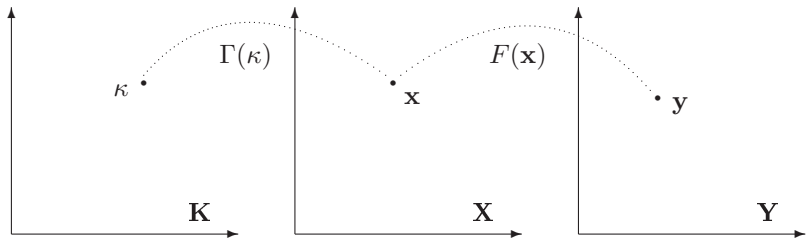


Figure 4.1: Different spaces of decisional processes.

that maps any particular state of reality κ from $\mathbf{K}$ onto $\mathbf{X}$. Thus, the decisional process comprises three spaces as is shown in fig. 4.1, viz. reality $\mathbf{K}$, the space of decision vectors $\mathbf{X}$ and the space of attributes or objectives $\mathbf{Y}$.

As a point of departure, let us assume an ideal state of complete and perfect information such that uncertainty no longer exists, neither in $\mathbf{X}$ nor in $\mathbf{Y}$. In such a situation Γ is an identity function, i.e. any possible state of affairs κ has a unique image in $\mathbf{X}$. Under this ideal scenario the DM has the power to know reality in great detail as well as the ability to designate and to implement any feasible setting of the decision variables. Different criteria are possible to define $F(\mathbf{x})$ but since $\mathbf{X}$ is perfectly related to $\mathbf{Y}$ and vice versa, due to the initial condition of no uncertainty, the discernment of alternatives is based on crisp-points comparisons. We shall see next how in real circumstances only coarse-points and set comparisons are possible.

There are two fundamentals situations that prevent Γ from being a bijective mapping onto a continuous space. The first one amounts to uncertainty whereas the second one to preferences. Decisional domain $\mathbf{X}$ is in fact a space of *nominal vectors* and this is partially due to the inability of DM to discern things beyond some thresholds. The external limitations imposed by our imperfect measurement devices and procedures, the scarce range of perception of our senses, cultural paradigms or a genuine state of mind in which two or more things are indifferent, amongst others reasons, may account for such thresholds. This yields a discrete space of decision $\mathbf{X}$ where each crisp vector $\mathbf{x}$ is a nominal representation of a subset of $\mathbf{K}$ as is depicted in fig. 4.2. Notice that should further level of discernment be practicable, DM's preferences establish the granularity of $\mathbf{X}$.

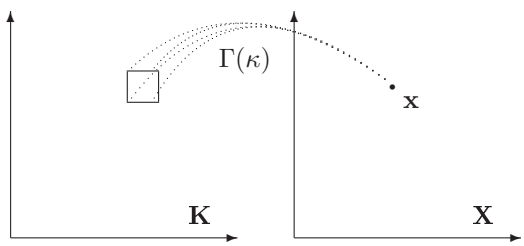


Figure 4.2: Coarse mapping from reality onto decisional space caused by: a) blunt preferences and b) aleatory uncertainty.

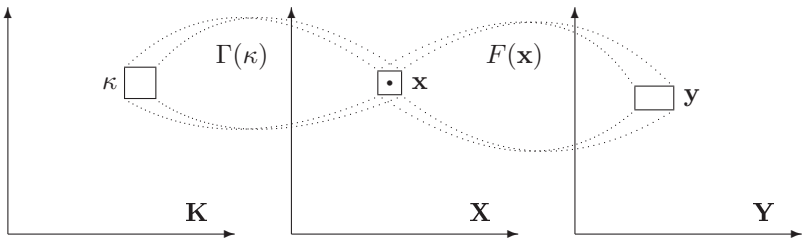


Figure 4.3: Propagation of aleatory uncertainty from reality to the decision space and then to the objective space. Notice the subset in $\mathbf{X}$ attached to the nominal point $\mathbf{x}$.

Aleatory uncertainty has the same effect sketched in fig. 4.2 but for reasons fundamentally different from those argued before. In aleatory uncertainty each decision vector $\mathbf{x}$ is a nominal representation of a subset of $\mathbf{K}$ due to the inherent variability of the system under consideration. This identification is by no means wished by the DM and cannot be attributed to their preferences. Since the DM nominate vectors in $\mathbf{X}$, aleatory uncertainty imposes an important obstacle on decision-making since now any vector $\mathbf{x}$ might represent a variety of realization of reality. Uncertainty should be therefore propagated from $\mathbf{K}$ to $\mathbf{X}$ and then to $\mathbf{Y}$. As a result, any nominal state of affair in $\mathbf{X}$ is associated with a subset of vectors in $\mathbf{Y}$ instead of a single crisp vector $\mathbf{y}$, as is shown in fig. 4.3

Epistemic uncertainty arises from imperfection in operations on information (cf. table 2.2), in particular on data collecting. Thus any particular state of affair κ of reality is identified to a subset of $\mathbf{X}$. The characteristics of such subsets depend on the quality and quantity of the data collected. Hence, as the procedures for information gathering are refined or repeated, the subsets associated with any particular κ change. Consequently, when uncertainty is propagated, a set of objective vectors in $\mathbf{Y}$, instead of a single one, represents a specific state of affair in $\mathbf{K}$ (see fig. 4.4).

Difficulties carried out by epistemic uncertainty are evident, as sketched in fig. 4.4 different sets like $\mathcal{X}_1$ and $\mathcal{X}_2$ may represent the same real state of affair κ. Later when those sets are propagated, κ is assessed in the objective space as $\mathcal{Y}_1$ and $\mathcal{Y}_2$ respectively. Both epistemic and aleatory uncertainty produce sets

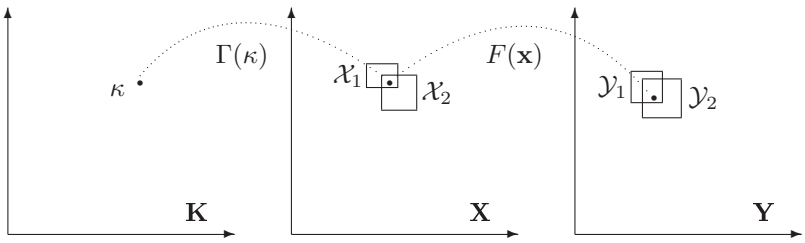


Figure 4.4: Effect of epistemic uncertainty in decision-making: a particular state of affair κ is associated to $\mathcal{X}_1$ and $\mathcal{X}_2$ and assessed in the objective space as $\mathcal{Y}_1$ and $\mathcal{Y}_2$ respectively.

into the space $\mathbf{Y}$ but while the sets generated by the latter type of uncertainty represent different values that the attributes can take, the former type induce sets that contain only one (unknown) true value each. Moreover, the two types can coexist –and often do– in some problems, generating a situation where the connections between the three aforementioned spaces are established in terms of sets instead of crisp points.

4.1.2 Criteria to compare sets

The question that follows the previous discussion is how to make decisions by comparing sets. The answer is far from being trivial and the problem turns out more difficult in EMO as the Pareto ranking methods entail a comparison of k intervals instead of single objectives' values for each alternative to consider. In order to tackle this problem, in this section we shall adopt the Cartesian principle of breaking up the problem into smaller pieces, by analyzing first single-dimension comparisons and afterwards extending the adopted principles to k-dimensional multiple-objective problems.

Two elements must be considered before setting the strategy for uncertainty handling upon which the comparisons of alternatives and consequently the EMO ranking procedure will rely, namely:

The decision-making problem: The characteristics of the problem to solve influence directly the policy on uncertainty handling and its implementation. For instance, discrete and continuous problems often require a different treatment. We shall exemplify this point along the discussion.

The available information: As we stated earlier on in sec. 2.1.1, uncertainty is in our view the result of the sum of imperfections (lack) of information about the elements related to the system of concern. From a decision-making viewpoint, the previous definition means that all the information available (namely the type of uncertainties, their sources, their impacts and the framework adopted for representing the uncertainty) should be considered, otherwise we are introducing more uncertainty into the problem. This may seem to be in contradiction to the spirit of EA, that tries to take advantage of powerful heuristic to design all-purpose algorithms, leaving aside many specificities of the mathematical problems. However, what the previous assertion means is that the EA designers should propose flexible algorithms that can be easily adapted to the peculiarities of the problem or a battery of algorithms suited for each class of uncertainty handling problem, instead of searching for a universal procedure.

At the same time, the ranking procedure should address these two questions in the presence of uncertainty:

1. How to determine dominance?
2. How to assess density?

Let us consider the different types of decision-making problems according to the amount of available information, beginning from single objective comparisons and then extending the analysis to k-dimensional comparisons. As we have mentioned so far in this chapter, decision-making in the presence of uncertainty

implies that every alternative in $\mathbf{X}$ is associated with a set in the objective space $\mathbf{Y}$. Such sets can contain a unique unknown value that corresponds to the true value of $F(\mathbf{x})$ or may denote the set of possible realizations of $\mathbf{y}$ due to the stochasticity associated to the function $F(\mathbf{x})$ or to the argument $\mathbf{x}$.

In our information perspective, the unknown 'unique value' scenario is considered the assumption of minimal prior information, and consequently it is to be assumed in absence of contrary evidence. To explain what motivates this assumption, suppose actor $\mathbf{a}$ only lets actor $\mathbf{b}$ to know that a set is associated with the outcome of $F(\mathbf{x})$; thus $\mathbf{b}$ has to admit that whether epistemic, aleatory or a mixture of uncertainties induced the set. However, if $\mathbf{b}$ assumes a priori that the set is pure aleatory or a mixture of uncertainty types, $\mathbf{b}$ implicity accepts that $\mathbf{a}$ has handled much more information formerly than if assumes the set to be pure epistemic, because one can obtain a set in the presence of epistemic uncertainty after a single evaluation of $F(\mathbf{x})$ (e.g. one measure), but it is not necessarily possible in the case of mixed uncertainty and is definitely impossible in the case of pure aleatory uncertainty.

On the other hand, notice that the least amount of information that is necessary to define a set is its membership function, so if we combine it with the assumption of minimal prior information, we can define the first class of problems to consider: *comparison of sets with no extra information*. Concomitantly, any additional piece of information promotes the problem to our next class: *comparison of sets with extra information*. Once the type of problem is identified, the next step consists in defining a policy to discriminate between intervals. One common procedure is to reduce the problem to crisp-points comparisons of some representative points of the set. More elaborated procedures that depend on more information compare many points. The aforementioned classes are analyzed in detail in the next sections.

4.1.3 Comparison of sets with no extra information

Let $\mathcal{A} = F(\mathbf{x}_1)$ and $\mathcal{B} = F(\mathbf{x}_2)$ denote two closed sets in the objective space $\mathbf{Y}$. Although sets can be defined in many ways, some conditions are desirable in order to be able to make decisions without additional information. In particular, to compare two discrete sets, both of them should comprise a countable collection of non-infinite numbers in $\mathbb{R}$, otherwise their extreme values, upon which the non-probabilistic criteria are formulated, cannot be identified. Likewise, continuous sets require both of them to be closed intervals.

From the definition of the problem, it is possible to know if $\mathcal{A}$ and $\mathcal{B}$ are discrete or continuous sets; however, for the sake of simplicity let $\mathcal{A} = [\underline{a}, \overline{a}]$ and $\mathcal{B} = [\underline{b}, \overline{b}]$ denote in the context of this discussion the intervals (continuous case) or the minimal convex hull (discrete case) of $F(\mathbf{x}_1)$ and $F(\mathbf{x}_2)$. From the uncertainty analysis the analyst can determine if each interval contains a unique unknown value, which is the case of pure epistemic uncertainty, or if on the contrary, some other assertions are plausible, as may be about the likelihood of the unique value to be placed in a particular subset of the interval or the possibility of different outcomes associated to the same decision vector $\mathbf{x}$. As was said before, the latter scenario entails extra knowledge and belongs to another category, and therefore shall be treated in the next subsection.

Since the real values in $\mathcal{A}$ and $\mathcal{B}$ are unknown, to choose one of those intervals when they overlap each other implies the risk of a wrong decision. According

to the DM's attitude towards risk, the following policies are considered:

Insufficient reason: also known as Laplace principle or criterion, it states that in absence of additional information, there is no reason to suppose that a particular possibility is more probable than the others, so it amounts to assume an uniform distribution inside the intervals (the principle of maximum entropy [103] relies upon the same assumption). Afterwards, the expected value or centroid of the interval is used as the representative value to make decisions.

It is very easy to implement this criterion by simply calculating the arithmetic mean of the interval. In fact, Teich [197] considered it as an alternative to his proposal (see sec. 3.4.1).

However, Laplace principle is often subject of criticisms since, on the one hand the expected value criterion may lead to incoherent conclusions (like the Ellsberg paradox) and on the other hand, this principle is not considered suitable for making decision under risk (cf. e.g. [47, 13]). It does not mean that this criterion is no longer valid, but that the analyst should use it with care, since the employment of average tends to underestimate losses and therefore risks.

Worst case optimization: also known as Wald criterion, this criterion is consonant with the risk-averse attitude which implies the decision are made on sure facts. Thus the worst possible case, that corresponds to the upper or the lower bound of the intervals when minimizing or maximizing respectively, serves as the representative point for the sake of comparisons.

Best case optimization: in contrast with the previous criterion, this principle corresponds to a risk-taker attitude that seeks the maximum possible payoff in any situation. Therefore it takes the lower or the upper bound of the intervals when minimizing or maximizing respectively.

Hurwicz criterion: this principle proposed by L. Hurwicz decides on a weighted average of the best and the worst case of the form $\alpha \cdot \overline{x} + (1-\alpha) \cdot \underline{x}$ with $\alpha \in [0,1]$. This principle shows a great versatility, thereby in a maximization case, $\alpha = 1$ corresponds to the best case optimization, $\alpha = 0$ to the worst case, $\alpha = 0.5$ yields the Laplace principle and any other value allows the DM to express their concern about the occurrence of a bad or a good payoff. However, the selection of α could be difficult in practice.

Maximal regret minimization (minimum regret): also known as Savage criterion; this principle aims at minimizing the maximal loss the DM can incur as a result of a wrong choice. Table 4.1 exemplifies the way this criterion is implemented in discrete problems. As only three scenarios can happen (1, 2 or 3), notice that the regret for every alternative (A, B and C) is calculated as the difference between the current cost and lower existence cost for every scenario. Then, the alternative that returns the minimal maximum regret (alternative A) is preferred.

Continuous intervals, however, entail infinite scenarios. For example, every value belonging $\mathcal{A} = [\underline{a}, \overline{a}]$ is possible in principle and its occurrence constitutes an scenario. The same is valid for $\mathcal{B} = [\underline{b}, \overline{b}]$. Now let us suppose a maximization case where the DM chose $\mathcal{B}$. In the case of a wrong

	Scenario 1		Scenario 2		Scenario 3	
Alternative	Cost	Regret	Cost	Regret	Cost	Regret
A	2322	345	2104	0	2615	**1072**
B	4220	**2243**	3847	1743	1543	0
C	1977	0	3485	1381	5151	**3608**

Table 4.1: Example of the minimum regret criterion for discrete problems: Cost matrix with regret values (maximal regret of each alternative in bold).

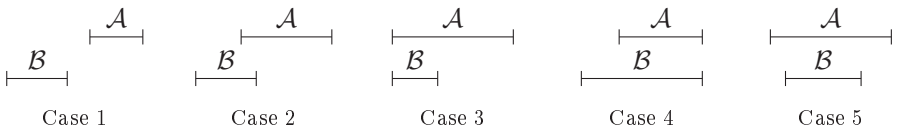


Figure 4.5: Five possible cases of intervals comparisons in MOEA ranking under uncertainty.

decision, the maximal loss is reached when the real value of $\mathcal{B}$ is $\underline{b}$ and the real value of $\mathcal{A}$ is $\overline{a}$, and its value amounts to $\overline{a}-\underline{b}$. Conversely, the maximum loss when choosing $\mathcal{A}$ amounts to $\overline{b}-\underline{a}$. Therefore, according to this criterion the interval associated with the lower loss should be preferred.

Now let us analyze what would happen if we compare two intervals in a MOEA ranking procedure using the precedent criteria. Fig. 4.5 shows five possible configurations of the intervals in a single-dimensional objective space. Case 1 is a risk-free case of sure dominance, whereas the others entail the risk of a wrong decision. Table 4.2 summarizes the results of applying the previous criteria to classify the five cases depicted in fig. 4.5.

The conclusions drawn by the first four criteria are straightforward. In the minimum regret criterion, cases 2 to 4 stand since $\underline{a} \geq \underline{b}$ and $\overline{a} \geq \overline{b}$, thus $\overline{a}-\underline{b} \geq \overline{a}-\underline{a} \geq \overline{b}-\underline{a}$ given that $\overline{a} \geq \overline{b}$. Case 5, however, can lead to different results depending on the relative position of the intervals under consideration.

The preceding definitions can be formulated mathematically as follows:

Definition 8 (SOI-dominance) *A single-objective interval (SOI) vector* $\mathcal{A} = [\underline{a}, \overline{a}]$ *dominates* $\mathcal{B} = [\underline{b}, \overline{b}]$*, denoted* $\mathcal{A} \succ_{SOI} \mathcal{B}$*, iff (maximization case):*

$$\mathcal{A} \succ_{SOI} \mathcal{B} \Leftrightarrow \begin{cases} \overline{a} > \overline{b} & \textit{Best Case} \\ \underline{a} > \underline{b} & \textit{Worst Case} \\ (\overline{a}+\underline{a}) > (\overline{b}+\underline{b}) & \textit{Laplace} \\ \alpha \cdot (\overline{a}-\overline{b}) > (1-\alpha) \cdot (\underline{b}-\underline{a}),\ \alpha \in [0,1] & \textit{Hurwicz} \\ \overline{a}-\underline{b} > \overline{b}+\underline{a} & \textit{Min Regret} \end{cases} \tag{4.1}$$

If neither $\mathcal{A} \succ_{SOI} \mathcal{B}$ *nor* $\mathcal{B} \succ_{SOI} \mathcal{A}$ *then both intervals are said to be incomparable:* $\mathcal{A}||\mathcal{B}$.

Let us consider now the expansion to a two-dimensional case in order to propose a k-dimensional rule of classification. Fig. 4.6 presents six possible configurations of intervals in a bi-dimensional space, whereas the classifications for those cases are reported in table 4.3.

Cases (Fig. 4.5)	Best case Criterion	Worst case Criterion	Laplace Criterion	Hurwicz Criterion	Min Regret Criterion
Case 1	$\mathcal{A} \succ \mathcal{B}$	$\mathcal{A} \succ \mathcal{B}$	$\mathcal{A} \succ \mathcal{B}$	$\mathcal{A} \succ \mathcal{B}$	$\mathcal{A} \succ \mathcal{B}$
Case 2	$\mathcal{A} \succ \mathcal{B}$	$\mathcal{A} \succ \mathcal{B}$	$\mathcal{A} \succ \mathcal{B}$	$\mathcal{A} \succ \mathcal{B}$	$\mathcal{A} \succ \mathcal{B}$
Case 3	$\mathcal{A} \succ \mathcal{B}$	$\mathcal{A}\|\|\mathcal{B}$	$\mathcal{A} \succ \mathcal{B}$	$\mathcal{A} \succ \mathcal{B}$	$\mathcal{A} \succ \mathcal{B}$
Case 4	$\mathcal{A}\|\|\mathcal{B}$	$\mathcal{A} \succ \mathcal{B}$	$\mathcal{A} \succ \mathcal{B}$	$\mathcal{A} \succ \mathcal{B}$	$\mathcal{A} \succ \mathcal{B}$
Case 5	$\mathcal{A} \succ \mathcal{B}$	$\mathcal{A} \succ \mathcal{B}$	$\mathcal{A}\{\succ, \|\|, \prec\}\mathcal{B}$	$\mathcal{A}\{\succ, \|\|, \prec\}\mathcal{B}$	$\mathcal{A}\{\succ, \|\|, \prec\}\mathcal{B}$

Table 4.2: Conclusions drawn by five non-probabilistic criteria in MOEA ranking under uncertainty with intervals (maximization case): Depending on the relation between intervals (Min Regret, Laplace) or on α (Hurwicz), different conclusions may arise ($\{\succ, ||, \prec\}$).

Cases (Fig. 4.6)	Best case Criterion	Worst case Criterion	Laplace Criterion	Hurwicz Criterion	Min Regret Criterion
Case 1	$\mathcal{A} \succ \mathcal{B}$	$\mathcal{A} \succ \mathcal{B}$	$\mathcal{A} \succ \mathcal{B}$	$\mathcal{A} \succ \mathcal{B}$	$\mathcal{A} \succ \mathcal{B}$
Case 2	$\mathcal{A}\|\|\mathcal{B}$	$\mathcal{A}\|\|\mathcal{B}$	$\mathcal{A}\|\|\mathcal{B}$	$\mathcal{A}\|\|\mathcal{B}$	$\mathcal{A}\{\succ, \|\|, \prec\}\mathcal{B}$
Case 3	$\mathcal{A} \succ \mathcal{B}$	$\mathcal{A}\|\|\mathcal{B}$	$\mathcal{A}\{\succ, \|\|, \prec\}\mathcal{B}$	$\mathcal{A}\{\succ, \|\|, \prec\}\mathcal{B}$	$\mathcal{A}\{\succ, \|\|, \prec\}\mathcal{B}$
Case 4	$\mathcal{A} \succ \mathcal{B}$	$\mathcal{A} \succ \mathcal{B}$	$\mathcal{A} \succ \mathcal{B}$	$\mathcal{A} \succ \mathcal{B}$	$\mathcal{A} \succ \mathcal{B}$
Case 5	$\mathcal{A} \succ \mathcal{B}$	$\mathcal{A}\|\|\mathcal{B}$	$\mathcal{A}\{\succ, \|\|, \prec\}\mathcal{B}$	$\mathcal{A}\{\succ, \|\|, \prec\}\mathcal{B}$	$\mathcal{A}\{\succ, \|\|, \prec\}\mathcal{B}$
Case 6	$\mathcal{A} \succ \mathcal{B}$	$\mathcal{A} \prec \mathcal{B}$	$\mathcal{A}\{\succ, \|\|, \prec\}\mathcal{B}$	$\mathcal{A}\{\succ, \|\|, \prec\}\mathcal{B}$	$\mathcal{A}\{\succ, \|\|, \prec\}\mathcal{B}$

Table 4.3: Conclusions drawn by five non-probabilistic criteria in MOEA ranking with multiple-objective intervals (maximization case): Depending on the relation between intervals (Min Regret, Laplace) or on α (Hurwicz), different conclusions may arise ($\{\succ, ||, \prec\}$).

For the first four criteria under consideration the same principle introduced in [126, 173] for multiple-objective intervals (see def. 16) can be applied, leading to the conclusions shown in table 4.3. In general we have:

Definition 9 (MOI-dominance) *A multiple-objective interval (MOI) vector* $\mathcal{A} = \{\mathcal{A}_i = [\underline{a}_i, \overline{a}_i] : i \in 1, \ldots, k\}$ *dominates* $\mathcal{B} = \{\mathcal{B}_i = [\underline{b}_i, \overline{b}_i] : i \in 1, \ldots, k\}$, *denoted* $\mathcal{A} \succ_{MOI} \mathcal{B}$, *iff (maximization case):*

$$\mathcal{A} \succ_{MOI} \mathcal{B} \Leftrightarrow \begin{cases} \mathcal{A}_i \succ_{SOI} \mathcal{B}_i \vee \mathcal{A}_i || \mathcal{B}_i : \forall i \in \{1, \ldots, k\} \\ \exists j \in \{1, \ldots, k\} : \mathcal{A}_j \succ_{SOI} \mathcal{B}_j \end{cases} \tag{4.2}$$

If neither $\mathcal{A} \succ_{MOI} \mathcal{B}$ *nor* $\mathcal{B} \succ_{MOI} \mathcal{A}$ *then both intervals are said to be incomparable:* $\mathcal{A}||\mathcal{B}$.

The application of the previous definition to MOEA ranking based on the first criteria studied is straightforward. Additionally, density control is possible utilizing the ordinary procedures described in sec. 3.3 using the representative values established by those criteria. However, the minimization of the maximal regret allows different implementations:

Direct implementation: Consists in using definition 9.

Weighted sum of regrets: The preceding option compels one interval to minimize the regret in every dimension to dominate. Instead of doing that, the overall regret can be considered based on a L_p norm of the form

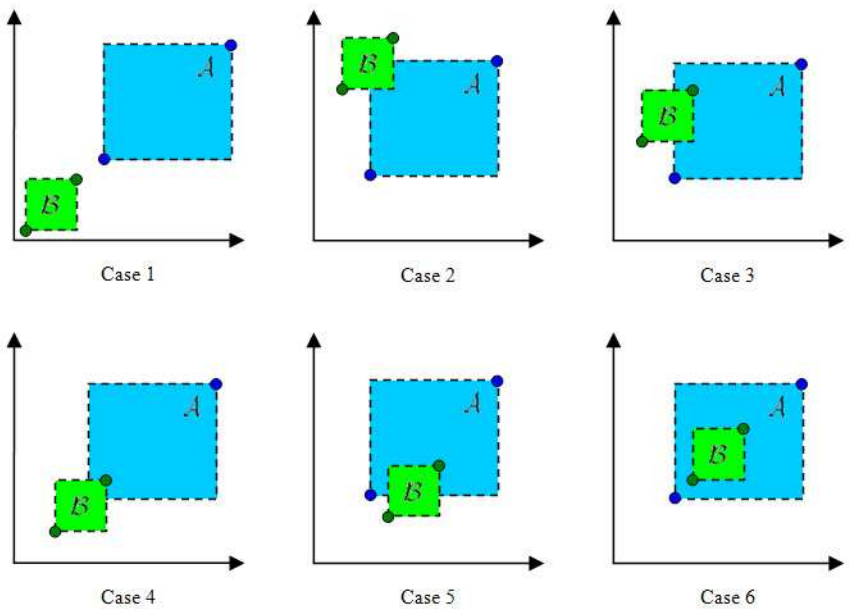


Figure 4.6: Examples of possible cases of bi-dimensional intervals comparisons in MOEA ranking.

$\left(\sum_i^k w_i|x_i|^p\right)^{\frac{1}{p}}$ where x corresponds to the amount of the regret caused by a wrong decision, and $w_i \geq 0$ are the weights given to each objective. In other words, the weighted distance between the best case of one alternative and the worst case of the other would serve to decide the dominance. It is evident that the results rely on the value of p and w_i and thus the procedure resorts to the introduction of new parameters. However, how difficult the definition of such parameters is depends on the problem itself.

Minimize maximal regret of k-objectives: A non-parametric version of the preceding alternative is to use the infinite norm, i.e. to minimize the maximal difference in a single objective between the worst and the best cases of the two alternatives under consideration respectively. This alternative is attractive not only because is simpler than the previous one, but for its definition bears an interesting resemblance to the additive ϵ-indicator $I_{\epsilon+}$(A,B) (see sec. 3.4.6 and eq. 3.5), for A and B being two intervals.

4.1.4 Comparison of sets with extra information

When the analyst handles additional evidence about the sets in $\mathbf{Y}$, such an evidence can take a possibilistic or probabilistic form that should be used to construct the ranking procedure of the MOEA. The type of uncertainty and the theoretical framework adopted to represent it shall determine the ranking procedures.

As we stated before in sec. 4.1.3, sets in $\mathbf{Y}$ can denote a unique unknown

value imprecisely defined or the range of possible realizations of $F(\mathbf{x})$ for a given $\mathbf{x}$. Let us consider again $\mathcal{A} = F(\mathbf{x}_1)$ and $\mathcal{B} = F(\mathbf{x}_2)$. In the presence of additional information, the former case entails that the analyst and/or DM believe that some subsets of $\mathcal{A}$ and $\mathcal{B}$ are more likely to contain the real values of $F(\mathbf{x})$ than others. This belief can take, for instance, the form of precise or imprecise probabilities. Ranking procedures should not neglect this information.

The case of many realizations of $F(\mathbf{x})$, on the other hand, may be the result of uncertainty in $\mathbf{X}$ (robustness problem) or may come from the evaluation of the objective function (noisy, time-varying or surrogate function) as we mentioned before in sec. 3.4. While robustness shall be treated in detail in sec. 4.3, we focus here on the commonalities of all categories. As now at least more than one value contained in $\mathcal{A}$ can happen, the same for $\mathcal{B}$, the information of the possibility (by fuzzy logic) or the probability (by precise or imprecise probability theories) of their occurrence should be the base to compare sets. Therefore, the issues to address in EMO are how can dominance be defined regarding the theories of uncertainty surveyed in chapter 2 and how to interpret density.

Ranking sets with probability theory

The easiest way to compare sets described by a PDF or a CDF is to use their expected values as representative values. This is in general the approach adopted so far in EMO (see sec. 3.4.1). Robust optimization as described in sec. 3.4.3 is also based on expected values comparisons. Besides, variance or other variability measures are often contemplated as an objective to be minimized in robust optimization and in some general decision-making problems under uncertainty. These measures are not treated here, but we will turn to them in sec. 4.3.

More sophisticated concept of probabilistic dominance has been proposed in the literature. S. Graves [65] surveys and compares six different probabilistic criteria for comparing uncertain alternatives. Amongst them stochastic dominance is considered the most general and less restrictive criterion:

Definition 10 (First-degree stochastic dominance [65, 80]) *let* $\mathbf{y} = F(\mathbf{x})$ *be an increasing objective function to be maximized* ($F'(\mathbf{x}) > 0$)[1]*, and let* $\mathtt{F}(\mathbf{y} = F(\mathbf{x}))$ *be the CDF of* $\mathbf{y}$*. An alternative* $\mathbf{x}_1$ first-degree stochastically dominates $\mathbf{x}_2$ *if* $\mathtt{F}(\mathbf{y} = F(\mathbf{x}_1)) \leq \mathtt{F}(\mathbf{y} = F(\mathbf{x}_2))$ *for all* $\mathbf{y}$ *and the inequality is strict for at least one* $\mathbf{y}$*.*

If the CDF of $F(\mathbf{x}_1)$ intersects the CDF of $F(\mathbf{x}_2)$, the first-order stochastic dominance is not conclusive. In this case, higher degree stochastic dominance can be used to decide:

Definition 11 (Second-degree stochastic dominance [65, 80]) *let* $\mathbf{y} = F(\mathbf{x})$ *be an increasing objective function to be maximized* ($F'(\mathbf{x}) > 0$) *to satisfy a risk-averse DM* ($F''(\mathbf{x}) < 0$)*, and let* $\mathtt{F}(\mathbf{y} = F(\mathbf{x}))$ *be the CDF of* $\mathbf{y}$*. An alternative* $\mathbf{x}_1$ second-degree stochastically dominates $\mathbf{x}_2$ *if* $\int_{-\infty}^{z} \mathtt{F}(\mathbf{y} = F(\mathbf{x}_1))d\mathbf{y} \leq \int_{-\infty}^{z} \mathtt{F}(\mathbf{y} = F(\mathbf{x}_2))d\mathbf{y}$ *for all* z *over* $\mathbf{y}$ *and the inequality is strict for at least one* $\mathbf{y}$*.*

[1] The original formulation makes reference to a utility function instead of an objective function (see [65, 80]).

The main limitations of the preceding concepts are the conditions ($F'(\mathbf{x}) > 0$) and ($F''(\mathbf{x}) < 0$). Other concepts of probabilistic dominance, like *Almost Stochastic Dominance* or the *Mean-Gini Analysis* can be found in [65]. The idea of using such dominance criteria is to enhance the power of discernment, since using only the expected value may lead to conclude incomparability in many comparisons of some problems. Nevertheless, in general probabilistic dominance criteria rely upon certain conditions that cannot hold in many problems. On the other hand, stochastic and almost stochastic dominance require the CDF of the alternatives in the objective space which can be very tricky and/or prohibited to determine. That may explain why EA designers recourse to mean value comparisons approximated by Monte Carlo methods. Besides, there is no simple way to measure the density of a Pareto approximation front beyond the use of single values like the first central moments. However if the conditions for their use can be satisfied, it may be worthwhile to build the ranking procedure on stochastic dominance concepts combined with a suitable density assessment perhaps based on first central moments.

Ranking sets with fuzzy logic

Fuzzy logic are often advocated in decision-making, and thus in EA and EMO, as a way of representing smoothly the transition of preferences between what is considered 'good' or 'acceptable' and what is 'rejectable' or 'bad'. This way, composite fuzzy indexes make possible to identify optimal solutions (see e.g. [231]) amongst crisp alternatives.

A pretty different problem comes up when the uncertainty in $\mathbf{Y}$ takes the form of fuzzy sets. When each objective has a range of possible values, criteria for ranking the dominance and for assessing the density should be established. The first obstacle is how to determine if a fuzzy set overcomes another regarding the objective values. Once more let us consider $\mathcal{A} = F(\mathbf{x}_1)$ and $\mathcal{B} = F(\mathbf{x}_2)$ but now as fuzzy sets. Membership functions should be considered in order to verify dominance, unless support sets $\mathcal{A}_0 = [\underline{\mathcal{A}_\alpha}, \overline{\mathcal{A}_\alpha}]$ and $\mathcal{B}_0 = [\underline{\mathcal{B}_\alpha}, \overline{\mathcal{B}_\alpha}]$ do not overlap. In such a case the conclusion is straightforward.

For example, amongst the different ranking options proposed in the literature (see [77, 57] and references therein), a simple and sound way for defining dominance is given by P. Fortemps and M. Roubens [57], based on the degree of compensation between the two fuzzy numbers:

Definition 12 (Ranking based on area compensation [57]) *Let $\mathcal{A}$ and $\mathcal{B}$ be two fuzzy numbers (maximization case), thus*

$$\begin{array}{lll} \mathcal{A} \succ \mathcal{B} & \textit{iff} & C(\mathcal{A} \geq \mathcal{B}) > 0 \\ \mathcal{A} \succeq \mathcal{B} & \textit{iff} & C(\mathcal{A} \geq \mathcal{B}) \geq 0 \end{array}$$

where the functional C can be defined in terms of a defuzzification function $\mathcal{F}$ as:

$$C(\mathcal{A} \geq \mathcal{B}) = \mathcal{F}(\mathcal{A}) - \mathcal{F}(\mathcal{B}) \tag{4.3}$$

$$\mathcal{F}(\mathcal{A}) = \frac{1}{2}\int_0^1 (\underline{\mathcal{A}_\alpha} + \overline{\mathcal{A}_\alpha})d\alpha \tag{4.4}$$

A similar alternative proposed by R. Groetschel and W. Voxman establishes that:

Definition 13 (Groetschel and Voxman ranking (cf. [25])) *Let $\mathcal{A}$ and $\mathcal{B}$ be two fuzzy numbers (maximization case), thus*

$$\mathcal{A} \succeq \mathcal{B} \Leftrightarrow \int_0^1 \alpha(\underline{\mathcal{A}_\alpha} + \overline{\mathcal{A}_\alpha})d\alpha \geq \int_0^1 \alpha(\underline{\mathcal{B}_\alpha} + \overline{\mathcal{B}_\alpha})d\alpha$$

where the strict preference is defined with a strict inequality.

The preceding definitions are very simple to calculate and constitute a complete ranking of fuzzy numbers. They can be used to construct a MOEA ranking procedure by simply using the MOI relation to determine dominance (def. 9) after equating the SOI relation (def. 8) to the ranking methods just described. However, the density assessment remain unsolved with this approach.

MOEA designers can use the existing approaches for density estimation by determining some representative points of the fuzzy sets in $\mathbf{Y}$. For instance, C. Carlsson and R. Fullér define the:

Definition 14 (Crisp possibilistic mean value [25]) *For a fuzzy number $\mathcal{A}$, the crisp possibilistic mean value is defined as*

$$\bar{M}(\mathcal{A}) = \frac{M_*(\mathcal{A}) + M^*(\mathcal{A})}{2} \tag{4.5}$$

where

$$\begin{aligned} M_*(\mathcal{A}) &= 2\int_0^1 \alpha\underline{\mathcal{A}_\alpha}d\alpha \\ M^*(\mathcal{A}) &= 2\int_0^1 \alpha\overline{\mathcal{A}_\alpha}d\alpha \end{aligned}$$

The crisp possibilistic mean value averages the arithmetic means of the different α-cuts weighted by the α values. Thereby the means closer to the core set have more importance to determine the reference point. Density assessment for the ranking and archiving procedures can be based on this measure or others of the like. Moreover, the whole MOEA can rely upon this kind of measures.

Ranking sets with imprecise probability theories

As imprecise probability theories, like those surveyed in chapter 2, work with upper and lower bounds for probabilities and PDF of events instead of precise values, those representative values used in MOEA to handle uncertainty like central moments become intervals. In consequence it is impossible to reduce the density assessment to an application of a procedure based on single-point comparison as the current MOEA do. On the other hand, criteria of dominance are inevitably referred to interval comparisons.

To the best of our knowledge, there is no practical or theoretical study of the use of imprecise probability theories in EMO, and the possibility of designing MOEA based on these frameworks has not even glimpsed yet, perhaps due to the practical difficulty to handle the calculations in an accurate and efficient way using EA.

The issue of dominance with imprecise probabilities in MOEA can be based on theoretical and practical results obtained regarding single-objective optimization. The theories of imprecise probability contemplate decision-making rules

for ranking alternatives: see [209, 199, 127] for theoretical issues of dominance criteria using imprecise probabilities and [23, 22] for applications in black-box single-objective optimization based on p-boxes. Other approaches rely upon a generalization of the expected value to imprecise probabilities, making decisions based on this quantities [224]. All these contribution should be adapted to define a multiple-objective optimization. Density assessment, on the other hand, would require to extent the concept of density by means of probabilistic reasoning, like it could be a probabilistic crowding distance in NSGA-II or the probabilistic density of the hyperboxes of an hypergrid (see sections 3.3.2 and 3.3.4). Some of these ideas shall be discussed in the next section.

4.2 Some Theoretical and Algorithmic Proposals for Uncertainty Handling in EMO

In this section we bring some proposals that can be used to tackle some of the problems posed by the presence of uncertainty in EMO-based MCDM. Such proposals are based on the topics discussed so far in this thesis. The first proposal is a MOEA that evolves using imprecise values (intervals) in the objective space. Afterwards, two suggestions about the use of the minimal regret criterion and the problem of density assessment with uncertain objectives are presented.

4.2.1 An Optimization Algorithm for Imprecise Multiple-Objective Problem Functions: IP-MOEA

In sec. 4.1.3 the problem of comparing sets without extra information was studied by analyzing different politics that can be applied to decide between pairs of uncertain objective vectors. Some other politics can be formulated according to the DM attitude regarding risk. In this section we bring an alternative approach to handle interval objective values. This algorithm, named IP-MOEA, was introduced by P. Limbourg & the author in [126, 173]. Some of the contents of this section are quite similar to those previously exposed in sec. 4.1.3, however, the discourse introduced by the authors is presented next to follow the subjacent reasoning.

Fig. 4.7 shows the three possible cases of one-dimensional interval comparison considered and the conclusions drawn according to the introduced operator $>_{IN}$. Type (a) represents the case in which it can be decided for sure that y is greater than y'. Thus, $y >_{IN} y'$ must hold. Type (b) is a much more critical case. If there exists a distribution inside y and y', it needs not necessarily be the same. On the other hand, the only fact certainly known is that the real values are somewhere inside the intervals. So it might well be that y' is greater than y. Yet, most rational decision makers would prefer y to y' since choosing y entails a better chance to get a better value (maximization case). Incomparability $y||y'$ is an error free possibility, but since it could lead to a large number of neutral conclusions and both the worst and the best cases of y dominate their counterparts in y', the authors tilted in favour of accepting the dominance. Finally, type (c) is the case of incomparability. One interval encloses the other. Without extra knowledge about the underlying distributions -if any- or the DM's risk attitude, incomparability should hold. The one-dimensional rule becomes:

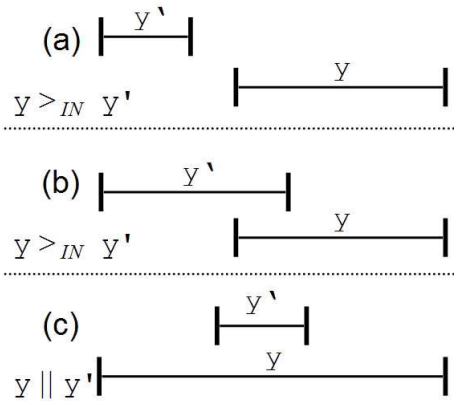


Figure 4.7: Examples of intervals comparisons (maximization case) [126, 173].

Definition 15 (IN-dominance) *An interval number* $y = [\underline{y}, \overline{y}]$ *dominates* $y' = [\underline{y'}, \overline{y'}]$*, denoted* $y >_{IN} y'$*, iff* $\underline{y} \geq \underline{y'} \wedge \overline{y} \geq \overline{y'} \wedge y \neq y'$.

Afterwards the authors extended the definition above to a multi-dimensional rule that establishes:

Definition 16 (IP-dominance) *An* k*-dimensional interval objective vector* $y = \{y_i = [\underline{y}_i, \overline{y}_i] : i \in 1, \ldots, k\}$ *dominates* $y' = \{y'_i = [\underline{y'}_i, \overline{y'}_i] : i \in 1, \ldots, k\}$*, denoted* $y >_{IP} y'$*, iff* $\forall i \in \{1, \ldots, k\} : y_i >_{IN} y'_i \vee y_i || y'_i$ *and* $\exists j \in \{1, \ldots, k\} : y_j >_{IN} y'_j$.

Figure 4.8 brings the four types of comparisons that might take place when applying definition 16. Case (a) corresponds to the case when both the worst and the best case of y' is dominated by their counterparts in y. In contrast, in case (b) neither the worst nor the best case of one interval dominate their counterparts, which corresponds to the typical case of incomparability or indifference in decision-making. Case (c), on the other hand, is also a case of incomparability but for another reasons. In this case there is no clear domination in neither objective, so it is not possible to choose without resorting to an additional criterion. Finally this situation is relaxed in case (d) where at least $y_i >_{IN} y'$ holds in one dimension, thus allowing to accept domination, although this decision could be considered as a case of incomparability by many DM.

Density is controlled using the hypervolume metric as in [45], but in this case both the certain contribution ($\underline{I_H}(y, P) = \underline{I_H}(P) - \underline{I_H}(P - y)$) and the plausible ($\overline{I_H}(y, P) = \overline{I_H}(P) - \overline{I_H}(P - y)$) contribution of each interval with respect to the whole population P is used as a selection criterion for further discriminations. The NSGA-II were then adapted to comply the following rules:

Selection criteria: Rank the population according to the following criteria

1. Compare individuals using $>_{IP}$

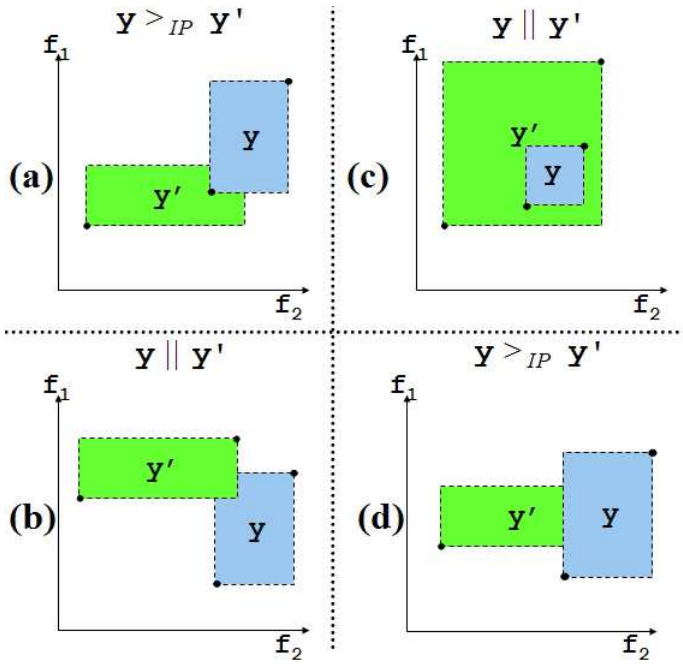


Figure 4.8: Examples of multiple-objective intervals comparisons (maximization case) [126, 173].

2. Compare hypervolume measures $[\underline{I_H}(y,P), \overline{I_H}(y,P)]$ using $>_{IN}$ criterion.

3. Choose at random.

The algorithm were tested against the NSGA-II operating on the mean value of the intervals. The quality of the results were measured with the hypervolume metric, adapted for the imprecise case (calculating $\underline{I_H}$) and $\overline{I_H}$), using four well known test functions (ZDT1, ZDT2, ZDT4, ZDT6 [236]) tailored for intervals instead of precise results, as shown in table 4.4. Notice that the intervals produced are reproducible, since no noise is introduced but an imprecise evaluation is emulated.

Both NSGA-II and IP-MOEA were run for 100 generations with a population size of 20 with a crossover probability of 0.9 and a mutation rate of 0.1. The number of decision variables was fixed on 30 for all test problems and the results were compared using the hypervolume metric. The reference point was fixed to the minimal objective values for all solution sets and the objective values were scaled to [0,1].

Table 4.5 reports the hypervolume of the worst-case and best-case of the approximation front after 100 generations. Besides the median and interquartile range (IQR) and the probability of the null hypothesis, $p(0)$ that both medians are equal is given. Significance testing was done by a Kruskal-Wallis test on the equality of medians. The results clearly support the thesis that imprecision propagation leads to better algorithm performance. Indeed, on all four test functions, both best and worst case hypervolume $(\underline{I_H}(y,P), \overline{I_H}(y,P))$ are higher and the null hypothesis can be safely rejected ($p(0) < 0.001$ in all tests). From

Imprecise-valued Functions $ZDT_I i, \quad i \in \{1, 2, 4, 6\}$

Functional representation

$$\underline{ZDT_I i(\mathbf{x})} = [\underline{ZDT_I i(\mathbf{x})}, \overline{ZDT_I i(\mathbf{x})}]$$
$$\overline{ZDT_I i(\mathbf{x})} = \max\{ZDTi(\mathbf{x}), ZDTi(\mathbf{x}) + (1/2)\sin\left(20\pi \sum_i x_i\right)\}$$
$$\underline{ZDT_I i(\mathbf{x})} = \min\{ZDTi(\mathbf{x}), ZDTi(\mathbf{x}) + (1/2)\sin\left(10\pi \sum_i x_i\right)\}$$

ZDT1

Functional representation

$$f_1(\mathbf{x}) = x_1$$
$$g(\mathbf{x}) = 1 + \frac{9}{n-1}\sum_{i=2}^{n} x_i$$
$$h(f_1(\mathbf{x}), g(\mathbf{x})) = 1 - \sqrt{f_1(\mathbf{x})/g(\mathbf{x})}$$
$$\mathbf{x} = (x_1, \ldots, x_{30}), \quad x_i \in [0, 1]$$

ZDT2

Functional representation

$$f_1(\mathbf{x}) = x_1$$
$$g(\mathbf{x}) = 1 + \frac{9}{n-1}\sum_{i=2}^{n} x_i$$
$$h(f_1(\mathbf{x}), g(\mathbf{x})) = 1 - (f_1(\mathbf{x})/g(\mathbf{x}))^2$$
$$\mathbf{x} = (x_1, \ldots, x_{30}), \quad x_i \in [0, 1]$$

ZDT4

Functional representation

$$f_1(\mathbf{x}) = x_1$$
$$g(\mathbf{x}) = 1 + 10(n-1) + \sum_{i=2}^{n}\left(x_i^2 - 10\cos(4\pi x_i)\right)$$
$$h(f_1(\mathbf{x}), g(\mathbf{x})) = 1 - \sqrt{f_1(\mathbf{x})/g(\mathbf{x})}$$
$$\mathbf{x} = (x_1, \ldots, x_{10}), \quad x_1 \in [0, 1], \quad x_2, \ldots, x_{10} \in [-5, 5]$$

ZDT6

Functional representation

$$f_1(\mathbf{x}) = 1 - \exp(-4x_1)\sin^6(6\pi x_i)$$
$$g(\mathbf{x}) = 1 + 9\left(\frac{1}{n-1}\sum_{i=2}^{n} x_i\right)^{1/4}$$
$$h(f_1(\mathbf{x}), g(\mathbf{x})) = 1 - (f_1(\mathbf{x})/g(\mathbf{x}))^2$$
$$\mathbf{x} = (x_1, \ldots, x_{10}), \quad x_i \in [0, 1]$$

Table 4.4: Imprecise-valued test functions ZDT_I (adapted from ZDT1, ZDT2, ZDT4, ZDT6) [126, 173].

this results, the authors concluded that the performance of the IP-MOEA was at least as good as the mean-assuming NSGA-II, and the IP-MOEA is a sound option to cope with imprecise values, i.e., values with epistemic uncertainty.

4.2.2 The ϵ-indicator revisited through the concept of minimal regret

The concept of ϵ-dominance (see sec. 3.3.3) has received an increasing attention in the EMO community due to the beneficial effects of its incorporation in MOEA. Hence, the ϵ-dominance improves the convergence capabilities of MOEA [117, 116] and, at the same time, provides a base for the construction of indicators to assess the quality of the approximation sets (see sec. 3.3.5).

In particular, the ϵ-indicator has been used to improve or design MOEA with and without uncertainty-handling. Regarding the former case, it was explored

Function	IP-MOEA		NSGA-II		
ZDT_I1	*Median*	*IQR*	*Median*	*IQR*	$p(0)$
Worst Case	0.7139	1.15E-02	0.7064	1.57E-02	9.60E-07
Best Case	0.8991	1.59E-02	0.8899	1.97E-02	3.23E-05
ZDT_I2	*Median*	*IQR*	*Median*	*IQR*	$p(0)$
Worst Case	0.6528	1.91E-02	0.6324	2.32E-02	7.25E-12
Best Case	0.8219	2.50E-02	0.7954	2.74E-02	1.04E-11
ZDT_I4	*Median*	*IQR*	*Median*	*IQR*	$p(0)$
Worst Case	0.7311	3.96E-02	0.7017	4.45E-02	1.04E-09
Best Case	0.9062	4.37E-02	0.8728	5.58E-02	1.74E-09
ZDT_I6	*Median*	*IQR*	*Median*	*IQR*	$p(0)$
Worst Case	0.4759	3.87E-02	0.4531	3.99E-02	7.12E-07
Best Case	0.7591	5.26E-02	0.7248	4.88E-02	2.92E-06

Table 4.5: Hypervolume values after 100 generations for imprecise-valued test functions ZDT_I [126, 173].

by M. Basseur & E. Zitzler in [11], where the authors propose to calculate the quality of each one of the elements that compose the approximation front using the expected value of the ϵ-indicator. Then, those candidates that show the best values are kept whereas the others are eliminated. Evidently, the proposed procedure entails the possibility of having or assuming some PDF or histogram for each uncertain candidate.

However, a non-probabilistic use of the ϵ-indicator seems appealing just by realizing that the additive ϵ-dominance and the concept of minimal regret have points in common. Indeed, let us consider the two cases sketched in fig. 4.9. Recall the ϵ-indicator $I_\epsilon(A, B)$ 'equals the minimum factor ϵ such that any objective vector in B is ϵ-dominated by at least one objective vector in A' [241, pg. 122]. Thus, if we apply the ϵ-indicator between two points (fig. 4.9 left) to determine which one is better, as in [11], the indicator will favour the point that is closer to the ideal point of the current set (red point), i.e. the solution with smaller ϵ. Notice that this comparison is equivalent to that of the size of the loss or regret of selecting the wrong alternative (see sec. 4.1.3), i.e. if $\mathbf{y}_1$ is chosen instead of $\mathbf{y}_2$, the loss in quality regarding f_1 is null or rather negative, which implies a gain. in contrast, the loss in regarding f_2 is positive and has size ϵ_2. Conversely, if we select $\mathbf{y}_2$ the regret is ϵ_1. Thus the option to choose is that that produces the minimal regret, i.e. the one with smaller ϵ, something that matches the result of using the ϵ-indicator.

Likewise, if we extend the analysis to compare one point with a set (fig. 4.9 right), the effect is the same. Nevertheless, if we want to compare two sets, the results of the ϵ-indicator and the minimal regret criterion are no longer equivalent since the latter amounts to decide in terms of the distance between the ideal points of the sets under consideration, whereas the former does not.

Then if we consider uncertain objective vectors, the regret is still an expression of the distance between the best and the worst cases of distinct alternatives, but now the distance is multidimensional and thus it can be calculated in several forms, as we discussed in sec. 4.1.3. In particular, if the distance is assessed with L_∞ norm, the regret between two alternatives (see fig. 4.10 left) and the

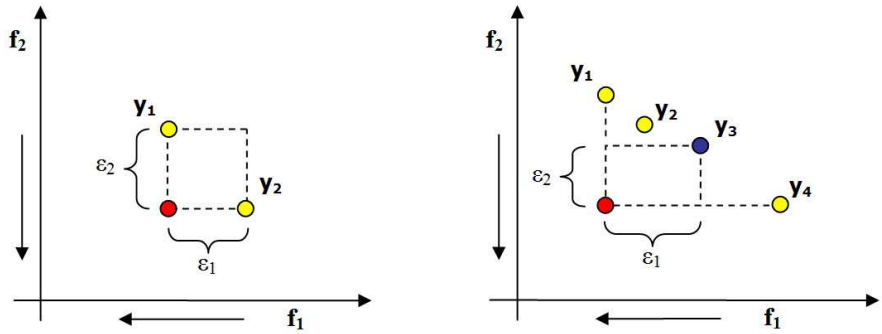


Figure 4.9: Comparison of the ϵ-dominance and the minimal regret concepts (minimization case): between two points (left) and between one point and a set (right).

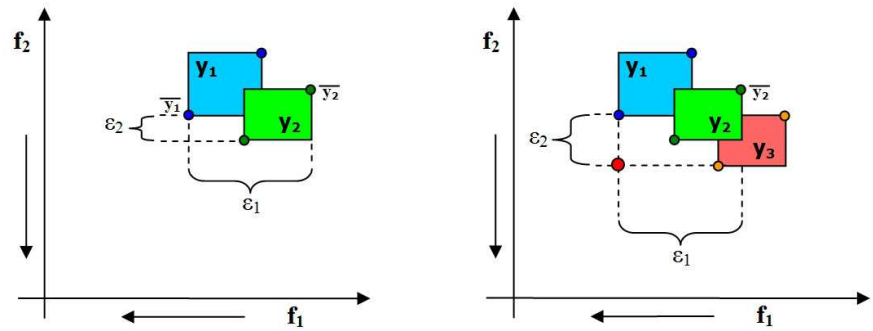


Figure 4.10: Comparison of the ϵ-dominance and the minimal regret concepts for uncertain objective vectors (minimization case): between two vectors (left) and amongst several vectors (right).

application of the additive ϵ-indicator bears the same results. Likewise, the regret of selecting one single vector instead of a set is equivalent to the maximal distance between the worst case of such vector and the ideal point of the set (see fig. 4.10 right). Hence, the construction of an additive ϵ-indicator compliant MOEA, similar to M. Basseur & E. Zitzler algorithm [11] but for handling non-probabilistic intervals, seems worthwhile.

4.2.3 A probabilistic interpretation of the hypercube's density

In sec. 4.1.4 it was hinted the possibility of reinterpreting the hypercube's density through a probabilistic reasoning. We offer a better description of this concept here, although the study of the implementation of this concept is not covered in this thesis.

Fig. 4.11 sketches a section of an hypergrid with the uncertain image $\mathbf{y_1}, \mathbf{y_2}, \mathbf{y_3}$ of three decision vectors. The centroid of each interval, which co-

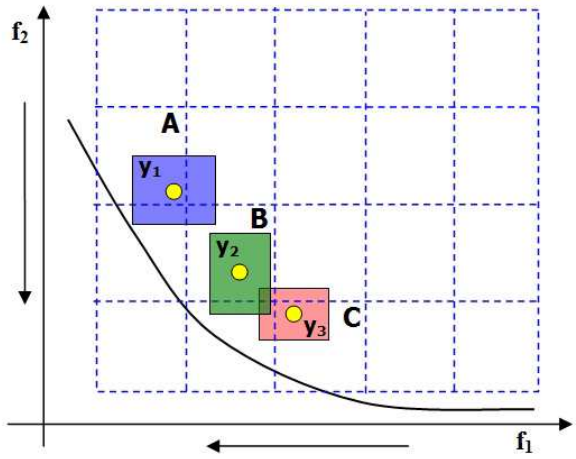


Figure 4.11: Example of an hypergrid with decision vectors with uncertain objectives.

incides with the expected value assuming e.g. a uniform or a normal PDF is represented as well.

Let us consider first the case of crisp objective values using the centroids. The ordinary procedure for truncating (recall `Truncation`(P_A) in sec. 3.3.2) should search for the most crowded hypercube, but since hypercubes **A**, **B** and **C** have density one, if the archive is to be reduced in one element, the algorithm should select at random the hypercube from which the the point should be deleted (in this example $\mathbf{y_1}$ and $\mathbf{y_3}$ are not treated as extreme points to be kept). Now if we resort the use of expected values when considering uncertainty, the result will be exactly the same although hypercube **B** might contain up to three points while the others may not. Hence, it seems logical to assign to hypercube **B** a higher density.

One possible strategy is to make density d_p a N_p-dimensional vector instead of a scalar, where N_p is the number of candidates points to remain in the archive. For example, the density of hypercube **A** in fig. 4.11 is

$$d_p(\mathbf{A}) = (P(|\mathbf{A}| = 1), P(|\mathbf{A}| = 2), \ldots, P(|\mathbf{A}| = N_p)) \tag{4.6}$$

with $N_p = 3$. $P(|\mathbf{A}| = 1)$ is the probability of hypercube **A** having one point, which coincides with $P(\mathbf{y_1} \in \mathbf{A})$. Since only $\mathbf{y_1}$ overlaps **A**, $P(|\mathbf{A}| = 2) = 0$ and $P(|\mathbf{A}| = 3) = 0$. Likewise, the density vector $d_p(\mathbf{C})$ of hypercube **C** is $(P(\mathbf{y_3} \in \mathbf{B}), 0, 0)$.

Let us consider now hypercube **B**. The density vector $d_p(\mathbf{B})$ has non-null components, viz.:

$$P(|\mathbf{B}| = 1) = \sum_{i=1}^{N_p} \left(P(\mathbf{y_i} \in \mathbf{B}) \prod_{j=1; j \neq i}^{N_p} P(\mathbf{y_j} \notin \mathbf{B}) \right) \tag{4.7}$$

$$P(|\mathbf{B}| = 2) = \sum_{i=1}^{N_p} \left(P(\mathbf{y_i} \notin \mathbf{B}) \prod_{j=1; j \neq i}^{N_p} P(\mathbf{y_j} \in \mathbf{B}) \right) \tag{4.8}$$

$$P(|\mathbf{B}| = 3) \quad = \quad \prod_{i=1}^{N_p} P(\mathbf{y_i} \notin \mathbf{B}) \tag{4.9}$$

Notice that no conditional probabilities are considered since the events are independent.

The CDF of $\mathbf{y_1}, \mathbf{y_2}, \mathbf{y_3}$ should be determine (analytically) or estimated (via simulation) to calculate the preceding probabilities. The use of CDF allows to handle imprecise probabilities as well, by just considering an upper and a lower bounds for density. On the other hand, the complete calculation of all components could be cumbersome since it entails determining 2^{N_p} terms. Hence for practical reasons N_p could be fixed e.g. on the maximal density of the crisp case (i.e. using centroids). Moreover, if one point is to be eliminated, the maximal probability that is relevant is that of having $N_p - 1$ points in a hypercube. In our example it means to calculate the combinatorics up to 2 points inside the same hypercube.

The procedure for truncating should consist in deleting that point whose contribution to the density is higher. One possibility is to emulate the ordinary procedure of the hypergrid, i.e. the first step is to select the hypercube of reference which is the hypercube with the higher non-null probability of having as many points as possible minus one. In our example it clearly corresponds to hypercube $\mathbf{B}$, due to $P(|\mathbf{B}| = 2) > 0$ while the counterparts of $\mathbf{A}$ and $\mathbf{C}$ are null. Afterwards, the point to be deleted should be that point whose elimination brings the maximal reduction in the probability under consideration. In our example the point to be eliminated is determined as

$$\mathbf{y_i} = \arg\min_i \{P(|\mathbf{B}| = 2 | \mathbf{y_i} \notin \mathbf{B})\}$$

As it was stated at the beginning of this section, the study of the implementation of this idea is out of the scope of this thesis. Other strategies are still possible and the issue is open to discussion. Nevertheless, we want to remark that the density assessment of uncertain approximation fronts has been scarcely studied. Consequently any additional effort about this topic, whether practical, theoretical or simply speculative, as the concept of probabilistic density introduced here, is not only worthwhile but necessary.

4.3 Robustness: from concepts to classes

In previous sections we have focused on possible alternatives to cope with uncertainty based on the characteristics of the outcome regarding space $\mathbf{Y}$. In this section we extend the analysis to consider the decisional space $\mathbf{X}$ as well through the concept of robustness.

This notion admits different interpretations that we strive for encompassing into the same analysis. In order to do so we shall begin by investigating what people conceptualize as robustness. Then the distinct versions of this notion are organized into a few classes that shall serve afterwards to build EMO-based procedures for MCDM under uncertainty.

4.3.1 What do we mean when we say that something is robust?

There are plenty of formulations of robustness in the literature[2] along with an innumerable quantity of citations evoking the intuitive meaning of this concept, which is, in the words of B. Roy, "*qui résiste à l'à peu près*"[169, pg. 1], something that R. Hites traduces as "*aptitude to resist to 'more or less' or to 'zones of ignorance' in order to protect from regrettable impacts*"[80, pg. 15].

The preceding definition is manifestly generic, and is so because, as R. Hites emphasizes later on in her thesis, "*the concept of robustness is very subjective*"[80, pg. 18]. Subjectivity here is far from meaning arbitrariness, but actually denotes the central role of the DM in defining the sort of impacts that are regrettable.

On the other hand, in an effort to provide a framework to understand robustness, P. Vincke [205] points out that four types of robustness problems derive from the type of system intended to be robust: 1) a decision in a dynamic context, 2) a solution in an optimization problem, 3) a conclusion or 4) a method. Naturally, in all these versions the uncertainty in the decisional space well may be associated with the decision vector or with the environmental parameters or both, i.e. with controllable and/or uncontrollable variables. In this thesis, we are particularly interested in the second category thus the rest of this discussion focuses on this topic; the third and fourth categories shall not be treated hereafter while the first one is briefly mention next.

Robust decisions in a dynamic context

Concerning the first category, J. Rosenhead [164, pg. 8–9] defines robustness as "*the ratio of the number of acceptably performing configurations with which a commitment is compatible, to the total number of acceptably performing configurations*". In other words, given a sequence of actions, the greater the number of available options to take once an action is done the better.

Robustness in optimization

For S. Sayin, "*robustness refers to the ability of a subject to cope well with uncertainties*"[182, pg. 6], a matter that can be tackled in different ways in optimization (see [16] for a comprehensive survey). It is worthwhile to notice that, in practice, it could be the goal of a DM to minimize the variability of an operating system's output by adjusting some input parameters, or simultaneously maximize its performance while hedging against uncertainty by modifying the parameter setting or the system design itself (see e.g. [1]). Both cases are referred to in the operation research and EA literature as *robust optimization* and in the engineering literature as *robust design*.

For example, according to Kouvelis and Yu (cited in [182, pg. 6]) robustness can be achieved by minimizing the maximum deviation from individual criterion best values (like the worst case optimization in sec. 4.1.3). This recours is, however, normally regarded as overly pessimistic so researchers have explored the use of central tendency moments to account for robustness. The easier

[2] A very interesting source of opinions that reflects the distinct facets of robustness is the Robustness Forum: http://www.inescc.pt/∼ewgmcda/Articles.html

fashion is to relate robustness with the expected value, an idea that has been embraced in EA by some researchers like S. Tsutsui and A. Ghosh [200] and M. Sevaux and K. Sörensen [185] (see sec. 3.4.3 for this concept in EMO).

Expected values, however, are not always well enough to characterize a system behaviour in the presence of uncertainty. Hence, the engineering tradition confers a great importance to variance, accounting for robustness e.g. by means of set of ratios of expected value to variance, as was advocated by G. Taguchi [38]. In this context robustness is understood as insensitiveness to input uncertainty. Therefore, mean and variance are treated in different fashions, viz. optimizing the mean while minimizing or constraining the variance, or simply minimizing the variance or mean's deviation from a target value (e.g. [149, 137, 29, 123, 143, 162]). Multiple-objective formulations have been explored as well, using classical and evolutionary algorithms to solve them [66, 132, 93, 15].

The issue that arises from the diversity of ways that robustness is assessed is if all the measures and procedures to find robust optima converge or not on the same results. The answer, as we shall see through the following example, is that they diverge at least partially:

Example 4.1 (Robust optima of a continuous function) *Consider the following function, taken from [194, pg. 72], to be maximized on the interval* $[0, 1]$*:*

$$f(x) = \begin{cases} \exp\left(-2\left(\frac{x-0.1}{0.8}\right)\right)\sqrt{|\sin(5\pi x)|} & 0.4 < x \leq 0.6 \\ \exp\left(-2\left(\frac{x-0.1}{0.8}\right)\right)\sin^6(5\pi x) & \textit{otherwise} \end{cases} \tag{4.10}$$

and suppose that x *is affected by a gaussian noise with* $\sigma = 0.0625$*, so that the expected value of* x *can be estimated as* $\frac{1}{n}\sum_{i=1}^{n} f(x+\delta_i)$*, with* δ_i *a random variable* $N(0, \sigma)$ *and* $n = 1000$*. Figs. 4.12 to 4.14 report the simulation results performed with Crystal Ball® and Microsoft Excel®.*

Fig. 4.12 brings the expected value and the standard deviation of $f(x)$*. As first observation, we should remark that our robust optimum is reached in* $x = 0.1$*, value that diverges from that reported in [194, pg. 73] fig. 5.3. which corresponds to* $x = 0.5$*. Thus, from our experiment, if we used the expected value to assess robustness as suggested in [200, 185], point* $x = 0.1$ *yields the robust optimum. On the contrary, if the variance is to be reduced, i.e. the goal is to minimize variability, point* $x = 0.9$ *becomes the optimal choice.*

On the other hand, if we adopt a bi-objective approach making mean and variance the objectives, the number of optimal solutions becomes four as shown in fig. 4.13. Point $x = 0.5$ *is the closest one to the ideal point so it can be finally selected according to the compromise programming principles (see sec. 3.1.1). On the other hand, if the coefficient of variation (COV) instead of the standard deviation is used to characterize the variability, the set of optimal solutions reduces to two, as shown in fig. 4.14.*

The preceding example evidences how the same generic notion realize in distinct forms in practice, leading to different conclusions. The same situation takes place in discrete and combinatorial problems. For instance, R. Hites [80] reports and introduces in total about ten definitions applicable to define a robust optimum in combinatorial problems. They are built with respect to some of the representative values discussed heretofore in this chapter, and some of them contemplate the introduction of acceptability thresholds that define a region

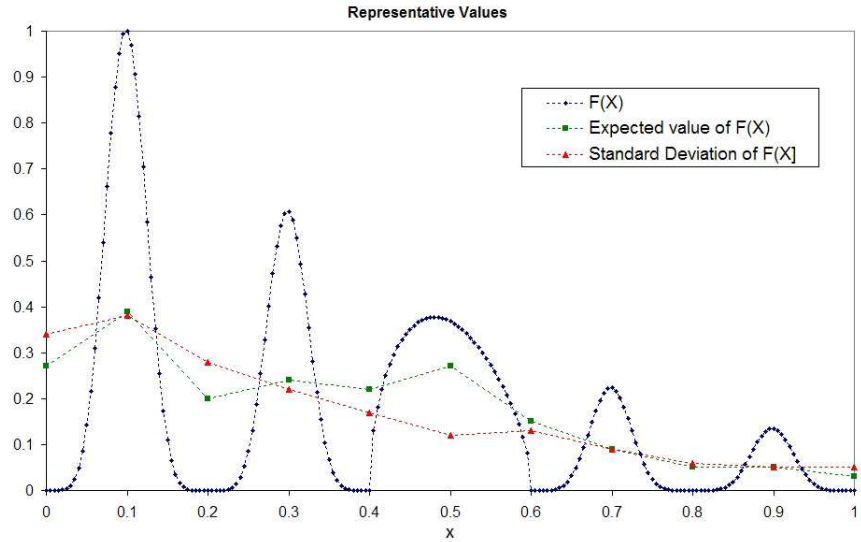


Figure 4.12: Image, expected value and standard deviation of $f(x)$ (example 4.1).

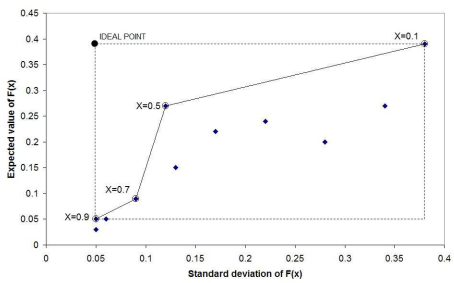


Figure 4.13: Trade-off between standard deviation and mean (example 4.1).

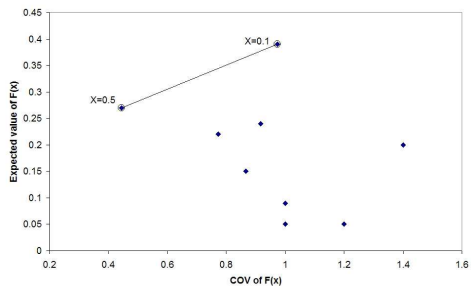


Figure 4.14: Trade-off between COV and mean (example 4.1).

within which the all the possible values that a variable takes should be placed in order to be considered as robust. Naturally, the use of one or another definition may yield in different sets of robust solutions.

In general, the current definitions of robustness can be classified in measures that do not rely upon possibility/probability information about the (finite or infinite) existing scenarios and those that do. To this last group belong most of the measures examined in this discussion. However, the problem of finding a robust optimum or design with the help of imprecise probabilities, evidence theory or fuzzy logic has been scarcely studied. For example in [17] the authors only cite a few applications of fuzzy logic and evidence theory in design optimization

Robustness through the effective domain

The wholly of the previous formulations have in common that robustness is assessed in the objective space $\mathbf{Y}$. In other words, these definitions are concerned about the impact of uncertainty on the performance or quality of candidate solutions or designs. A pretty different form for concerning with robustness focus directly on the decisional space $\mathbf{X}$, assessing the robustness of a nominal vector $\mathbf{x}$ in terms of the subset of values that $\mathbf{x}$ may adopt that abide with the DM's performance requirements. This problem have some similarity to the *constraint satisfaction problem* (see e.g. [20]) as both problems are defined on the decisional space, but the robustness case demands not only the feasibility in terms of the natural constraints of the mathematical problem (recall the condition $G(\mathbf{x}) \leq 0$ in sec. 3.1.1) but normally impose additional constraints of performance. This is called '*effective domain assessment*' by us in [177]. For example, E. Hendrix et al. [76] offer different formulations of robustness in a continuous space, based on the maximization of the L_p with $p \in \{1, 2, \infty\}$ around a centrepoint $\mathbf{x}$, thereby assessing robustness as the value of the norm. This approach to robustness in continuous spaces has been successfully studied in the context of EA by other authors [155, 156] and in EMO by us [158, 178, 179, 176, 180] as we shall see in the next chapters. Notice that this view of robustness also resembles the interpretation given by the info-gap theory (see section 2.2.8).

Robustness measures for discrete domains are also possible as we shall see in the next section. Afterwards, in sec. 4.3.3 we strive for offering a unified view of robustness concerning the many concepts just surveyed.

4.3.2 Robustness concept: A survey

The diversity of meanings and formulations about robustness motivates this fundamental question: *Is it possible to exist consensus around this concept?* If not so, *what are, at least, the elements that people associate with this notion?* This subsection deals with this issue. In order to investigate the aforementioned questions, we prepare a short questionnaire with 14 problems of pairwise comparisons. This questionnaire was sent by email to 15 relevant researchers[3] in decision-making, uncertainty theory and/or robustness. For convenience, the sent email and the questionnaire are reported in appendix . In this section we make general comments about the answers.

Table 4.6 brings the answers of two consulted researchers, Bilal M. Ayyub and Jürgen Branke[4]. The preferences are expressed in terms of the binary relations **strict preference** (P), **weak preference** (Q), **indifference** (I) and **incomparability** (J).

Responses

Answers in table 4.6 show discrepancy in most of the opinions. They are apparently motivated by differences in the features that are considered as repre-

[3]The identities of the non-responding researchers is not revealed here. For those who replied the option of anonymity was offered. None of them took this option, thus their responses are reported with their identities

[4]His preferences were not expressed using the pairwise comparisons operators suggested. Instead, Prof. Branke provided a list of the most robust solution in each case (see appendix A). We chose to translate such preferences as weak relations.

Case	Bilal M. Ayyub's preferences	Jürgen Branke's preferences
0:	AQB (Smaller coefficient of variation)	AQB
1:	AQB (Assuming smaller coefficient of variation and a smaller range)	AQB
2:	BQA (Assuming smaller coefficient of variation)	BQA
3:	AQB (Assuming smaller coefficient of variation)	AQB
4:	AIB (Skewness might have an importance depending on the application domain)	AQB
5:	BQA (Cardinality is smaller)	AQB
6:	AQB (Same cardinality, but smaller range)	AQB
7:	BQA (Cardinality is smaller)	Unclear
8:	BQA (Greater entropy)	AQB
9:	BQA (Greater entropy)	Unclear
10:	AQB (Greater range)	BQA
11:	AIB (Depending on the application domain)	BQA
12:	AQB (Smaller range)	Not sure what the difference is between distributions and fuzzy sets.
13:	BQA (Tighter shape around central value)	Not sure what the difference is between distributions and fuzzy sets.

Table 4.6: Answers to the questionnaire about robustness.

sentative to define robustness by the researchers consulted. In is interested to note that the coincidence of opinions occurs only with continuous quantities represented by PDF, perhaps due to the familiarity of this form of uncertainty representation to our 'traditional' mean-variance ratios for measuring robustness. A more detailed comparison of these results is given in appendix A.

A totally different answer but by no means less relevant was that given by B. Roy (see appendix A), who proposed three definitions of robustness [170], in the same spirit of the effective domain assessment notions recently described, but for discrete domain problems. Let $b \geq 0$ be a reference value established by the DM and let $w \leq \max_{\mathbf{x}} \min_s(f(\mathbf{x}, s))$ be the lower bound of the performance of $f(\mathbf{x})$ on a family of scenarios $s = \{s_i\}$, for which $f^*(\mathbf{x}, s_i)$ is the optimum for scenario s_i. Then a robust solution $\mathbf{x}$ is every solution that maximize the following measure, given in three versions:

Definition 17 ((b,w)-Absolute robustness)

$$r_{bw}(\mathbf{x}) = \begin{cases} 0 & \textit{if} \quad \exists s_i \ : \ f(\mathbf{x}, s_i) < w \\ |\{s_i\}| \ : \ f(\mathbf{x}, s_i) \geq b & \textit{otherwise} \end{cases}$$

with $b = \max_{\mathbf{x}} \min_s(f(\mathbf{x}, s))$. *Thus the robustness is assessed as the cardinality of the family of scenarios that assure the performance abides by the constraints.*

Definition 18 ((b,w)-Absolute deviation)

$$r_{bw}(\mathbf{x}) = \begin{cases} 0 & \textit{if} \quad \exists s_i \ : \ f(\mathbf{x}, s_i) < w \\ |\{s_i\}| \ : \ f^*(\mathbf{x}, s_i) - f(\mathbf{x}, s_i) \geq b & \textit{otherwise} \end{cases}$$

with $b = \min_s \max_{\mathbf{x}} |f^*(\mathbf{x}, s_i) - f(\mathbf{x}, s_i)|$. *Thus the robustness of each solution is the cardinality of the family of scenarios for which the amount of deviation from the optimum does not exceed the boundary after the elimination of those solutions with* $r_{bw}(\mathbf{x}) = 0$.

Definition 19 ((b,w)-Relative deviation)

$$r_{bw}(\mathbf{x}) = \begin{cases} 0 & \text{if} \quad \exists s_i \ : \ f(\mathbf{x}, s_i) < w \\ |\{s_i\}| \ : \ \frac{f^*(\mathbf{x},s_i) - f(\mathbf{x},s_i)}{f^*(\mathbf{x},s_i)} \geq b & \text{otherwise} \end{cases}$$

with $b = \min_s \max_{\mathbf{x}} \frac{f^*(\mathbf{x},s_i) - f(\mathbf{x},s_i)}{f^*(\mathbf{x},s_i)}$. *Likewise, the robustness of each solution is the cardinality of the family of scenarios for which the relative amount of deviation from the optimum does not exceed the boundary after the elimination of those solutions with* $r_{bw}(\mathbf{x}) = 0$.

The results of this survey reinforce our conviction in the central role that information plays in the definition of robustness as we argued in [177]. Indeed, it is not possible to assess robustness in the sense given by B. Roy without a previous knowledge of the performance of the alternatives along the family of scenarios. Likewise, deficiencies in information may hamper the propagation of uncertainty and the consequent analysis of the outcome upon which the robustness is typically assessed in optimization.

In the next section we recourse to the concept of classes in order to give a unified view of robustness in the realm of optimization.

4.3.3 Robustness: many concepts and a few classes

Consider a model for a system's performance, denoted as $F(\mathbf{x})$, where $F : \mathbb{R}^n \to \mathbb{R}^k$ is a mapping from a decisional space $\mathbf{X} \subseteq \mathbb{R}^n, n \geq 1$ onto an attribute space $\mathbf{Y} \subseteq \mathbb{R}^k$ composed by $k \geq 1$ figures of merit of such performance. By 'system' we mean the most generic sense so the preceding definition conveys multiple cases, in such a way that $\mathbf{x} \in \mathbf{X}$ may represent e.g. structural changes to be applied in an open design, a sequence of actions in a planning or simply a setting of controllable parameters of a device. In every case, a realization of $\mathbf{x}$ is deemed as a realization of the system, i.e. a different system.

The performance assessment consists in evaluating $F(\mathbf{x})$, while the optimization of such performance leads to solve a mathematical program of the type

$$\begin{aligned} & Opt\,(F(\mathbf{x})) \\ \text{s.t.:} \quad & \\ & G(\mathbf{x}) \leq 0 \end{aligned} \tag{4.11}$$

where the value of k determines the single or multiple objective character of the optimal solution. $G(\mathbf{x})$ is a vector of constraints that can be formulated as inequality or equality constraints that arise from physical limitations of the system (hard constraints) and desired levels of performance attainment (soft constraints) imposed by the DM.

Consider now a more informative model that takes into account the effect of uncontrollable parameters denoted by a vector $\mathbf{p}$, usually called environmental parameters. Such a model can be expressed as $F(\mathbf{x}, \mathbf{p})$ and now the figures of performance of any particular system are open to change as $\mathbf{p}$ varies. Thus the

concern for robustness appears when the outcome of the system is uncertain due to $\mathbf{x}$ and/or $\mathbf{p}$ are subject to aleatory or epistemic uncertainty or even to both types (see sec. 4.1). Evidently, a uncertain performance is problematic, for instance in optimization, since it hampers the identification of the optimal solutions. Likewise, even if no optimization is pursued, an uncertain outcome can carry violations of hard or soft constraints.

Let us try now to account for robustness departing from the numerous concepts discussed in the previous sections. We formulate robustness in the more general way as:

$$\begin{aligned} & R\left(F(\mathbf{x}), \mathbf{x}, \mathbf{p}, \delta_x, \delta_p, \gamma\right) \\ \text{s.t.:} & \\ & \mathbf{x} \in \mathbf{X}\ ,\ \mathbf{p} \in \mathbf{P} \\ & \underline{\delta_x} \leq \delta_x \leq \overline{\delta_x} \\ & \underline{\delta_p} \leq \delta_p \leq \overline{\delta_p} \end{aligned} \tag{4.12}$$

which can be read as *the robustness of a system determined by its nominal decisional vector* $\mathbf{x}$ *in the decisional space* $\mathbf{X}$*, is a criterion defined in terms of the variation of its performance* $F(\mathbf{x})$*, regarding a target quantity* γ*, and the uncertainty associated with* $\mathbf{x}$*, namely* δ_x*, and its counterpart* δ_p *in the environmental parameters' space* $\mathbf{P}$.

Before analyzing the role played by $\mathbf{p}, \delta_x, \delta_p$ and γ, let us focus on $R(\cdot)$. It seems to us that robustness is a criterion and therefore $R(\cdot)$ is susceptible to be defined in different ways by different DM and even by the same DM in distinct occasions, a fact sufficiently evidenced in the previous sections. Hence eq. 4.12 can be understood as a special criteria class whose realizations partially rely upon the DM. In that sense we agree with R. Hites in recognizing the necessity of having some input from the DM regarding their preferences[5]. This view also entails that robustness can be improved and thus searching for robust systems, namely designs or solutions, means that $R(\mathbf{a}) > R(\mathbf{b}) \Leftrightarrow \mathbf{a} \succ \mathbf{b}$, otherwise robustness is not the unique criterion sought or $R(\mathbf{x})$ is not well defined[6].

It is easy to realize that eq. 4.12 is general enough so it is straightforward to make many of the definitions present in the literature to comply with it, no matter if the problem has continuous or discrete domain. Function $R(\cdot)$ can be defined in many ways, e.g. as a classifier into categories of robustness (like *not robust* and *robust*) with $R : \mathbb{R}^k \to \{0, 1\}$ or as a continuous function with $R : \mathbb{R}^k \to [0, 1]$. The actual problem in practice is to determine the most suitable way to quantify or qualify 'variation' or 'insensitiveness to uncertainty'. For instance, an analyst can hesitate about what is better between variance or interfractile distance. In other words, if $R(\cdot)$ admits different definitions, say $R_1(\mathbf{x})$ and $R_2(\mathbf{x})$, there is an actual concern about if $R_1(\mathbf{a}) > R_1(\mathbf{b}) \Leftrightarrow R_2(\mathbf{a}) > R_2(\mathbf{b})$ holds. It is out of the scope of this thesis to study what measures of robustness amongst those typically used abide by the preceding relation. Nevertheless the matter is not only interesting but important.

For simplicity, the uncertainty associated with $\mathbf{x}$ can be expressed as $\mathbf{x} + \delta_x$ for epistemic and aleatory uncertainty. Thus, in the epistemic case, $\mathbf{x} + \delta_x$ says

[5] According to R. Hites "*The concept of robustness is very subjective, and, in our opinion, a robust approach should allow the decision maker to give input regarding his preferences to find a solution that is satisfactory for him.*"[80, pg. 18-19]

[6] For convenience and without a loss of generality, we formulate $R(\mathbf{x})$ as a function to be maximized. This way, e.g. if robustness makes reference to variance reduction then $R(F(\mathbf{x})) = -\text{var}(F(\mathbf{x}))$.

that the actual value can deviate from the nominal value $\mathbf{x}$ in δ_x units, yielding an interval $[\mathbf{x}+\underline{\delta_x}, \mathbf{x}+\overline{\delta_x}]$. Likewise, for aleatory uncertainty the formulation is similar but with the addition of a PDF that governs the behaviour of $\mathbf{x}$. The same reasoning applies to $\mathbf{p}$.

Now, in the most general way, the search for robustness can be expressed mathematically through a program of the type:

Definition 20 (Robustness-seeking program) *let $F(\mathbf{x})$ be a measure of performance of a system determined by the decisional vector $\mathbf{x}$ and influenced by a vector of environmental parameters $\mathbf{p}$, each of which is subject of uncertainties δ_x and δ_p respectively. Let $G(\cdot)$ be a vector of constraints (inequality or equality) defined for the optimization of $F(\mathbf{x})$, and finally let $I(\cdot)$ be a vector of robustness constraints imposed upon the performance. The optimization of robustness consists in solving the following program*

$$\begin{aligned} & Opt\left(R\left(F(\mathbf{x}), \mathbf{x}, \mathbf{p}, \delta_x, \delta_p, \gamma\right)\right) \\ s.t.: & \\ & \mathbf{x} \in \mathbf{X}\ ,\ \mathbf{p} \in \mathbf{P} \\ & \underline{\delta_x} \leq \delta_x \leq \overline{\delta_x} \\ & \underline{\delta_p} \leq \delta_p \leq \overline{\delta_p} \\ & G(\mathbf{x}, \mathbf{p}, \delta_x, \delta_p) \leq 0 \\ & I(F(\mathbf{x}), \mathbf{x}, \mathbf{p}, \delta_x, \delta_p, \gamma) \leq 0 \end{aligned} \tag{4.13}$$

The amount of information available determines the particular formulation of $R(\cdot)$ and $I(\cdot)$ and therefore the class of robustness-seeking program to solve. There are fundamentally two classes of definitions of robustness with well defined solving procedures and a third mixed class that combines the reasoning of the preceding classes in order to solve problems with the least amount of information [177].

Class 1: Uncertainty propagating programs

This class, also known as *stochastic programming*, is characterized by a suitable description of $\underline{\delta_x} \leq \delta_x \leq \overline{\delta_x}$ and $\underline{\delta_p} \leq \delta_p \leq \overline{\delta_p}$ in such a way that the uncertainty can be propagated through $F(\mathbf{x})$. If the uncertainty is aleatory, δ_x and δ_p have associated PDF. For instance, δ_x could be a normally distributed number $N(0, \sigma)$ such that $\underline{\delta_x} = -\infty$ and $\overline{\delta_x} = \infty$. On the contrary, if the uncertainty is epistemic, $\underline{\delta_x}$ and $\overline{\delta_x}$ are finite scalars.

Robustness is typically assessed using mean and variance as representative values within this class. Table 4.7 shows some of the formulations present in the literature. All these formulations rely upon classical probability theory or upon non probabilistic reasoning. For instance, discrete optimization considers multiple scenarios for the same alternative; we model such scenarios as realizations of the environmental parameters denoted by $\mathbf{p}$. However, in practice $\mathbf{p}$ can exhibit a probabilistic behaviour as well.

We are not aware of any extension of Class 1 robustness-seeking programs to fuzzy logic or to imprecise probabilities theories. However, applications of this type are conceivable in practice via interval dominance (recall def. 8 and 9). Indeed, different approaches for estimating the mean and the variance of imprecise or fuzzy quantities that can serve as the basis for practical calculations

Functions	Authors and works
Expected value as representative value:	
• $R(F(\mathbf{x}), \mathbf{x}, \delta_x) = \frac{1}{n}\sum_{i=1}^{n} F(\mathbf{x} + \delta_{x,i})$	S. Tsuitsui & A. Ghosh [200], M. Sevaux & K. Sörensen [185, 193], Y-S. Ong, P.B. Nair & K.Y. Lum [147], I. Paenke, J. Branke & Y. Jin [148].
• $R(F(\mathbf{x}), \mathbf{x}, \delta_x) = \left(f_1^{\text{eff}}(\mathbf{x}, \delta_x), \ldots, f_k^{\text{eff}}(\mathbf{x}, \delta_x)\right)^t$ s.t.: $I(F(\mathbf{x}), \mathbf{x}, \delta_x) \equiv \frac{\lvert f_i^{\text{eff}}(\mathbf{x}+\delta_{x,i}) - f_i(\mathbf{x})\rvert}{\lvert f_i(\mathbf{x})\rvert} \leq \eta$, where $f_i^{\text{eff}}(\mathbf{x}, \delta_x) = \frac{1}{n}\sum_{i=1}^{n} f_i(\mathbf{x} + \delta_{x,i})$ and $F(\mathbf{x}) = (f_1(\mathbf{x}), f_2(\mathbf{x}), \ldots, f_k(\mathbf{x}))^t$	K. Deb & H. Gupta [34].
Expected value and variance as representative values:	
• $R(F(\mathbf{x}), \mathbf{x}, \delta_x) = \frac{\sigma[F(\mathbf{x})]}{\overline{\sigma}[\mathbf{x}]}$ where $\sigma[\cdot]$ is the standard deviation of its argument and $\overline{\sigma}[\mathbf{x}]$ is the average of $\sigma[x_i]$.	Y. Jin & B. Sendhoff [93].
• $R(F(\mathbf{x}), \mathbf{x}, \delta_x) = w_1 E[F(\mathbf{x})] - w_2 \sigma^2[F(\mathbf{x})]$ where $E[\cdot]$ is the expected value of its argument.	D.W. Coit, T. Jin & N. Wattanapongsakorn [29].
• $R(F(\mathbf{x}), \mathbf{x}, \delta_x) = \left(E[F(\mathbf{x})], -\sigma^2[F(\mathbf{x})]\right)^t$ where $\sigma^2[\cdot]$ is the variance of its argument.	D. Greiner, J.M. Emperador, B. Galván & G. Winter [66], M. Marseguerra, E. Zio, L. Podofillini & D.W. Coit [132], D.W. Coit, T. Jin & N. Wattanapongsakorn [29], I. Paenke, J. Branke & Y. Jin [148].
Extreme values as representative values:	
• $R(F(\mathbf{x}), \mathbf{x}, \delta_x) \equiv$ worst case$\{F(\mathbf{x} + \delta_{x,i}) : \forall i\}$	Y-S. Ong, P.B. Nair & K.Y. Lum [147].
• $R(F(\mathbf{x}), \mathbf{x}, \delta_p) \equiv$ worst case$\{F(\mathbf{x} + \delta_{p,i}) : \forall i\}$ where the scenarios are defined as realizations of vector $\mathbf{p}$.	P. Kouvelis & G. Yu (cf. [80]).
• $R(F(\mathbf{x}), \mathbf{x}, \delta_p) \equiv$ Estimated worst case for some confidence level and probability content.	S. Martorell et al. [136], J.F. Villanueva et al. [204].

Table 4.7: Some examples of Class 1 $R(\cdot)$ realizations in the literature.

of these representative values, as well as the construction of MOEA, exist in the literature. Table 4.8 surveys some of these contributions[7].

Class 2: Robust domain-seeking programs

Some situations may keep the uncertainty associated with the input variables from being propagated through $F(\mathbf{x})$ and therefore handled as a Class 1 robustness problem. Whether the uncertainty is purely epistemic or mixed in nature, any assertion about $\underline{\delta_x} \leq \delta_x \leq \overline{\delta_x}$ or $\underline{\delta_p} \leq \delta_p \leq \overline{\delta_p}$ entails making some assumptions that could lead to inadequate or even artificial estimations of some representative parameters in $\mathbf{Y}$ with the consequent misclassification of the optimal solutions of a Class 1 $R(\cdot)$.

For example, if the DM know that the real value of the nominal vector $\mathbf{x}$ is susceptible to vary but ignore the range of such variation, to assume the set of

[7] Daniel Berleant's web page contains many other contributions in interval uncertainty by different authors: http://class.ee.iastate.edu/berleant/home/ServeInfo/Interval/intprob.html

Realm of study	Authors and works
Fuzzy logic:	
Estimation of the mean and variance.	R. Kröger [113], D. Dubois and H. Prade [42], A.G. Bronevich [21], C. Carlsson & R. Fullér [25].
Dempster-Shafer:	
Computing mean and variance.	V. Kreinovich, G. Xiang & S. Ferson [112]
p-boxes:	
Estimation of the expected value.	M. Bruns, C.J.J. Paredis & S. Ferson [22, 23].
Intervals:	
Fast computation of statistics for intervals and bounding.	G. Xiang [218], S. Ferson, L. Ginzburg, V. Kreinovich & J. Lopez [46].
Monte Carlo simulation:	
Techniques for processing interval uncertainty in engineering.	V. Kreinovich, J. Beck, C. Ferregut, A. Sanchez, G. R. Keller, M. Averill & S. A. Starks [111].

Table 4.8: Some contributions to the calculation of the mean and variance of fuzzy or imprecise probability quantities that can help towards the resolution of Class 1 $R(\cdot)$ functions.

values that $\mathbf{x}$ can take or their likelihood may underestimate uncertainty. The DM are therefore compelled to maximize the range of 'acceptable' realizations of $\mathbf{x}$ in order to hedge against regrettable consequences. In that sense, robustness is sought by widen the range of possible inputs, regarding the nominal vector $\mathbf{x}$, that comply with quality requirements imposed on $F(\mathbf{x})$ which define what is deemed as 'acceptable'.

Class 2 is then characterized by the existence of constraints of the type $I(F(\mathbf{x}), \mathbf{x}, \mathbf{p}, \delta_x, \delta_p, \gamma) \leq 0$ that constitute desired performance levels of attainment (quality requirements), and a robustness function $R(\cdot)$ that aims at maximizing the range of variation of the input variables.

Class 2 $R(\cdot)$ admits different formulations, but in every case it defines a domain of allowable deviations of the input variables. On the other hand, it differs from the constraint satisfaction problem in that the system constraints $G(\cdot)$ are to be strictly satisfied, no matter the level of robustness of the solution. In other words, feasibility is a hard demand upon which robustness is sought.

It is a matter open to research the functional expression that $R(\cdot)$ should show. One possibility present in the literature is to assess robustness in terms of the size of the allowed deviation in the input. For instance, as we mentioned earlier on, $R(\cdot)$ can be defined as the L_p distance between the maximal deviation from the centrepoint $\mathbf{x}$ [76, 140, 139], then by changing p the DM obtain domains of different shapes. Likewise, robustness can be assessed as the maximal α in a info-gap model (see sec. 2.2.8).

By contrast, other approaches measure the level of robustness in terms of the size or the cardinality of the allowed deviation domain. Hence, B. Roy measures robustness as the cardinality of the set of scenarios that comply with $I(\cdot)$ type constraints in discrete problems (see def. 17 to 19). Likewise, for continuous problems the cardinality can be assessed, e.g. as the Lebesgue measure of the domain. For instance, C. Barrico and C.H. Antunes [10] indirectly recours to such measure by considering the integer number of times a square domain can

be enlarged without missing $I(\cdot)$-type constraints.

In chapter 5 we shall bring several examples of Class 2 problems where $R(\cdot)$ is calculated as the maximal inner box (MIB) (see [158, 155, 156] and references therein), according to the following definitions:

Definition 21 (Inner hyper-Box) *Consider the Robustness-seeking program given by def. 20. An inner hyper-box* **IB** *of a nominal vector* $\mathbf{x}$*, denoted as* $\mathbf{IB}(\mathbf{x})$ *is a subset defined as:*

$$\mathbf{IB}(\mathbf{x}) = \{\mathbf{z} \in \mathbb{R}^n | z_i \in [x_i + \underline{\delta_{x,i}}, x_i + \overline{\delta_{x,i}}]\} \tag{4.14}$$

such that $G(\cdot) \leq 0$ *and* $I(\cdot) \leq 0$ *hold.*

Notice that the inner box[8] coincides with the larger convex hull that can be defined around $\mathbf{x}$ regarding the Cartesian product $[x_1 + \underline{\delta_{x,1}}, x_1 + \overline{\delta_{x,1}}] \times \ldots \times [x_i + \underline{\delta_{x,i}}, x_i + \overline{\delta_{x,i}}] \times \ldots \times [x_n + \underline{\delta_{x,n}}, x_n + \overline{\delta_{x,n}}]$, an hypercube described from the set of n intervals $[x_i + \underline{\delta_{x,i}}, x_i + \overline{\delta_{x,i}}]$, each of which is defined for the i-th component of $\mathbf{x}$ in such a way that $G(\cdot) \leq 0$ and $I(\cdot) \leq 0$ and $\mathbf{x} \in \mathbf{X}$ are satisfied. This definition is a sound and intuitive manner to express robustness in practice since it defines close intervals for each component in such a way that every combination defined from such intervals is valid. Moreover, the condition of symmetry can be imposed as a logical addition that reinforces its practical convenience[9]. In contrast, other convex subdomains no longer allow for every possible combination within the Cartesian product, as those proposed in [76] from L_p metrics which amount to domains with complex shapes (e.g. with $p = 2$).

The cardinality of an inner box corresponds to its Lebesgue measure:

$$|\mathbf{IB}(\mathbf{x})| = \prod_{i=1}^{n} |\overline{\delta_{x,i}} - \underline{\delta_{x,i}}| \tag{4.15}$$

and if additionally we demand symmetry, i.e. if $x_i = (\overline{\delta_{x,i}} - \underline{\delta_{x,i}})/2$, the preceding expression can be written as

$$|\mathbf{IB}(\mathbf{x})| = 2^n \prod_{i=1}^{n} |\overline{\delta_{x,i}} - x_i| = 2^n \prod_{i=1}^{n} |\underline{\delta_{x,i}} - x_i| \equiv \prod_{i=1}^{n} |\overline{\delta_{x,i}} - x_i| \tag{4.16}$$

which corresponds to the definition in terms of the centrepoint $\mathbf{x}$. Thus, considering symmetry, to maximize the Class 2 robustness is equivalent to find that setting with maximal inner box's cardinality. If the centrepoint is fixed a priori (fig. 4.15 left), it amounts to solve:

Definition 22 (Maximal Inner hyper-Box with centre specified program)

[8] The inner box can be viewed as a special lower set (see fig. 2.6).

[9] A limit case of symmetry is the hyper-cubic domain, as is used by C. Barrico and C.H. Antunes [10]. As this demand could be very restrictive, it suffices to us that $\mathbf{x}$ be the centroid.

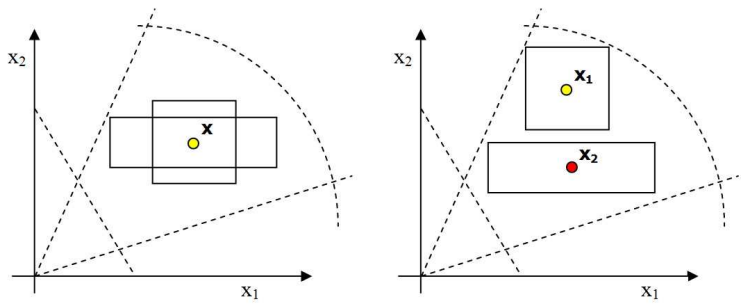


Figure 4.15: Examples of different inner hyper-boxes for specified (left) and unspecified (right) centres ($G(\cdot)$ and $I(\cdot)$ constraints are represented by dotted lines).

Find the lower bounding nominal vector $\underline{\delta_x}$ such that:

$$\begin{aligned} &\underline{\delta_x} = \arg\max_{\underline{\delta_x}} R\left(F(\mathbf{x}), \mathbf{x}, \delta_x\right) \\ s.t.: \quad & R\left(F(\mathbf{x}), \mathbf{x}, \delta_x\right) = |\mathbf{IB}(\mathbf{x})| \equiv \textstyle\prod_{i=1}^{n} |\underline{\delta_{x,i}} - x_i| \\ & \mathbf{x} \in \mathbf{X} \\ & G(\mathbf{x}, \delta_x) \leq 0 \\ & I(F(\mathbf{x}), \mathbf{x}, \delta_x, \gamma) \leq 0 \end{aligned} \tag{4.17}$$

likewise if the centrepoint is also a decision variable (fig. 4.15 right) we get:

Definition 23 (Maximal Inner hyper-Box with centre unspecified program) *Find the nominal vector* $\mathbf{x}$ *such that:*

$$\begin{aligned} &\mathbf{x} = \arg\max_{\mathbf{x}, \underline{\delta_x}} R\left(F(\mathbf{x}), \mathbf{x}, \delta_x\right) \\ s.t.: \quad & R\left(F(\mathbf{x}), \mathbf{x}, \delta_x\right) = |\mathbf{IB}(\mathbf{x})| \equiv \textstyle\prod_{i=1}^{n} |\underline{\delta_{x,i}} - x_i| \\ & \mathbf{x} \in \mathbf{X} \\ & G(\mathbf{x}, \delta_x) \leq 0 \\ & I(F(\mathbf{x}), \mathbf{x}, \delta_x, \gamma) \leq 0 \end{aligned} \tag{4.18}$$

In practice, an ideal procedure to find the MIB should check that none of the constraints is violated. This is possible, e.g., by resorting to the use of interval arithmetic (see sec. 2.2.1), as in [179, 176, 158, 178, 156, 155] or via simulation [180]. The former has the advantage of propagating the $\mathbf{IB}(\mathbf{x})$ with a single calculation; this way the image of $F(\mathbf{IB}(\mathbf{x}))$ and the constraints $I(\cdot) \leq 0$ can be checked easily. However, the use of interval arithmetic can overestimate the size of such image, with the consequent misclassification of some inner boxes as not $I(\cdot)$-compliant. In compensation, the analyst is certain that the $I(\cdot)$-compliant inner boxes are rightly classified. In contrast, when interval arithmetic is not applicable due to the characteristics of the problem, the use of Monte Carlo simulation allows estimating the image of the inner boxes but at the expense of computation time and without a complete confidence in that $I(\cdot) \leq 0$ holds.

Class 2 robustness and uncertainty theoretical frameworks

Let us try to account for the different interpretations of the Class 2 robustness program according to some of the theoretical frameworks for representing uncertainty surveyed in chapter 2. For that matter, let us call $\mathcal{U}(\mathbf{x})$ the set of 'uncertain' values that $\mathbf{x}$ will take regarding its nominal value. Likewise let be $\mathcal{D} = \{\mathbf{x}|\mathbf{x} \in \mathbf{X} : G(\mathbf{x}) \leq 0\}$ be the set of feasible solutions.

The inclusion of new constraints $I(\mathbf{x}) \leq 0$ redefines the domain of acceptable solution vectors, so now we want the values that $\mathbf{x}$ will take to belong to $\mathcal{D}_{\text{eff}} = \{\mathbf{x} \in \mathcal{D} : I(\mathbf{x}) \leq 0\}$. Thus, if $\mathcal{U}(\mathbf{x})$ would be known, it suffices to verify that $\mathcal{U}(\mathbf{x}) \in \mathcal{D}_{\text{eff}}$ to accept the system defined by $\mathbf{x}$ as acceptable. However, since $\mathcal{U}(\mathbf{x})$ is unknown or ill-defined (recall this is a characteristic of Class 2 robustness problems), we wish to find that $\mathbf{x}$ such that $\mathcal{U}(\mathbf{x}) \in \mathcal{D}_{\text{eff}}$ is certain, which is the same to say that maximizing $R(\cdot)$ is equivalent to maximize $N_{\mathcal{U}(\mathbf{x})}(\mathbf{IB}(\mathbf{x}))$ (see eq. 2.20). In simple words, we attempt to define the system in such a way that the allowable deviations from the nominal value $\mathbf{x}$ would be as larger as possible to contain $\mathcal{U}(\mathbf{x})$.

For (binary) classic possibility theory, it is easy to see that the maximization of $\mathbf{IB}(\mathbf{x})$ as defined in def. 22 and 23 converges to the maximum of $N_{\mathcal{U}(\mathbf{x})}(\mathbf{IB}(\mathbf{x}))$. However, when $\mathcal{U}(\mathbf{x})$ has an associated distribution, whether it be of possibility or probability, the maximization of the inner box's hypervolume ceases to be the goal, instead it becomes the density or mass enclosed within it.

Example 4.2 (Comparing the MIB with distributions) *Let* $\mathbf{x} = (x_1 = 2, x_2 = 2)^t$ *be a nominal decisional vector whose components might be subject to epistemic or aleatory uncertainty. Suppose that, with respect to* $\mathbf{x}$*, the former type of uncertainty is represented as a number uniformly distributed in* $[-0.5, 4.5]$*, while the latter corresponds to a* -unknown- *gaussian (symmetric) noise with the quartiles* $x_{5\%} = -0.5$ *and* $x_{95\%} = 4.5$*. Thus, when we associate a uniform distribution to any component, we say that the maximum possible variation in the corresponding dimension is* $[-0.5, 4.5]$*. In contrast, a normal distribution implies that the range of variation is larger (actually infinite) since the noise is gaussian.*

Now let us consider two inner boxes, viz. $\mathbf{IB_1}(\mathbf{x}) = \{\mathbf{z} \in \mathbb{R}^2 | 1 \leq z_1 \leq 3, 1 \leq z_2 \leq 3\}$ *and* $\mathbf{IB_1}(\mathbf{x}) = \{\mathbf{z} \in \mathbb{R}^2 | 0 \leq z_1 \leq 4, 1.5 \leq z_2 \leq 2.5\}$*. The following table brings the probability density enclosed within them when their components are affected by distinct uncertainties:*

x_1	x_2	$\mathbf{IB}(\mathbf{x})$	$P(z_1 \in \mathbf{IB}(\mathbf{x}))$	$P(Z_2 \in \mathbf{IB}(\mathbf{x}))$	$P(Z_1, z_2 \in \mathbf{IB}(\mathbf{x}))$
Uniform	*Uniform*	*1*	*0.4*	*0.4*	*0.16*
		2	*0.8*	*0.2*	*0.16*
Uniform	*Normal*	*1*	*0.4*	*0.4894*	*0.1958*
		2	*0.8*	*0.2578*	*0.2062*
Normal	*Uniform*	*1*	*0.4894*	*0.4*	*0.1958*
		2	*0.8118*	*0.2*	*0.1624*
Normal	*Normal*	*1*	*0.4894*	*0.4894*	*0.2395*
		2	*0.8118*	*0.2578*	*0.2093*

As expected, if the two uncertainties are uniform, which corresponds to epistemic uncertainty for every component, same cardinality implies same robust-

ness. Thus, e.g. if $\mathbf{IB_1}$ *and* $\mathbf{IB_2}$ *represent the MIB of two different designs centred in the same point, both designs are equally robust. It can be interpreted as the amount of uncertainty (expressed as enclosed joint probability density) that they both can absorb is the same.*

However, the two shapes cease to be equivalent due to the non-uniformity in the probability density distribution, and therefore now the hyper-volume is no longer suitable to assess robustness.

In the previous example notice that, if the gaussian noise were known and sufficiently characterized, there is no reason to use a Class 2 $R(\cdot)$ formulation instead of Class 1 one for the normal-normal case. On the other hand, mixed uncertainties represent a challenge and its treatment will depend upon the available information. If the types of uncertainties associated with the domain are completely unknown, the maximization of the inner box's hyper-volume is well enough. On the contrary, if the aleatory part can be known, the MIB can be defined as the box that encloses the greatest density.

Class 2 $R(\cdot)$ formulations can be extended to consider imprecise probabilities and fuzzy logic, but the issue remains unexplored. For instance, consider a problem when the specifications of the $I(\mathbf{x})$ constraints are given in a fuzzy way, so now the search for the most robust alternative should consider not only the range of $\mathbf{IB}(\mathbf{x})$ but its membership function. Indeed, by means of the extension principle (see sec. 2.2.2) every inner box can be propagated; therefore the goal is still to find that system whose inner box, once propagated, complies with the fuzzy performance requirements $I(\mathbf{x})$, as is the general case of Class 2 formulations.

However, the difficulty that arises is the way to define the $\mathbf{IB}$, or in other words, the meaning of robustness in fuzzy theory. Notice that a reverse application of the extension principle yields a membership function instead of a range of variation, as is the case of crisp values. Hence, the early idea of maximizing the allowed variation should be adapted to the fuzzy case regarding the shape of the membership function. Then, the comparison of two inner boxes might be tackled, for instance, by averaging the α-cuts:

$$\overline{\mathcal{A}_\alpha} = \int_\alpha \alpha \mathcal{A}_\alpha d\alpha \tag{4.19}$$

which can be estimated as $\sum_{\alpha_i} \alpha_i \mathcal{A}_{\alpha_i}$ for a suitable partition α_i of α. For example, if we adopt a step size of 0.1 for α_i, the average α-cuts of the three sets shown in fig. 2.8 are $\overline{\mathcal{A}_\alpha} = 3.3$, $\overline{\mathcal{B}_\alpha} = 4.214$ and $\overline{\mathcal{C}_\alpha} = 7.15$ respectively. Then, if these three sets were candidate solutions for a Class 2 robustness-seeking program, the alternative with the larger value, i.e. set $\mathcal{C}$ in this case, is the more robust one. Eq. 4.19 is just a proposal inspired in Yager's first index [221, 220] and its properties as a robustness measures are to be thoroughly studied.

We will not return to this point in the remainder of this thesis, since it lays out of the scope of this work; however it is worthwhile to point out that the issue of formulating Class 2 programs with fuzzy theory or imprecise probabilities and the evaluation of its theoretical and practical relevance, if any, remains open.

Class 3: Mixed robust-seeking procedure

To close this section, let us consider again the general robustness-seeking program formulated in def. 20. The two classes of robustness studied heretofore,

derive from a partial knowledge of the elements constituting such definition. Hence, Class 1 involves an amount of information about the uncertainty associated to the input enough to propagate it and to assess its impact in the output. Class 2, on the other hand, entails the ability of formulating performance levels of attainment that make possible to investigate what is the alternative or design with the highest resistant to the input uncertainty, expressed as the size of the domain of allowed variation covered.

However, if the DM and the analyst are unable to characterize the input uncertainty nor the desired performance levels of attaintment, or alluding def. 20, if they cannot set δ_x, δ_p and $I(F(\mathbf{x}), \mathbf{x}, \delta_x, \gamma)$, the actual definition of the robustness function $R(\cdot)$ is not possible.

It is therefore mandatory to generate information to help the DM to make their minds about δ_x, δ_p or about $I(\cdot)$, in such a way that the problem collapses into a Class 1 or Class 2 program. For instance, A. O'Hagan and J.E. Oakley [146] remark the role of the elicitation procedure and outline a methodology to put it into practice that could be applied to describe δ_x, δ_p effectively, albeit not necessarily in an easy way, should a Class 1 program be chosen. On the contrary, if the DM opt to solve a Class 2 program, an initial assumption about δ_x, δ_p is necessary to roughly approximate the Pareto frontier, allowing to set $I(\cdot)$ and to solve the corresponding Class 2 program afterwards. This procedure shall be exemplified later on in sec. 5.3.

4.4 Information based Perspective for Robustness Analysis

In this section we strive to put all the issues discussed heretofore together in an information-based perspective named *Analysis of Uncertainty and Robustness in Evolutionary Optimization* (AUREO). The AUREO constitutes an analytic methodology that encompasses theoretical and practical facets of uncertainty representation, EA design and MCDM. It is intended for practical purposes, as to help the selection or design of MOEA for particular problems, as well as to outline, analyze and envisage further developments in the state of the art in EMO with uncertainty handling.

The AUREO comprises two stages or levels of analysis (see fig. 4.16): the first one focuses on the types, sources and effects of uncertainty regarding the original functions or figures of performance that are to be optimized and/or controlled. Besides it covers the options to transform the original problem into one of uncertainty handling. The second stage, on the other hand, focuses on implementation issues as the selection of the metaheuristic, the alternatives for propagating the uncertainty, the efficiency of the procedure and the comparison of results. The flow of information through these two stages is not necessarily linear but it might turn out a loop. However, dividing the problem into two stages helps the analyst/DM to realize what difficulties come from uncertainty and what come from the algorithm.

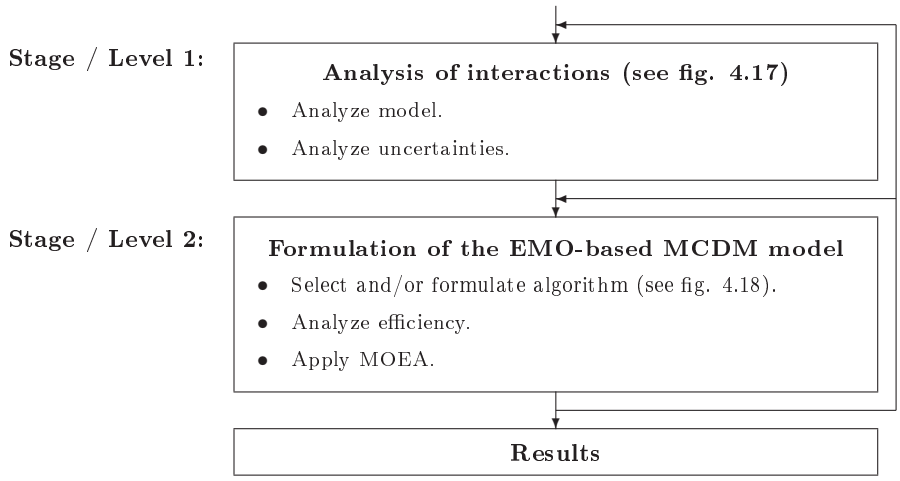


Figure 4.16: AUREO: The two levels or stages of analysis.

4.4.1 AUREO Stage 1: From the analysis of interactions to the formulation of a new program

The main task in the first stage is to determine what are the alternatives to reformulate the original problem in order to handle uncertainty. Therefore, the analysis is intended to identify the sources and the areas of influence of the present uncertainty as well as its effects. Fig. 4.17 shows the three instances upon which this analysis focuses and brings some question to guide it.

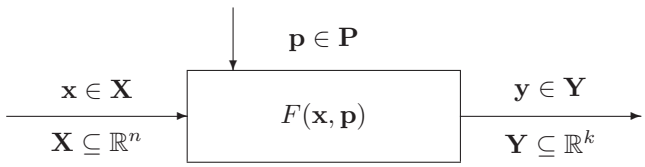
The first concern is about the original model itself. The system's performance is suppose to be suitable described by means of a function (with multiple objectives in the realm of this thesis). If this function does not fit well enough that thing intended to be described, further efforts to handle uncertainty might turn out useless.

On the other hand, the characteristics of the domain and the objective functions determine not only the applicability of one or another solving method but the reformulation of the problem. Recall e.g. the different forms of the Class 2 robustness seeking program for discrete (def. 17 and 18) or continuous domains described earlier on.

The second issue to analyze, once the suitability of the model is verified, is the presence of uncertainty associated to the objective functions. Noisy or dynamic (see sec. 3.4) functions have uncertainty naturally attached and thus their outcomes vary after several evaluations of the same argument. A further source of uncertainty appears when a function that is not uncertain in nature is approximated by simulation or surrogate models. The necessity of reducing the computational effort and hence the time of calculation often comes along even before worrying about the selection and implementation of the heuristic. Then, a trade-off between accuracy of the calculations and making decisions under uncertainty with less computational time is one of the issues to consider here.

Finally the third element to take into account is the presence of uncertainty

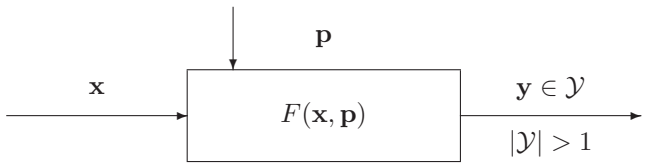
1. Analyze the model:



1.1. Check for model adequacy.

1.2. Consider characteristics of the domain and objective functions.

2. Check for uncertain objective functions:



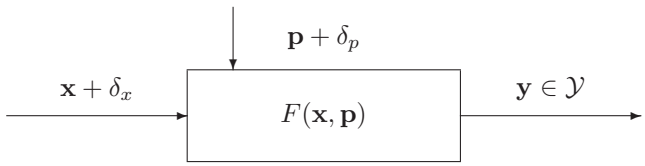
2.1. Do many evaluations of the same argument produce different outcomes?

2.2. Is $F(\mathbf{x}, \mathbf{p})$ a dynamic or stochastic function?

2.3. How is $F(\mathbf{x}, \mathbf{p})$ to be evaluated (surrogate model, approximation, simulation)?

2.4. Is the cardinality of $\mathcal{Y} \subseteq \mathbf{Y}$ reducible to the unit?

3. Check for input uncertainties:



3.1. Is $\mathbf{x}$ subject to uncertainty (δ_x)? If so, what type?

3.2. Are the environmental parameters $\mathbf{p}$ subject to change (δ_p)?

3.3. Is the objective function sensitive to uncertain inputs ($|\mathcal{Y}| > 1$)?

Figure 4.17: AUREO Stage 1: Analysis of interactions between model and uncertainties.

Input: $\mathbf{x}+\delta_x, \mathbf{p}+\delta_p$		**Output:** $\mathbf{y} \in \mathcal{Y}, \lvert\mathcal{Y}\rvert > 1$	
δ_x	δ_p	$I(\cdot)$ definable	$I(\cdot)$ undefinable
None	None	**Comparison of sets ≡ Class 1**	
None	Definable	**Class 1**	**Class 1**
None	Undefinable	**Class 2**	**Class 3**
Definable	None	**Class 1**	**Class 1**
Definable	Definable	**Class 1**	**Class 1**
Definable	Undefinable	**Class 2**	**Class 3**
Undefinable	None	**Class 2**	**Class 3**
Undefinable	Definable	**Class 2**	**Class 3**
Undefinable	Undefinable	**Class 2**	**Class 3**

Table 4.9: AUREO Stage 1: Classes of uncertainty-handling formulations (see sec. 4.1.2 and 4.3.3) according to the available information.

attached to the input variables. If this uncertainty actually exists, it indicates a robustness-seeking program. However, the type and the impact of such uncertainty are crucial factors. A preliminary sensitivity analysis could help to evaluate if the effects of such uncertainty could be neglected or not.

Once the uncertainties have been identified, the analyst should determine the final formulation of the uncertainty-handling program that derives from the current level of information. Indeed, as we have pointed out formerly, the alternatives to handle uncertainty depend on the available information and therefore it is such information the definitive element that defines the class of problem to solve. Hence, regarding table 4.9, a scenario of a uncertain function with a fully defined input (third row) is a case of comparison between sets (see sec. 4.1.2). In practice it is equivalent to Class 1 robustness problems in that function $R(\cdot)$ is meant to lead to the optimization of some selected representatives values of $F(\mathbf{x}, \mathbf{p})$ outcomes, as is in uncertainty-handling programs with uncertain functions. Nevertheless, the procedures to determine such values differ from one case to another, since the robustness program implies some kind of uncertainty propagation (see fig. 4.18) that does not take place with fully defined inputs.

In contrast, be $F(\mathbf{x}, \mathbf{p})$ or not a uncertain function, uncertain inputs always induce robustness-seeking programs, yet its treatment can be neglected for practical reasons after a sensitivity analysis. Hence, the rest of the cases shown in table 4.9 require the reformulation of the original problem to one of robustness.

4.4.2 AUREO Stage 2: From the uncertainty-handling program to an efficient implementation

The link with the second stage is given by a detailed analysis of the elements that will characterize the final implementation of the multiple-objective heuristic. Fig. 4.18 presents three questions that summarize the targets to identify in this step. The sequence for answering such questions may vary from one problem to another.

For every specific problem the analyst should identify the type of uncertainty-handling formulations prescribed in tab. 4.9 for the case under study. Afterwards, the analysis should lead to identify the existing MOEA to use, with the eventual adaptations to the particular problem under study, or the envisaged

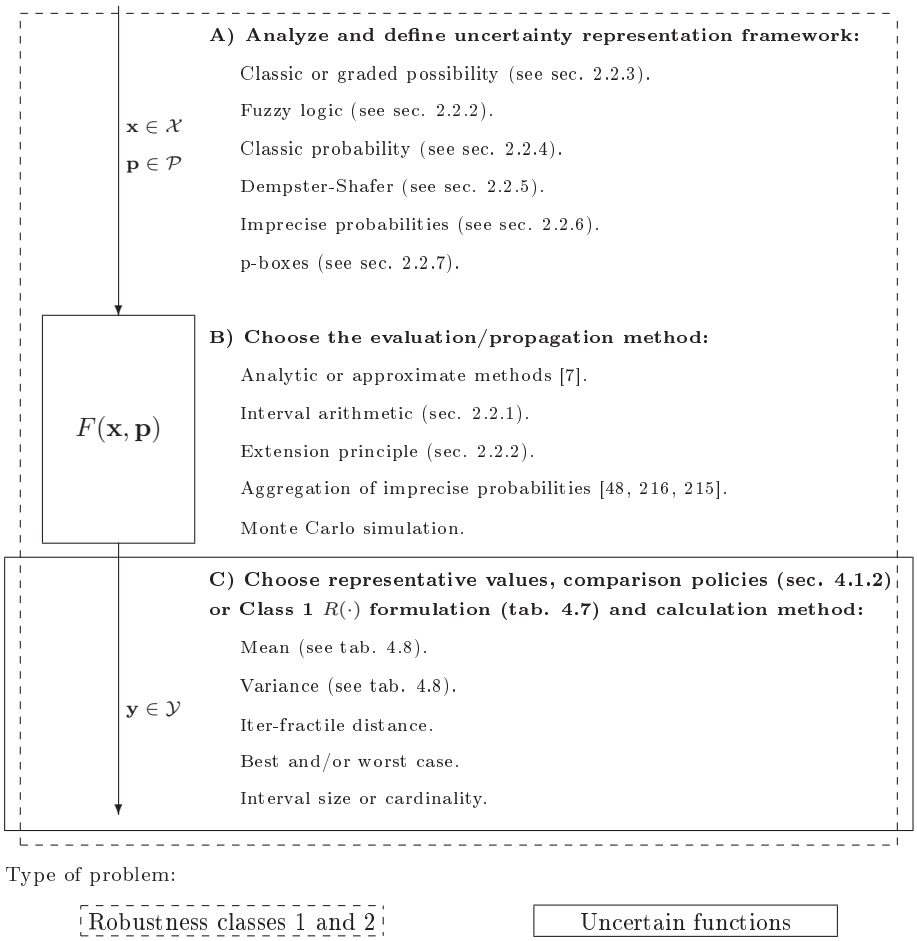


Figure 4.18: AUREO Stage 2: Elements of model formulation for uncertainty handling.

algorithm to design. Table 4.10 brings some existing alternatives as well as some possible extensions of them that can be used to solve the uncertainty-handling formulations prescribed in tab. 4.9.

If the problem is one of comparison of sets or a Class 1 robustness problem, the MOEA ranking procedure should optimize a $R_1(\cdot)$ type function without missing the $G(\cdot)$ and eventual $I(\cdot)$ constraints. Class 2 robustness problems require the same treatment but for functions of type $R_2(\cdot)$. However, the implementation and efficiency issues vary from one kind of problem to another.

In practice, robustness problems entail the use of propagation procedures and/or interval analysis but for different reasons. In Class 1 problems, it suffices to propagate the uncertainty only once for each alternative, which does not mean that the objective functions are evaluated once but the procedure for assessing the resulting uncertain vector is applied only one time. Such procedure should allow to determine the representative values as well as for checking constraints violations. In contrast, Class 2 problems require verifying $I(\cdot)$

Class of uncertainty handling formulation	Algorithms and references	Some possible extensions
Comparison of sets:		
• Without extra information (sec. 4.1.3):		
Optimization with epistemic uncertainty:	• P. Limbourg [125], • P. Limbourg & D.E. Salazar A. IP-MOEA [126, 173] (see sec. 4.2.1)	• Adaptation of existing MOEA (see e.g. sec. 3.3.4) to work with representative values. • Construction of $I_{\epsilon+}$ compliant algorithm (see sec. 4.2.2).
• With extra information (sec. 4.1.4):		
Optimization with interval fitness value	• J. Teich [197]	
Optimization with noisy fitness function	• E. Hughes [89, 88], • J.E. Fieldsen & R.M. Everson [51]	• Control of statistical significance to handle cloned chromosomes and to improve convergence and diversity.
Indicator-based optimization	• M. Basseur & E. Zitzler [11]	
Robustness-seeking programs (see sec. 4.3.3):		
• Class 1 programs:		
Optimization with uncertainty propagation:	•See comparison of sets in this table and references, • K. Sörensen [194], • T. Ray [152], • K. Deb & H. Gupta [34]	• Adaptation of existing MOEA (see e.g. sec. 3.3.4) to work with $R_1(\cdot)$ function of tab. 4.7. • Construction of $I_{\epsilon+}$ compliant algorithm (see sec. 4.2.2). • Extensions of existing algorithms to propagate different types of uncertainty using the procedures shown in tab. 4.8. • Control of statistical significance to handle cloned chromosomes and to improve convergence and diversity.
• Class 2 programs:		
Domain-seeking:	• C. Barrico and C.H. Antunes [10]	• Adaptation of existing MOEA (see e.g. sec. 3.3.4) to handle discrete domains (see def. 17 to 19).

Table 4.10: AUREO Stage 2: Existing alternatives and possible extensions in EMO to solve the uncertainty-handling formulations prescribed in tab. 4.9.

constraints compliance every time that the domain of allowed deviations (or the maximal $\mathbf{IB}(\mathbf{x})$) is redefined by the heuristic procedure, for each particular alternative $\mathbf{x}$. It means the propagation time is critical for both classes of robustness-seeking program but it seems especially important for the latter.

On the other hand, the statistical significance of the representative values determined for $R_1(\cdot)$ type functions could affect the performance of the MOEA. Consider that cloned chromosomes in the population will get different fitness due to different factors such as:

- inherent noise of the objective function,
- the use of simulation procedures to propagate uncertainty.

Both cases result in different samples thereby assigning distinct fitness to totally equal alternatives. Similarly, dominance in a mathematical sense not necessarily means dominance in a statistical one. Notice that two representative values

of two different alternatives might be close enough to statistically reject its difference, but they still are mathematically different, and therefore one might dominate the other. The influence of this phenomenon does not appear thoroughly reflected in the state of the art of uncertainty handling in EMO (see tab. 3.3), albeit it might have been considered by some authors. In any case, analyst and MOEA designers must not disregard this issue.

The implementation and efficiency analysis covers many features of the MOEA structure and its final form clearly depends on the problem. Nonetheless, it is possible to draw some guidelines to help such analysis, viz.:

Representation: Find the best way to encode the problem. Consider the type of domain (discrete or continuous). Does the encoding fit well the type of uncertainty under consideration?

Propagation of uncertainty: Is propagation necessary? If so, what are the characteristic of the objective function? Study the analytical ways to propagate uncertainty. Are they applicable? In case of simulation, choose suitable parameters.

Fitness assignment: Does the ranking procedure handle cloned alternatives? Does it consider statistical significance? How does it assess density?

Number of objectives: How many representative values are considered? Are all the representative values to be considered at the same time or there are room for lexicographic approaches?

Amongst the existing categories shown in table 4.10, Class 2 robustness-problems have received the less attention in EMO and hence the second part of this thesis is focused on the application of the AUREO to different kinds of real problems. In particular some particular efficiency issues not treated here are tackled in the applications.

It is pertinent to remark that the efforts of C. Barrico and C.H. Antunes [10] in Class 2 robustness-problems were developed independently and in parallel with some of the algorithmic applications reported in the second part of this thesis. It means that, to the best of our knowledge, this type of problems as not been considered in an EMO context prior this thesis.

4.5 Summary of the chapter and concluding remarks

In this chapter we have introduced the main contribution of this thesis, which is an information-based perspective tool named *Analysis of Uncertainty and Robustness in Evolutionary Optimization* or AUREO for short. The AUREO encompasses the many aspect of uncertainty and robustness in EMO-based MCDM, ranging from the different forms for representing uncertainty to the implementation and the efficiency issues of the solving MOEA to use. AUREO is conceived to help analysts and MOEA designers as well to give them a broader view of the different facets of working with uncertainty and the role of MOEA in real life problems.

The first stage of the AUREO focuses on the analysis of the initial problem and the associated uncertainties. Hence, by means of the information-based

analysis the analyst is guided to identify the class of uncertainty-handling program that translates the problem into one to be treated with MOEA. Afterwards, the implementation aspects of the MOEA like the selection and the adaptation of the algorithm are tackled during the second stage.

Within this framework we have analyzed different categories of problems, identifying many aspects that have not been considered yet in EMO, like the use of theoretical frameworks for uncertainty handling beyond classic probability theory and some attempts in fuzzy logic, and the theoretical and practical work that could underpin such task. The discussion around the AUREO presented in this chapter offered insights into the salient discussion about uncertainty and robustness, from a practical and utilitarian view.

In addition, in this chapter a MOEA named IP-MOEA was presented as an example of non-probabilistic approaches. Afterwards we suggested the use of the ϵ-indicator to design new non-probabilistic MOEA and introduced the concept of probabilistic density to be considered in further EMO research.

This chapter concludes the first part of this thesis. The second part is focused on the application of the AUREO to dependability problems. The dependability of a system is its ability to deliver specified services to the end users or accomplish specified missions so that the users can justifiably *depend* on the services provided by the system. This kind of problems is characterized by the presence of one or many sources and types of uncertainty so it makes the dependability a suitable field to test the AUREO.

Part II

Applications of the AUREO + MCDM in Dependability: Theoretical and Algorithmic Issues

Chapter 5

AUREO + MCDM in Reliability and Scheduling

"Let the problem drive the analysis: This is a very straightforward commandment that means exactly what it says. What you do and the models you build should driven by specifics of the problem you are working on and not by such extraneous factors as what tools you like to use or what software you already have built."

2nd commandment for good policy analisys [142]
by M. Granger Morgan and Max Henrion.

5.1 Introduction

In this chapter we illustrate the AUREO through its application to some reliability and scheduling optimization problems where the notion of robustness plays a central role to hedge the system against uncertainty.

The reliability of a system is a dependability attribute related to the likelihood of the system accomplishing its mission under specified conditions. This measure is particularly important to some kind of systems like non-reparable ones, safety and protective devices and in critical system in general, where the occurrence of failures may carry out severe consequences and thus their likelihood is to be minimized. The first problem studied in this chapter is related to the concept of robust design and aims at finding the maximal range of variation for the reliability of the components that constitute a system. The resolution of this problem exemplifies the notion of Class 2 robustness described in 4.3.3 while remarks the second stage of the AUREO by the introduction of the percentage representation to enhance the efficiency of the MOEA employed.

The second problem tackles a particular flow-shop problem where the sequence of tasks is hampered by a critical machine that blocks the flow. The model studied here is taken from a real life problem and was studied first in [59] but without considering uncertainty. Additionally, the obtained results are featured by many equivalent optimal solutions that hinder the decision-making. The application of the AUREO to this problem leads to identify a Class 3 robustness-seeking program whose resolution required some assumptions to make the problem collapse into one of Class 2. As a result, the application

of the AUREO allowed to differentiate the equivalent solutions found in [59] in terms of their robustness.

5.2 AUREO in Reliability

5.2.1 Basics of Reliability Engineering

In Reliability Engineering theory the *reliability* of a system, denoted as R_S in this chapter, is a measure of the likelihood of a system accomplishing its mission or intended function under stated operation conditions or environment and within a time framework. In most of the cases the classical probability theory (see sec. 2.2.4) supports the reasonings and the reliability is expressed as precise probability, although fuzzy [86], possibilistic [87] or imprecise probabilistic reasonings [202, 150] are possible as well. Hence, the reliability takes values within [0,1]. Likewise, the *unreliability* of a system, denoted as Q_S, is the complementary measure of reliability, thus $Q_S = 1 - R_S$.

When two or more elements compose a system, the reliability of the latter depends upon the reliability of the formers as well as the functional interactions amongst them. The basic modes of interaction are series and parallel: in the former case a failure in one component results in a failure of the whole system, thus the reliability of a series system is expressed by $R_S = \prod_i R_i$. Conversely, in parallel systems, all the components must fail to make the system fails, therefore the reliability is assessed as $R_S = \prod_i (1-R_i) = \prod_i Q_i$. Some systems with many components are not single series nor parallel but a combination of both types. In such a case systems are modelled as a group of subsystems, each of which might be parallel, series or a combination of both. Besides, other features like capacity requirements or k-out-of-n parallel configurations amongst others can be considered, but the formulae given before constitute the basis for reliability calculus. There is an extensive literature on reliability the reader can consult for more details, e.g. [2].

Some systems do not bear series-parallel structures and therefore cannot be assessed directly using the previous formulae. Instead, the analyst should resort to other analytic tools like Fault Tree Analysis (see [2] for details) in order to identify minimal combinations of elements whose simultaneous failure doom the whole system to failure or *minimal cut sets* (MCS).

Every system can be seen as a collection or series of MCS. Some components of the system can be present in many MCS so that the probability of failure of such MCS is not independent but rely upon the replicated components. When the MCS are disjoint, i.e. when no element is present in more than one MCS, the failure probabilities of the MCS are independent. Hence, the sum of all disjoint MCS reliabilities gives the system reliability. However, in practice this addition might bear too many terms and therefore some simplifications are done, like considering MCS composed of up to m elements, where m gives the order of the cut set.

5.2.2 Statement of the problem

This problem considers the robust design of a Nuclear Power Plant safety system: The Residual Heat Removal system (RHR) is a low pressure system (400

psi) directly connected to the primary system which is at higher pressure (1200 psi). The RHR constitutes an essential part of the low-pressure core flooding system which is part of the Emergency Core Cooling System of a nuclear reactor. Its objectives are:

1. to remove decay and residual heat from the reactor so that refueling and servicing can be performed,
2. to supplement the spent fuel cooling capacity,
3. to condense reactor steam so that decay and residual heat may be removed if the main condenser is unavailable following a reactor scram.

In fig. 5.1 a schematic of the system is shown. The unreliability of the system Q_S, i.e. the probability that that systems fails in a specified period of time, can be obtained from the simplified fault tree shown in fig. 5.2 in which the 16 most important (third-order) cut sets, originated by the combination of 8 basic events, are considered [131]. The simplified disjoint expression for Q_S yields:

$$\begin{aligned} Q_S \approx \; & Q_1Q_3Q_5 + Q_1Q_3Q_6R_5 + Q_1Q_3Q_7R_5R_6 \\ + \; & Q_1Q_3Q_8R_5R_6R_7 + Q_1Q_4Q_5R_3 + Q_1Q_4Q_6R_3R_5 \\ + \; & Q_1Q_4Q_7R_3R_5R_6 + Q_1Q_4Q_8R_3R_5R_6R_7 + Q_2Q_3Q_5R_1 \\ + \; & Q_2Q_3Q_6R_1R_5 + Q_2Q_3Q_7R_1R_5R_6 + Q_2Q_3Q_8R_1R_5R_6R_7 \\ + \; & Q_2Q_4Q_5R_1R_3 + Q_2Q_4Q_6R_1R_3R_5 + Q_2Q_4Q_7R_1R_3R_5R_6 \\ + \; & Q_2Q_4Q_8R_1R_3R_5R_6R_7 \end{aligned} \tag{5.1}$$

where Q_i and R_i are respectively the unreliability and reliability of each of the 8 basic events. The reliability robust design optimization problem aims at obtaining the maximum allowed ranges of variation of the reliability of each basic event R_i, $i = 1, 2, \ldots, 8$ such that the system reliability $R_S = 1 - Q_S$ remains bounded as $0.99 \leq R_S \leq 1$. The component reliability values are subject to the constraint $0.80 \leq R_i \leq 1$, i.e. no component can bear a reliability lower than 0.80 since low values entail small intervals of maintenance, a situation that is not convenient in practice. For practical reasons the allowed deviations are conditioned to be symmetrically distributed around the centroid. Hence, the reliability nominal values are initially assumed to be centred in 0.90, thus the centroids of allowed deviations are $c_i = 0.90, i = 1, 2, \ldots, 8$. Furthermore, normally the system design problem seeks also to constrain the system cost, here taken equal to a nonlinear combination of the components' reliability $C_S = 2\sum_i K_i R_i^{\alpha_i}$. In this problem, the vector of coefficients of the nonlinear combination is $K = \{100, 100, 100, 150, 100, 100, 100, 150\}$ and the exponents are $\alpha_i = 0.6, \forall i$.

5.2.3 AUREO: Stage 1

In order to apply the AUREO, we will follow step by step the principles proposed in fig. 4.17:

1. Analyze the model

The problem described is one of 'Robust Design' in the sense of Class 2 robustness. The idea is to determine how much can the reliability of the components

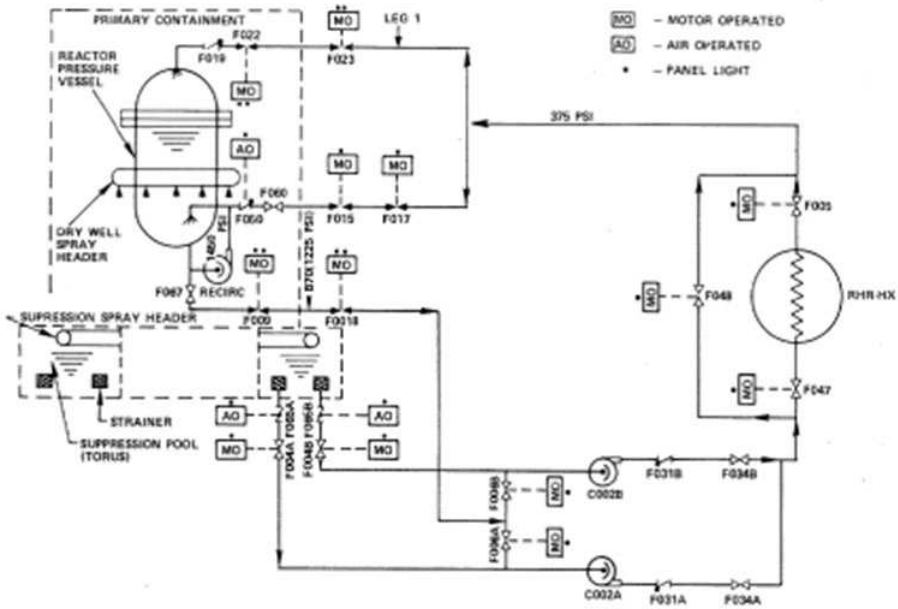


Figure 5.1: Schematic of the Residual Heat Removal system (RHR) of a Nuclear Power Plant (taken from [131]).

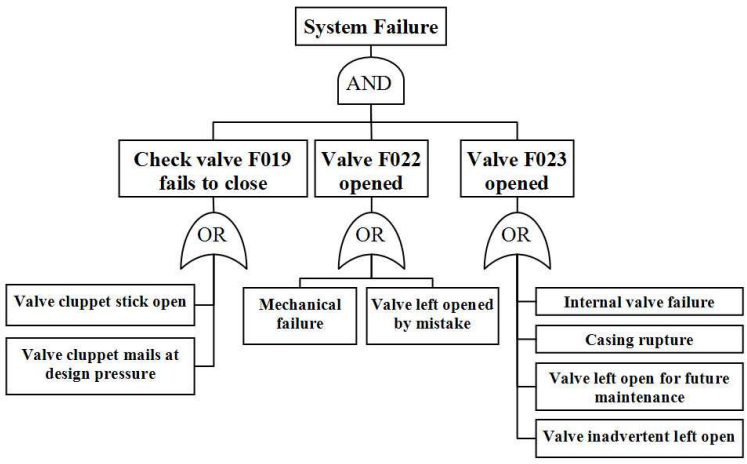


Figure 5.2: Schematic of a Residual Heat Removal system [131].

deviate from the centroid (or nominal value) without missing the performance requirements. The larger the allowed deviation the more robust the system.

If the model requires the nominal values to be fixed, the problem corresponds to the mathematical program described in def. 22. Alternatively, when one introduces more degrees of freedom allowing the nominal values to be set, the program becomes of the type described in def. 23.

2. Search for uncertain objective functions

The calculation of R_S through eq. 5.1 does not entail any type of uncertainty since it can be performed in an accurate manner, although such equation is an approximation. In other words, the only source of uncertainty in the assessment of R_S comes from the fact that its exact value cannot be known through eq. 5.1. However, the impact of the uncertainty induced by such approximation is assumed numerically negligible by the DM in this example and thus the objective function is considered free of uncertainty in a practical sense.

3. Check for input uncertainties

The problem does not contemplate a vector $\mathbf{p}$ of environmental parameters in the calculation of R_S. The only source of input uncertainty is the possible variations in the values of reliabilities R_i. Notice that no distinction is made between epistemic and aleatory uncertainty since, in principle, both types could be present but it is assumed that they cannot be characterized a priori. It is the case, e.g. of a system whose components could be no longer available in stock and they have to be replaced by others with presumably different reliability. Likewise, a component whose nominal value differs from the real one could be integrated into the design if its range of reliability lies within the maximal $\mathbf{IB}(\mathbf{x})$ determined by solving the programs described in def. 22 and 23.

The final robustness-seeking program

Following def. 22 and 23 we will define two problems: the first one consists in finding the MIB for $\mathbf{x}$ with minimal cost C_S, where $\mathbf{x}$ is the vector of components reliability values such that $\mathbf{x} = (R_1, \ldots, R_8)^t$, subjected to the performance requirements $I(R_S) \equiv 0.99 \leq R_S \leq 1$ and the constraints $0.80 \leq R_i \leq 1$ where each reliability R_i is centred in $c_i = c, \forall i$ and c is a constant. The last constraint entails vector $\mathbf{x}$ is centred a priori and all nominal values are the same as well.

The second problem is identical to the first except for $\mathbf{x}$ is not centred a priori but the centroid is a decision variable as well. Recall the range of deviations should be symmetrical regarding the centre or nominal value in both problems.

5.2.4 AUREO: Stage 2 and results

The implementation of MOEA for the first problem is straightforward and the only relevant detail is the constraints handling. On the one hand, every $\mathbf{IB}(\mathbf{x})$ analyzed should meet the constraints $0.80 \leq R_i \leq 1$ and $I(R_S) \equiv 0.99 \leq R_S \leq 1$ to be valid. The former is easily met by defining the search space for each component as $R_i = [0.8, 1]$. The latter one, however, requires a method to propagate the $\mathbf{IB}(\mathbf{x})$ such that condition $R_S(\mathbf{IB}(\mathbf{x})) \in [0.99, 1]$ can be verified. For that matter, R_S has to be evaluated as an interval function according to the

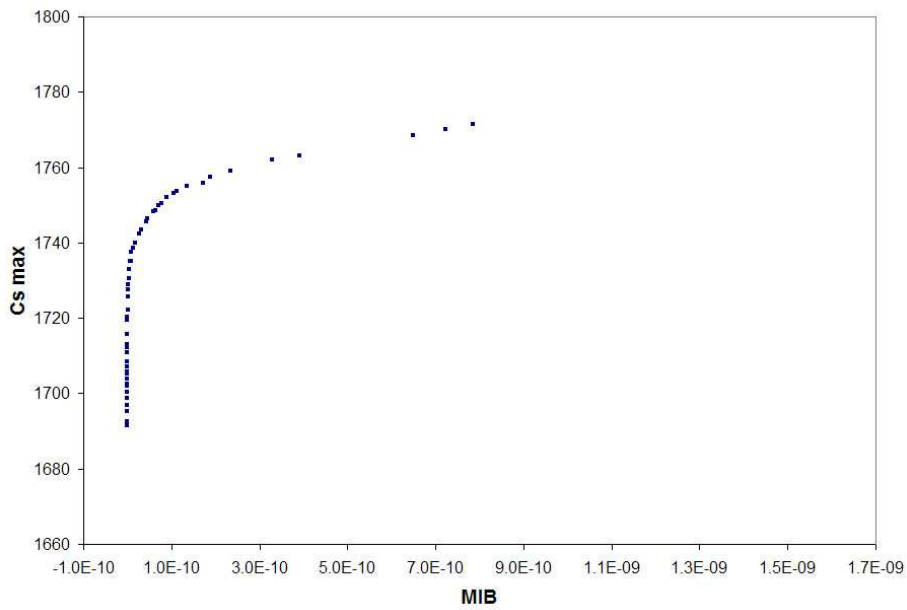


Figure 5.3: Pareto approximation frontier robustness-cost for a design centred in $c_i = 0.90$ [179].

rules surveyed in sec. 2.2.1. On the other hand, the MOEA should implement a constraint-handling mechanism in order to ensure an efficient handling of solutions violating the constraints.

Fig. 5.3 brings the results of using NSGA-II with the following characteristics:

- Algorithmic implementation: see algorithm 4 in fig. 3.16.
- Representation: real representation using the mixed chromosome described in sec. 3.2.4.
- Recombination operators: crossover and mutation operators described in sec. 3.2.4.
- Constraint handling technique: Deb's [36].
- Population size $M = 50$, archive size $N = 50$, number of generations $t^{max} = 250$
- Crossover rate $\rho_c = 0.9$, crossover weight $\alpha = 0.8$, mutation rate $\rho_m = 0.3$.
- Number of function evaluations: 12550.

Concerning the AUREO the results shown in fig. 5.3 mean that the first problem is completely solved. The implementation of the MOEA is able to approximate the Pareto frontier in an efficient way, providing the DM with valuable information about the possible deviations in R_i in terms of the maximal inner box (MIB) and the maximal cost that ones could incur if the top values for

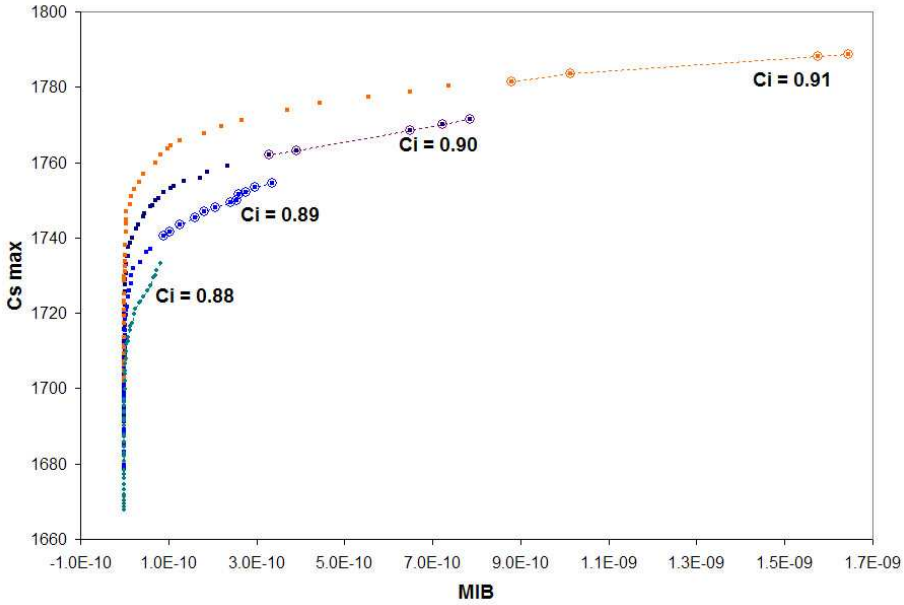


Figure 5.4: Pareto approximation frontiers robustness-cost for non-specified centre using a double-loop approach (non-dominated points are highlighted with rings) [179, 178].

the different R_i are selected. In practical terms, the robustness-seeking program (AUREO Stage 1) and the MOEA implementation (AUREO Stage 2) give the designer a methodology for determining allowed deviation reliability tolerances of the components, thereby making possible to select components from stock that do not necessarily have reliability 0.9, ensuring the investment will not surpass C_s max and the system reliability will not fall below 0.99.

By contrast, when the vector of nominal values is not fixed a priori, the preceding procedure is not sufficient but an additional strategy must be formulated to determine the optimal $\mathbf{x}$; that is the case of the second problem. Indeed, the procedure implemented formerly allows to find the robustness-cost frontier once vector $\mathbf{x}$ is fixed. Therefore a natural alternative is to control the value of $\mathbf{x}$ through an external loop while applying the MOEA to determine the Pareto approximation front for each one of the values assigned to $\mathbf{x}$, this way adopting a double-loop configuration (see sec. 3.2.5 and 3.2.6).

Fig. 5.4 presents the results of solving the second problem using a double-loop. For simplicity the constraint $R_i = c_i, \forall i$ was kept so one single loop is necessary to vary c_i, i.e. to set $\mathbf{x}$ while the NSGA-II is performed to approximate the Pareto frontier for the given $\mathbf{x}$ with the same setting described before. The outcome for every value of $\mathbf{x}$ can be seen in fig. 5.4 as small dots, whereas the Pareto frontier obtained after merging all the outcomes is highlighted with rings and dotted lines.

Notice the approximation front is not found in a particular value of c_i but different fractions of the efficient frontier correspond to different values of $\mathbf{x}$. This suggests the initial condition of centering $\mathbf{x}$ in $c_i = c, \forall i$ should be relaxed

since the optimal solution does not lie on such constraint. Consider as well the low efficiency of the double-loop configuration to solve this problem. Indeed, for exploring four values of c_i it was necessary to perform four runs of the NSGA-II, amounting to $12500 * 4 = 50000$ evaluations. It is possible to reduce the number of generations by tuning the algorithm, nevertheless the complication remains. The search space is continuous and varying $\mathbf{x}$ entails nine nested loops (eight for fixing the R_i and one more for finding the MIB) if one wants to explore the whole space. For instance, if one selects a stepsize of 0.01, the number of evaluations becomes $(0.2/0.01)^9 * t^{max} * M * N$, i.e. the number of evaluations performed by the MOEA to find the MIB times 5.12E11. This clearly makes the nested-loop strategy not affordable.

Percentage representation

At this point, let us consider how a small change in the formulation of the robustness-seeking program can have a noticeable impact upon the efficiency of the implementation, even when the Stage 1 of the AUREO is substantially the same for both problems. During the Stage 2, analysts are concerned with the efficiency of the algorithm and thus they are compelled to develop alternatives approaches to boost the algorithmic performance.

One of the targets during the Stage 2 is the effect of the encoding on the algorithmic implementation. Consider for instance the limitations of using binary or real encoding in the traditional way: In the most general case of a the robustness-seeking program for a n-dimensional design with unspecified centre, the components of $\mathbf{x}$ are the centres c_i and the coordinates of the inner box $\mathbf{IB}(\mathbf{x})$ which can be expressed in terms of the lower vertexes $\underline{x_i}$, where $\underline{x_i} \in [\underline{c_i}, \overline{c_i}], 0 \leq \underline{c_i} \wedge \overline{c_i} \leq 1, i = 1, 2, \ldots, n$. For each c_i, the following constraints stand: the lower vertex must verify $\underline{c_i} \leq \underline{x_i} \leq c_i$ whereas the upper vertex, which can be calculated as $\overline{x_i} = 2c_i - \underline{x_i}$ due to the symmetry condition imposed, is restricted to $c_i \leq \underline{x_i} \leq \overline{c_i}$. Thus, there exists a dependency between the bounds of $\underline{x_i}$ and the value of c_i. Such dependency prevents the algorithm from breed optimal chromosomes found for a particular set of values of centres c_i with chromosomes defined for different c_i, since invalid chromosomes can be produced (recall the $\mathbf{IB}(\mathbf{x})$ is a set that must lie within the search space), in a similar way to the example brought in sec. 3.2.6. In simple words, you cannot mix boxes with centroids without the risk of constraints violation.

Hence, if one adopts a percentage (see sec. 3.2.5) instead of a crisp representation, the inner boxes can be expressed in relative terms regarding $\mathbf{x}$, unifying the search space into an augmented one but with the advantage of ensuring that no offspring will violate the search space. In particular for the second problem described here, each individual was codified as a group of 8 pairs $(c_i, \%x_i^{\max})$ where $\%x_i^{\max}$ represents the percentage of the maximal allowed distance $x_i^{\max}$ between c_i and its bounds (recall $c_i \in [0, 1]$). Therefore, the mathematical relation to determine the value of the vertex x_i that guarantees that only feasible individuals can be produced when the recombination operators are applied is [176]:

$$\%\overline{x_i} = \frac{|c_i - x_i|}{\%x_i^{\max}} = \frac{|c_i - x_i|}{\min\{\overline{c_i} - c_i, c_i - \underline{c_i}\}} \tag{5.2}$$

Notice that in principle, x_i could be either the lower or the upper vertex. Thus, by simply introducing a decision rule during the decoding, we obtain the

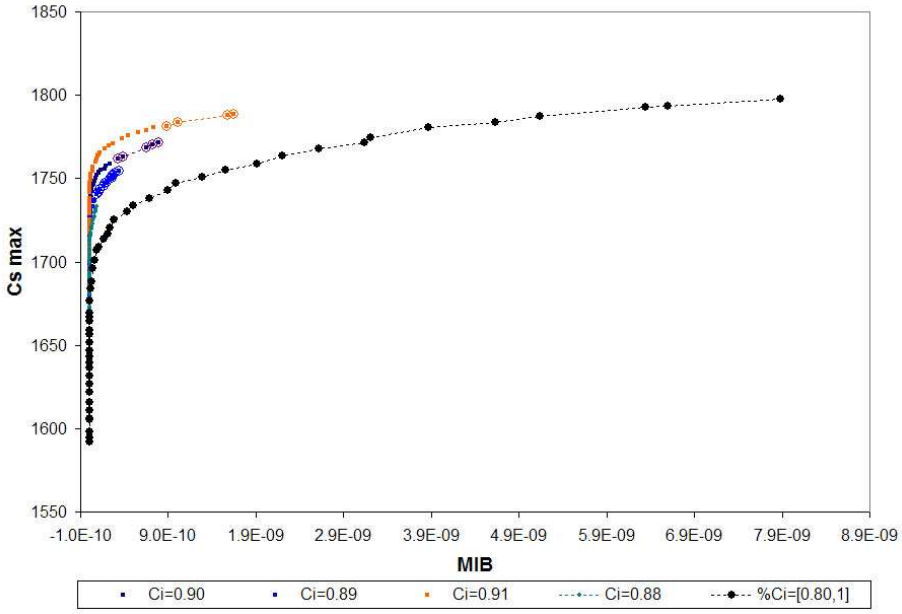


Figure 5.5: Pareto approximation frontiers robustness-cost for non-specified centre using percentage representation.[179, 178]

value of $\underline{x_i}$ as [176]:

$$\underline{x_i} = \begin{cases} c_i - \%x_i^{\max} \cdot (c_i - \underline{c_i}) & \text{if} \quad c_i - \underline{c_i} < \overline{c_i} - c_i \\ c_i - \%x_i^{\max} \cdot (\overline{c_i} - c_i) & \text{otherwise} \end{cases} \tag{5.3}$$

Fig. 5.5 brings the result of using NSGA-II with percentage representation and the same setting employed for the first problem. Observe that all the solutions obtained using this encoding dominate those obtained with the double-loop approach but with noticeable lesser effort. Moreover the algorithm was able to extend the size of the Pareto frontier reaching highly (about eight times more) robust solutions at the same cost of those found for $c_i = 0.91$.

The results presented here are a very good example of the philosophy behind the AUREO: the Stage 1 aims at making the best use of the available information in order to identify the class of uncertainty handling problem that better matches the problem, while the Stage 2 strives for making the best possible implementation using the available heuristic tools or creating new ones.

5.3 AUREO in Scheduling

5.3.1 Basics of scheduling theory

In words of H. Hoogeven "*the basic scheduling problem can be described as finding for each of the tasks, which are also called jobs, an execution interval on one of the machines that are able to execute it, such that all side-constraints are met*"[85, pg. 592]. Hence, a schedule is characterized by a physical sequence,

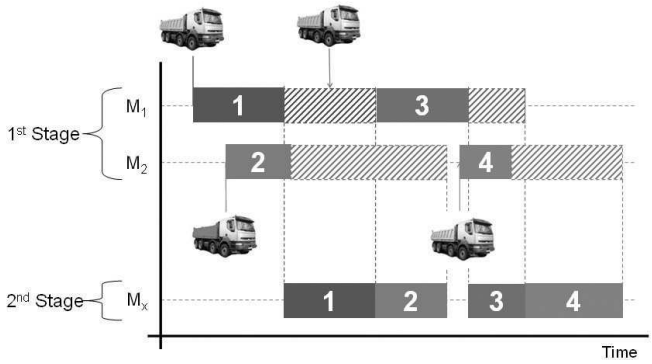


Figure 5.6: RCb scheduling problem in a waste treatment plant (two-stage example): Machines M_m constitute the 1st stage (unloading) while the mixer M_x constitutes the 2nd stage (waste processing) [180].

given by the assignments job-to-machine and a temporal sequence which indicates the intervals of execution of such assignments with respect to a zero time. Every job has a particular *processing time*, which means the total time the job is being executed by a particular machine and can vary from one machine to another. In some cases it is possible to interrupt the execution of a job, something called *preemption*. Thus, a job begins to be executed at the *starting time* and finishes after the *completion time*, which is equal to the sum of the starting and the processing times if no preemption is allowed, or is greater otherwise.

The machine configuration determines the type of scheduling problem to tackle. Here we are concerned with multiple-objective flowshop problems (to learn more about other types of problems see [85] and references therein). According to J. Gupta and E. Stafford a flowshop problem "*is characterized by more or less continuous and uninterrupted flow of jobs through multiple machines in series. In such a shop, the flow of work is unidirectional since all jobs follow the same technological routing through the machines.*"[67, pg. 700]. In flowshop problems all machines work independently of each others. Likewise, the processing time of a job can vary amongst the machines. Finally, not all the jobs have to follow the same sequence of machines [67].

5.3.2 Statement of the problem: The RCb hybrid flowshop bi-objective scheduling problem

This problem was studied first in [59] and was originated in a waste treatment plant. In this problem a job is the complete treatment of the waste transported by a truck, a task accomplished in two operations: the shop consists of a 1st stage composed of a set of $m \geq 1$ silos that are considered identical parallel machines, where a group of trucks unload the waste (1st operation), and a 2nd stage with a unique mixer (critical machine) that processes the waste to dispose it of (2nd operation). There is no storage capacity, thus once a silo accepts the load it cannot be released until the waste is completely transferred to the mixer and the processing ends. Subsequently, the sequence strictly begins at the 1st stage and finalizes at the 2nd with no preemption (see Fig. 5.6).

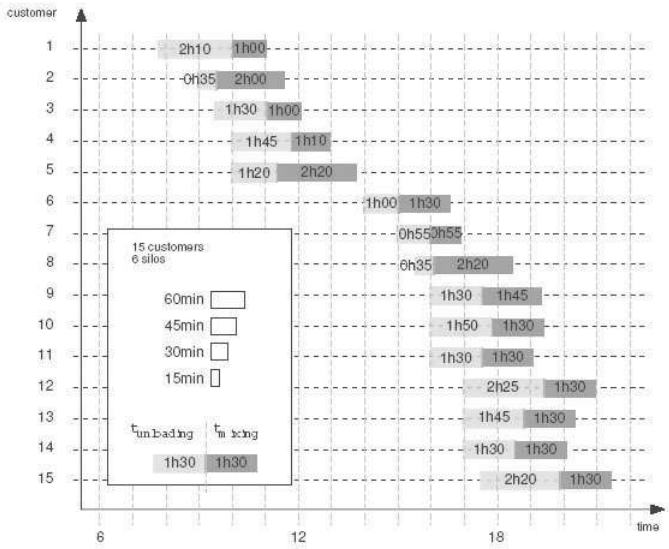


Figure 5.7: A real instance of ready and processing times for each truck [180].

The problem raises two objectives: on the one hand the maximum completion time or *Makespan* is to be minimized to save resources. On the other hand, transportation is provided by an external entity that penalizes delays in the unloading operations beyond the expected arrival times. Hence, the *Total Waiting Time* (TWT) between the ready (arrival) times and the starting times of unloads is to be minimized as well.

The optimization of these two objectives was tackled in [59] with a tailored version [59] of the Multiple-Objective Simulated Annealing algorithm (see algorithm 6 in fig. 3.17) solving two real instances (see fig. 5.7) corresponding to a waste treatment problem, along with several random instances . The procedure was built using MNEH, an heuristic proposed by Nawaz, Enscore and Ham and modified by Martínez [133] (see [59] for more details). Contrary to what was expected, the output is featured by few or even only one non-dominated point and a large number of equivalent schedules, no matter what instance is chosen (see fig. 5.8) [59]. In other words, there are several different schedules that show the same Makespan and the same Total Waiting Time: different alternatives for the DM that yield the same performance in terms of objectives values.

The results obtained in [59] show that the algorithm is capable of solving the optimization problem, so one could say that the problem itself does not represent a technical challenge. Nevertheless, the described situation raises an unfavourable panorama for decision-making: how does one select the right alternative from the numerous set of optimal equivalent schedules? Moreover, are all those really equivalent?

Incorporation of uncertainty

After trying to solve the RCb flowshop problem with all the available (classical and heuristics) tools, it was necessary to consider a different perspective: prior

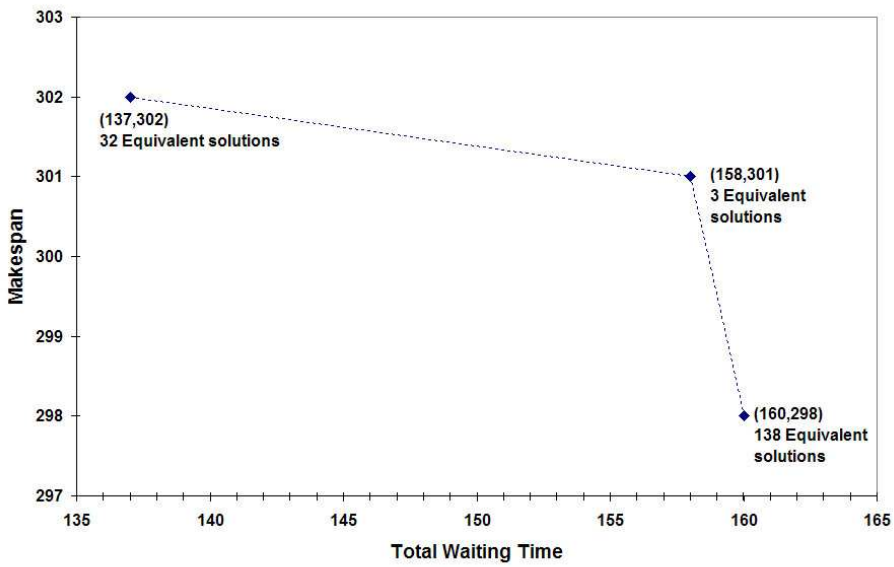


Figure 5.8: Example of using MOSA and MNEH over a real instance (in minutes) [59].

analyses were performed neglecting uncertainty for the sake of simplicity, while the following study proposes to consider some uncertain aspects of the original decision-making problem through the AUREO. In order to do so, the next step is to realize the new conditions that derive from working with uncertainty and the DM's attitude towards them. The results are:

- The amount of information about the problem is scarce and the DM is averse to give detailed information or make assumptions.
- Online scheduling will not be provided, hence the physical sequence will be strictly respected.
- Trucks may arrive earlier or later regarding the expected time.
- Delays in processing times will not be considered.

5.3.3 AUREO: Stage 1

Once again, we apply the AUREO following the principles proposed in fig. 4.17 in a convenient order:

1. Analyze the model

The problem consists in a two-stage shop composed of six silos at the first stage and a mixer at the second one. All the silos are considered identical machines so the processing times are independent of the silo but relies upon the truck load. The same stands for the mixing time. Hence, the processing times are

parameters to the model. Likewise, the arrivals of trucks are expected to happen according to a distribution that relies upon several factors like the time when the waste is available, the time to load the truck with the waste and the route to pick it up along the region. Since the DM has no inherence on such arrival times, they turn out to be parameters to the model as well. Fig. 5.7 shows an example of a real instance consisting of $|J| = 15$ trucks, each one with a ready time (the coordinate of the lower bound of the intervals), an unloading time (size of the light shaded section) and a mixing time (size of the dark shaded section).

Decisions have to be made on the starting times of each job at every stage and how the machines will be allocated amongst the jobs. The decision vector therefore constitutes the schedule itself. Let $S(\mathbf{r}, \mathbf{u}, \mathbf{m})$ denotes a schedule defined from the aforementioned parameters, namely vector $\mathbf{r}$ of ready times, vector $\mathbf{u}$ of unloading times and vector $\mathbf{m}$ of mixing times. A schedule is a collection of vectors of the type $(J_j, M_i, t_{j,s})$, where a machine M_i is allocated to execute job J_j beginning at starting time $t_{j,s}$, with indexes j, i, s denoting jobs, machines and stage respectively, such that the following constraints stand:

- Jobs cannot be executed before their ready times, thus $t_{j,1} \geq \mathbf{r}_j$ (at 1st stage) and $t_{j,2} \geq t_{j,1} + \mathbf{u}_j$ (at 2nd stage).
- Preemption is not allowed and storage is not possible, thus the completion time equals $t_{j,1} + \mathbf{u}_j$ at 1st stage and $t_{j,2} + \mathbf{m}_j$ at 2nd stage.
- Finally, to ensure sustained service, the makespan should not exceed 24 hours.

The quality of the schedule is measured through its makespan and the TWT. The original model studied in [59] can be formulated as: find the feasible schedule $S(\mathbf{r}, \mathbf{u}, \mathbf{m})$ such that

$$S(\mathbf{r}, \mathbf{u}, \mathbf{m}) = \arg\{\min \mathrm{TWT}(S(\mathbf{r}, \mathbf{u}, \mathbf{m})) \wedge \min \mathrm{Makespan}(S(\mathbf{r}, \mathbf{u}, \mathbf{m}))\}$$

where:

$$\mathrm{TWT}(S(\mathbf{r}, \mathbf{u}, \mathbf{m})) = \textstyle\sum_j (t_{j,1} - \mathbf{r}_j)$$
$$\mathrm{Makespan}(S(\mathbf{r}, \mathbf{u}, \mathbf{m})) = \max_j \{t_{j,2} + \mathbf{m}_j\}$$

2. Check for input uncertainties

In the presence of uncertainty and according with the assumptions mentioned in sec. 5.3.2, part of the input turns out to be uncertain. Notice that the arrival times are assumed now to vary around the expected arrival time, which means that vector $\mathbf{r}$ is now subject to epistemic uncertainty. Indeed, one can presume the arrivals are random and follow some PDF, a situation of type I or aleatory uncertainty. However, the actual level of information does not allow to sustain such assumption. Hence the analyst is forced to realize an actual presence of epistemic uncertainty and a presumable presence of randomness that would suggest mixed uncertainty.

On the contrary, vectors $\mathbf{u}$ and $\mathbf{m}$ are considered free of uncertainty due to the assumption of no variation in processing times. Likewise, the decisional vector, i.e. the schedule comprises no uncertainty since the assignments $(J_j, M_i, t_{j,s})$ made by the heuristic MNEH [133, 59] are crisp.

3. Search for uncertain objective functions

In scheduling the presence of uncertainty is often addressed by means of online rescheduling that allow adjusting the pending activities according to the events occurred not considered in the original schedule (see [32]). The absence of online scheduling, however, entails in practice a strict obedience to the physical sequence of the schedule. It means that the order of precedence from the original schedule is to be kept, even when some variations on vector $\mathbf{r}$ take place. The consequence is that the TWT and the makespan could change along the time in response to eventual earlier or later trucks arrivals. Thus, as the input uncertainty is propagated, the objectives values, measured as crisp quantities in the original problem, become uncertain.

Nevertheless, the key question is not about the outcomes, which are obviously uncertain, but about the strategy to propagate the input through the model. In the case of the application described here, Monte Carlo simulation was chosen due to the use of the heuristic MNEH and the lack of information about how $\mathbf{r}$ varies.

The objectives are now to be estimated by generating N samples of the uncertain vector as $\mathbf{r} + \delta_r$, where δ_r represents the uncertainty associated to $\mathbf{r}$ according to the definitions introduced in sec. 4.3.3. For every sample of ready times -denoted $\delta_{r,k}$- and with the physical sequence of the schedule under consideration, the TWT and the makespan are to be calculated. Afterwards the representative values are assessed. For that matter, the expected value was selected to represent the uncertain objectives, viz.:

$$F_1 = \frac{1}{N}\sum_{k}^{N} \mathrm{TWT}(S(\mathbf{r} + \delta_{r,k}, \mathbf{u}, \mathbf{m})) \tag{5.4}$$

$$F_2 = \frac{1}{N}\sum_{k}^{N} \mathrm{Makespan}(S(\mathbf{r} + \delta_{r,k}, \mathbf{u}, \mathbf{m})) \tag{5.5}$$

Notice that, no matter what representative values be chosen, the use of Monte Carlo simulation entails that different evaluations of the objectives will produce different outcomes. Hence, the implementation should consider this new source of uncertainty during the ranking procedure.

The final robustness-seeking program

Let us now analyze the type of formulation that derives from the available information making use of table 4.9. According to the statements of sec. 5.3.2, the original problem does not consider uncertainty, thus the uncertainty δ_r is not defined and the reluctancy of the DM to provide more information limits the analyst to investigate it. In consequence, programs of Class 1 are not applicable according to the principle of the AUREO. Likewise, Class 2 programs are either unapplicable since the DM did not provide precise information to define $I(\cdot)$ constraints. Thereby the problem is one of Class 3 robustness.

In order to solve the problem, it is necessary to make the problem collapses into a Class 1 or Class 2 robustness-seeking program. For that matter, a maximal deviation of 10 minutes around the expected arrival times was proposed to the DM as a reasonable point of departure to asses the maximal inner box. Then, using the results obtained in [59] as a reference, the DM was asked to point

out a threshold of indifference, i.e. the maximal negligible difference between two values, instead of defining the $I(\cdot)$ constraints explicitly. Such a value was set to five minutes, thus any variation smaller than this value is not considered significant in terms of decision-making.

The information obtained this way along with the original constraints that comprise the original formulation made possible to define a Class 2 program of the type: *find the schedule S with with the greater MIB that minimizes F_1 and F_2 such that the value $|F_1 - TWT(S)| \leq 2$ min and $|F_2 - Makespan(S)| \leq 2$ min (performance constraints) and $\overline{F_2} \leq 24$ hrs (operability constraint, the remainder constraints are considered when adapting schedule S to match $\mathbf{r}+\delta_r$).* More details on this models are given in the next section

5.3.4 AUREO: Stage 2 and results

Implementation

The computational tool developed in [59] to solve the original problem was modified to allow the adaptation of the physical sequence of the schedules obtained using MOSA and MNEH to the different realizations of $\mathbf{r} + \delta_r$ in order to estimate F_1 and F_2. From here the technical issues to address were how to perform the simulation, how to handle the uncertainty induced by such simulation during the ranking and how to calculate the MIB of each schedule.

Since the heuristic is time consuming, the Monte Carlo simulation was restricted to 30 samples (the minimum recommended in [165]) with the verification of the confidence interval size as an additional acceptance criterion. Every ready time was sampled using a uniform distribution. Afterwards, in order handle the uncertainty induced by the simulation procedure about the real value of the expectancies of TWT and makespan, the following procedure was proposed: the objective space was divided into boxes inspired on the concept of (additive) ϵ-dominance surveyed in sec. 3.3.3, but giving ϵ an interpretation of threshold of indifference, which means that every point that lies into the same box is considered the same (see fig. 5.9). Thus, the numerical variations between objectives of distinct schedules carried out by the simulation are disregarded in some amount.

Boxes were built of side one minute, hence the performance constraints are verified if the expected value of a schedule lies at most one box away from the Pareto approximation front. In practical terms, the DM is able to accept numerically suboptimal solutions if and only if those solutions do not deviate from the reference more than two minutes.

The use of boxes has here a particular purpose. Notice that if too many different schedules map the same point in absence of uncertainty, it makes sense to expect that many points gather in clusters around the same region of the objective space. In principle, all those solutions are potential members of the final optimal outcome, and it is the size of their MIB the criterion to filter such clusters. Having boxes, the many candidates can be reduced to (at most) one solution per box for the non-dominated boxes.

The archiving strategy could filter dominate solutions in a steady-state fashion or might be performed in two stages, the first one based on finding as many solutions complying the constraints as possible, and a final filtering stage in terms of the MIB after the MOEA finished. The first procedure amounts to

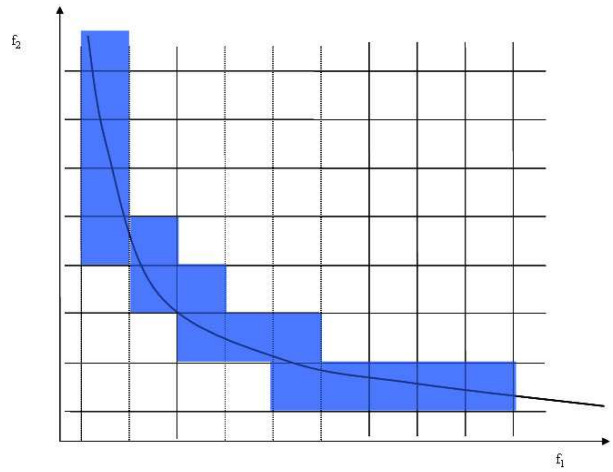


Figure 5.9: Example of ϵ-dominance used as threshold of indifference: solutions located inside a box are considered the same in terms of quality.

solving a three-objective program, namely F_1, F_2 and MIB, whereas the second resembles a lexicographic procedure, optimizing F_1 and F_2 simultaneously and MIB in the second step.

For the assessment of the MIB a genetic algorithm was proposed: the chromosome is a vector $(t_1, t_2, \ldots, t_j, \ldots, t_{|J|})^t$ such that the uncertainty $\mathbf{r} + \delta_r$ is represented as $(r_1 \pm (10 \text{ min} + t_1), \ldots, r_j \pm (10 \text{ min} + t_j), \ldots, r_{|J|} \pm (10' + t_{|J|}))$ and $t_j \geq 0$ min$\forall j$, whereas the fitness function is MIB $= \prod_j t_j$. In words, the goal is to determine the maximal allowed deviation above ten minutes for every ready time that the schedule can absorb without violating the constraints. For that matter a population of ten individuals is recombined along ten generations and the MIB assigned to the schedule is that of the fittest individual.

Results

Fig. 5.10 shows the result obtained when the candidates are filtered in two stages. Original outcomes are represented by dark rhomboids while the expected value of the robust schedules appear like squares (with their confidence interval).

Fig. 5.11 reports the results of solving the problem as a three-objective problem. Notice the solutions at the right bottom violate the performance constraint on TWT, however their representation is useful to show how two schedules can nearly overlap each other and show a very different level of robustness at the same time. It means that, both schedule absorb well a variation of ten minutes in the ready times, but only one of them is useful if the such variations in the arrival times increase. Likewise, the cluster at the upper left show several schedules with similar performances, all belonging to the same hyperbox. All those solutions are considered non-dominated in a three-objective space, but that with MIB 0.149 will be chosen for decision-making purposes as the more robust schedule, since it complies all the constraints and also have have ability to absorb late arrivals amongst the valid schedules.

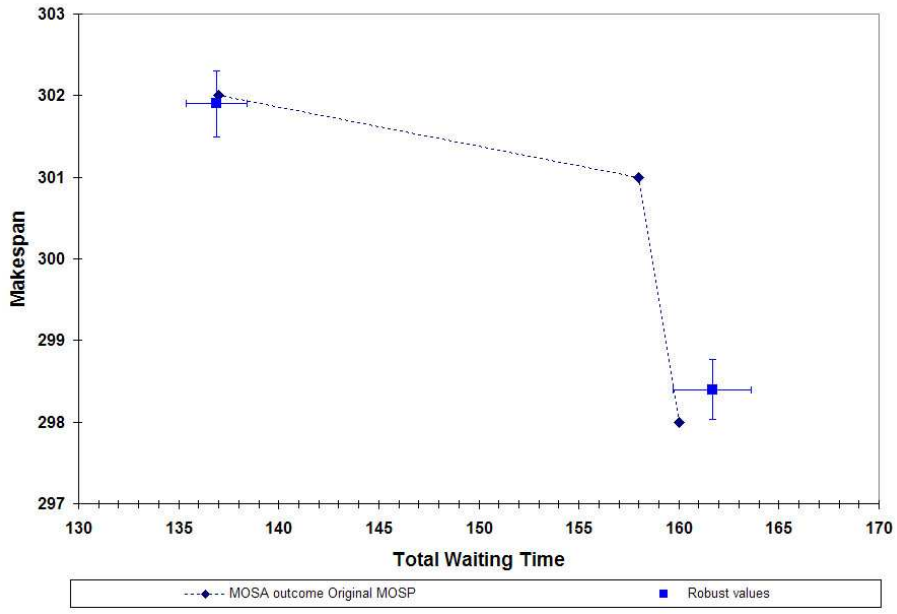


Figure 5.10: Results filtering the optimal solutions in two stages (in minutes).

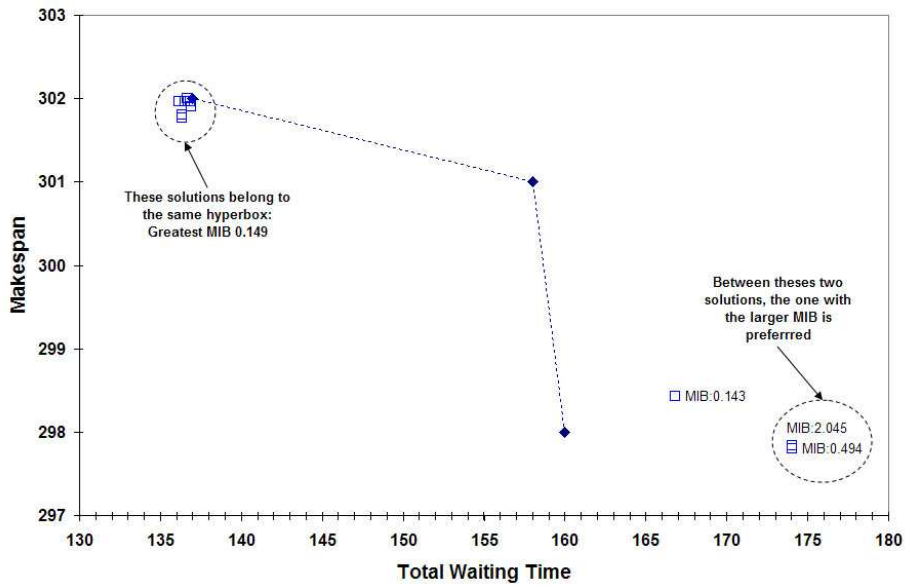


Figure 5.11: Results of solving the problem as a three-objective program (in minutes).

These results exemplify the many aspect of a Class 3 problem and how the AUREO can be used to assist analyst and DM in solving real life problems.

5.4 Summary of the chapter and concluding remarks

In this chapter we have described the application of the AUREO to robust design problem drawn from the Reliability Engineering field in the realm of nuclear power plants safety. The analysis indicated a Class 2 formulation which provide acceptable level of variation in the components specifications comprising a safety system. Results of that kind are of applicable in real life, for example to know a priori how many alternative components can be used as replacements in case that the original components are not available.

While the analysis of uncertainty did not contribute very much since the original problem were one of robust design, the efficiency analysis performed in the stage 2 of the AUREO method resulted in a noticeable enhancement of the MOEA due to the introduction of the percentage representation.

Afterwards the AUREO was applied on a real multiple-objective scheduling problem whose analysis indicated a Class 3 robustness problem. In order to solve it the DM was asked to provide more information to define wether a Class 1 or a Class 2 program. However, the reluctancy of the DM to define clearly performance constraints and the impossibility to elicitable properly the input uncertainty made necessary to recourse to different strategies. The DM was asked about the size in time units considered indifferent. With this information it was possible to define performance constraints and the solve the problem as a Class 2 robustness-seeking program.

Chapter 6

AUREO + MCDM in Vulnerability Analysis

"The good fighters of old first put themselves beyond the possibility of defeat, and then waited for an opportunity of defeating the enemy. To secure ourselves against defeat lies in our own hands... Thus the good fighter is able to secure himself against defeat, but cannot make certain of defeating the enemy. Security against defeat implies defensive tactics."

Excerpts of *The Art of War* (Chapter IV)
by Sun Tzu.

6.1 Introduction

Modern societies become more and more complex as their reliance on interconnected systems increase. On the one hand, the technological nuance of modern life has risen our quality of life but paradoxically, our strong dependance on technology makes us more vulnerable as it opens the door to emerging types of hazard, as computer viruses or electronic swindles amongst others. On the other hand, the growth of population demands more from classical services, posing new challenges to transportation, health and food services, social safety and security, etc.

Natural disasters and terrorist attacks occurred during the last years have put concepts like vulnerability, survivability and resilience in the target of public debate. The scientific community has made thorough reflections from both theoretical and speculative [61, 69, 114] and practical viewpoints [4, 62, 68], and has come to recognize the tremendous challenge of modelling, assessing and handling vulnerability, security and safety of nowadays critical infrastructures, due to their complexity and to our level of uncertainty regarding these new hazards [94].

In this chapter we bring the contributions of using AUREO to vulnerability analysis, more specifically to the vulnerability assessment of complex networks problem introduced in [233, 232]. The present discussion remarks the use of our information perspective and the research question emerged from the use of AUREO rather than other interesting facets related to risk analysis, network analysis or the details of the simulation tool used all along the research.

Nonetheless, the chapter is of interest for researchers and practitioners interested in the AUREO or in vulnerability and survivability analysis as well.

6.2 Statement of the problem

The concept of critical infrastructures is nowadays subject of many studies, discussions and debates in scientific, academic, politic and social spheres. Even when the perception of criticality may change from time to time and amongst societies, every infrastructure of such is "*understood as vital for economic performance and social welfare*"[19, pg. 859]. Hence, the protection of critical infrastructures is a top priority task.

The importance of critical infrastructures makes societies especially sensitive to partial or complete incapacitation of such infrastructures that may come from internal or external sources. While reliability engineering and risk analysis provide tools and procedures for estimating, preventing and handling many events that occur at random, emerging risks like intentional attacks constitutes a challenge due to the presence of "*a malevolent intelligence directed toward maximum social disruption*"[4, pg. 361].

Many efforts have been devoted in recent years towards a new understanding of the way to analyze safety and security of critical infrastructures and how to protect them against natural disasters and/or intentional attacks [18, 73, 109, 120, 121, 122]. Moreover, amongst the current critical infrastructures (see [19] and references therein for examples), networks are perhaps the structure that prevails. Hence, many researchers have recently focused on the protection of complex networks against terrorist or intentional attacks to e.g. the water supply network [191, 206] or the electricity network [82, 83, 94].

Complex networks are susceptible to at least two modes of attack. On the one hand, such attack may be directed to damaging the infrastructure itself by impacting its components or, on the other hand, an antagonist could take advantage of the infrastructure as a vector of propagation of a hazard to the people and the environment (e.g. a contaminant, a poison, a virus, injected for transmission into a water supply network). Propagation modelling is therefore a quite relevant task for providing the necessary information to devise effective countermeasures to the attacks.

In [233, 232] E. Zio and C. Rocco proposed the use of cellular automata [217] and Monte Carlo simulation [130] to investigate the dynamic of attacks propagation though networks, thereby allowing to assess some indexes of impact based on such dynamic. The early stage of that research was based on single-objective analyses and focused on the development of the simulation tool, thereby considering only environmental uncertainties. Once the simulation tool was fully developed, the research interest widened to consider the optimal allocation of protective countermeasures. Since the efficacy of such countermeasures depends not only on the network but on the antagonist's actions which remain, at least in some degree, uncertain in real life, the AUREO perspective was embraced to investigate the possible models to handle uncertainty and the evolutionary alternatives to solve them.

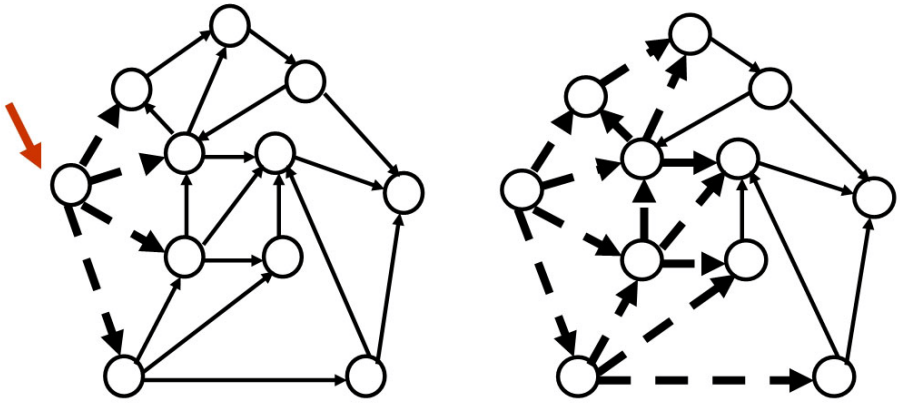


Figure 6.1: Example of propagation through networks: the attack begins at the pointed node (left) and propagates to adjacent nodes through links (right) (network taken from [225]).

6.2.1 The decision-making problem

A generic network $G(N, E)$ is composed of a set $N = \{n\}$ of nodes linked by a set of edges $E = \{e_{i,j}\}$, each of which connects two generic nodes n_i with n_j in a directed or undirected manner [189]. This abstraction can be applied to model numerous types of interconnected systems. In particular the model proposed in [233, 232] associates nodes to sets of entities (beings or assets) that can be simultaneously damaged by an attack and can propagate it, and edges to propagation channels. Several features can be considered in a model of this kind, like the edges' transmission capacity, the strength of the hazard propagated or the existence of different modes of attack amongst others. However, the basic considerations are the number of entities at each node, the number of nodes attacked by the antagonist and the time to propagate their hazardous effects through links.

Consider the example sketched in fig. 6.1; the pointed node (left) is initially set on by the antagonist with the consequent impact on the entities associated to such node. Afterwards, the propagation begins through all the edges adjacent to the point of attack in such a way that, after a period of time, all the neighbour nodes are completely reached and affected (right). The propagation continue with a similar pattern but now from the many newly affected nodes until the wholly of them in the network are hit by the hazard.

From a defender point of view, the possible strategies are to keep the antagonist from performing the attack or to implement a set of countermeasures to neutralize or mitigate its impact once it is performed. The decision-making problem studied here corresponds to the last option. Consider the decision variables of the defender: there is a set of possible countermeasures $P^c = \{p_i^c\}$ that can be implemented on each node, each one with different cost $c(p_i^c)$ and different protective effect. The defender will try to minimize the impact on the network subject to his amount of available resources R_D. On the other hand, P_N (the powerset (see sec. 2.2.1) of the set of nodes N) constitutes the set of all possible combination of nodes that the antagonist might attack simultaneously.

The antagonist is assumed to be rational so the selection of targets would try to maximize the impact subject to their amount of available resources R_A.

The defender problem can be formulated as *find the assignment of protections* $\{(n, P^c(n))\}$ *from* P^c *to the nodes of* N *that minimizes the impact on the network, s.t.* $\sum_{n \in N} \sum_{p_i^c \in P^c(n)} c(p_i^c) \leq R_D$. Naturally, the pattern of attack is uncertain, even when some information about the preferences of the attacker and their resources R_A can be estimated through intelligence gathering.

The key issue is to define the expression of impact to minimize. For instance, in [109, 120, 121, 122] the authors aim at minimizing a utility function which can be, e.g. the expected damage. By contrast, in [233, 232] E. Zio and C. Rocco proposed two raw indexes of impact, viz.:

Time To Reach All network Destination nodes (TTRAD) is the time to affect the whole network. This measure is similar to the 'all-terminal network' [30] evaluation, often performed within network reliability analysis. For this index, shorter times indicate higher impacts.

Average number of persons affected (ANPA) or Average number of entities affected (ANEA), is the average speed of affecting people or entities during the propagation. It is proportional to the area behind the cumulated curve of people affected by the propagation of the attack in function of time. For this index, larger figures indicate higher impacts.

Both measures bear resemblance, however they are not completely equivalent, but their relation depends on the distribution of entities around the point of attack. Thereby, if most of them are gathered near the node where the attack begins, the cumulated curve of people affected saturates fast and the ANAP or ANEA value is high, even when the TTRAD be high too.

If the impact is assessed according to the principles of probabilistic risk analysis, the goal might be formulated as to minimize an expression of the type

$$R = p \times q \tag{6.1}$$

where p denotes the probability of an event and q its consequences. Hence, to reduce the risk R of attacks, the defender should find the protection pattern that minimize the overall risk. Since the analysis should consider at least $|P_N|$ scenarios, then the expected risk of attack for a particular set of implemented protection $\{(n, P^c(n))\}$ may be preliminarily assessed as $\sum_{a \in P_N} p(a)q(a)$, where a denote the scenario of attack, and $p(a)$ and $q(a)$ its probability and its consequence respectively (cf. [5]). Such an expression of risk is compatible with the utility function approach adopted in [109, 120, 121, 122]. However, the previous expression relies upon a single scalar quantification of consequences, situation that is easily handled when one has a unique descriptor of impact but, e.g. in the case of the two indexes used in [233, 232] (ANPA and TTRAD), could require additional treatment, like aggregation with some kind of averaging procedure, e.g. ordered weighted averaging aggregation operators (OWA) [222]. Moreover, other approximations to the assessment of risk that allow many incommensurable descriptors of consequence are possible, like the use of 'triplets' to describe risk [97, 96] as in [62]. As we can see, the discussion about the proper way to assess risk facing intentional attacks is far from being settled and, although interesting, it lies out the scope of AUREO; nevertheless, it suffices to

remark that, in terms of applying AUREO, the preceding words indicate that the suitability of model remains uncertain.

6.2.2 Instances analyzed

Let us formalize now the statement of the problem according to the problem introduced in [233, 232]. Consider a network $G(N, E)$ compose by a set N of nodes n and a set E of edges $e_{i,j}$. Every node has associated a number of entities (assets or beings) that can be affected either by direct attack on the node or indirectly via propagation through the network. Likewise, every edge has a velocity of transfer whether for directed or undirected nodes; same velocity for both directions in the latter case. In general, all these quantities, namely numbers of entities and velocities of transfer, are assumed to be aleatory.

Resources R_D and R_A are assumed to be limited and for simplicity, but without a loss of generality, the constraints are expressed in terms of the number of allowed (defensive and offensive) actions rather than in terms of the sums of costs described in the previous section. Additionally, only one-off attacks and no cascade effects are studied.

Hence, the problem is to find the optimal assignment $\{(n, P^c(n))\}$ of protections to nodes that minimizes the overall impact, in terms of the raw impact indexes ANPA (or ANEA) and TTRAD.

6.2.3 The simulation tool

Cellular automata are mathematical models of dynamic systems. The dynamic of cellular automata unfolds at discrete time steps on a discrete lattice of cells, typically assumed homogeneous (all cells bear the same properties) and with fixed updating rules that assign to each cell a new value which is function of the current value of such cell and its neighbourhood [217]. The tremendous potential of CA has been exploited up to the point of defining Turing machines with just a few rules.

In [157] the authors introduced the definition of cellular automata adapted to analyze basic features of network reliability. Later on in [160], the automata were merged to Monte Carlo sampling to yield a hybrid methodology that makes possible the analysis of advanced network reliability problems.

Basics of cellular automata [161, 234]

Consider for example a three-dimensional cellular state space; the state at the discrete time t of the generic cell ijl, of coordinates x_i, y_j, z_l with $i, j, l \in Z$, is described by the state variable $s_{ijl}(t)$. Each cell of L is a finite automaton which can assume one of a finite number of discrete values in a local value space $S \equiv \{0, 1, 2, , k-1\}$.

The generic cell ijl interacts only with a fixed number n of cells that belong to its predefined local neighbourhood N_{ijl}. At the next discrete time $t+1$, the cell ijl updates its state $s_{ijl}(t+1)$ according to a transition rule $\Phi : S^n \to S$, which is a function of the state variables at time t of the n cells in N_{ijl}, viz. $s_{ijl}(t+1) = \Phi[s_{rsp}(t), rsp \in N_{ijl}]$. Notice that the homogeneity assumption implies that the functional form of the rule is assumed to be the same everywhere in the cellular state space, i.e. there is no space index attached to Φ. Differences between what

is happening at different locations are due only to differences in the values of the state variables of the local neighbourhood, not to the update rule. The rule is also homogeneous in time. One 'iteration step' of the dynamical evolution of the automata is achieved after the simultaneous application of the rule to each cell in the lattice L. The temporal evolution of this cellular automata is obtained by:

1. specifying the finite size of the lattice L;
2. specifying the boundary conditions;
3. specifying the initial condition $\vec{s}(0) = [s_1(0), s_2(0), \ldots, s_M(0)]$ and,
4. simultaneously applying the rule Φ to each of the L lattice cells, in an iterative manner.

Outline of the tool

The type of problem addressed by the tool in [232, 161, 234] considers that the activation of a node is delayed by the time required to propagate the attack from node to node. Indeed, a cell is activated if it is connected to and receives input from at least one active cell or node in its neighbourhood. When accounting for the hazard propagation process, the cell activation concerning the hazard also depends on the time required to propagate the attack: the arrival time of the propagated attack is determined as the sum of the current time plus the time delay. If several nodes can propagate the attack to a given node, the arrival time of the attack at such node is determined by the minimum of the times of propagation from all connected nodes in its neighbourhood.

Assume now that the generic connecting element (edge) ij from node i to j can be in two states, active ($w_{ij} = 1$) or passive ($w_{ij} = 1$). The ij edge state variable w_{ij} defines the 'operational' state of the edge. Initially w_{ij} holds for all state variables. As soon as node j is reached by the attack, the state of w_{ij} changes from 0 to 1, for $t = t + TD_{ij}$. The transition rule governing the evolution of the generic cell j consist of the application of the following rule:

$$s_j(t) = [w_{pi} \vee w_{qi} \vee \ldots \vee w_{ri}], \quad p, q, \ldots, r \in N_i \tag{6.2}$$

The basic algorithm proceeds as is described in fig. 6.2. To consider the evaluation of the time to reach every node in the network, an additional node is introduced that is activated only when all nodes are activated. The network under analysis is assumed to be operational, so there is at least one spanning tree. [161, 234]

6.3 Application of the AUREO

In this section we exemplify the AUREO by means of two problems of optimal allocation of protections.

Algorithm 7 Simulation tool of [161, 234]

1: $t = 0$ /* Initial time step */
2: Set all the cells state values and $w_{ij}(t)$ to 0 /* (passive) */
3: Set the source node $s_s(0) = 1$ and w_{si}(source activated and edges from s activated at $t = t + TD_{si}$)
4: Update each cell state by means of rule eq. 6.2 and update all exiting edges from each cell
5: $t = t + 1$
6: If the destination node $s_D(t) = 1$ then stop (destination activated), else go to 4

Figure 6.2: Algorithm 7: Simulation tool based on cellular automata and Monte Carlo simulation introduced in [232, 161, 234].

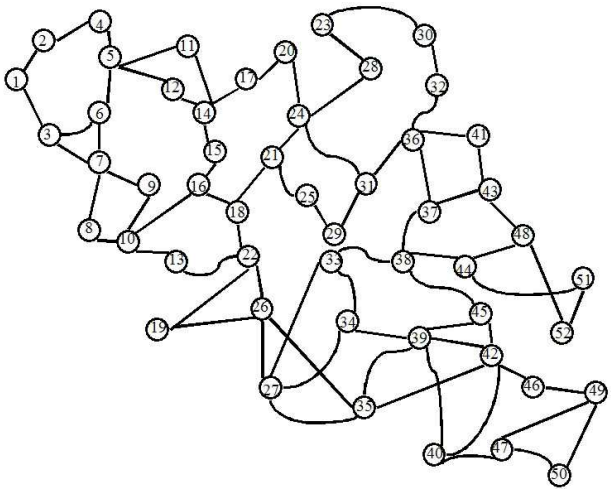


Figure 6.3: Topology of Network 1: a real telephone network [128].

6.3.1 Description of the problems

The first problem studies the protection of a small network composed by 52 nodes and 73 bi-directional edges, whose topology corresponds to a real telephone network topology [128] (fig. 6.3); this network shall be called Network 1 hereinafter. The second one studies a bigger network of 332 nodes and 2126 bi-directional links whose topology was taken from the US airports network (source: http://vlado.fmf.uni-lj.si/pub/networks/pajek/data/gphs.htm) (fig. 6.4); this network shall be called Network 2 hereinafter. It is important to remark that only the topologies relate the problems studied here to the original (telephone and airport) networks and no other aspect connected to such original systems were considered.

Each problem is characterized not only by its topology but its parameters. Both networks are composed by undirected links, so the propagation of the attack at any node spreads immediately to the neighbors. Nevertheless the time to accomplish such propagation is considered to be variable from link to link. In the first case, time delays in the propagation through the network are randomly selected, without loss of generality, using a discrete uniform distribution U(0,10). This assumption stands for the second network as well. On the other hand, the

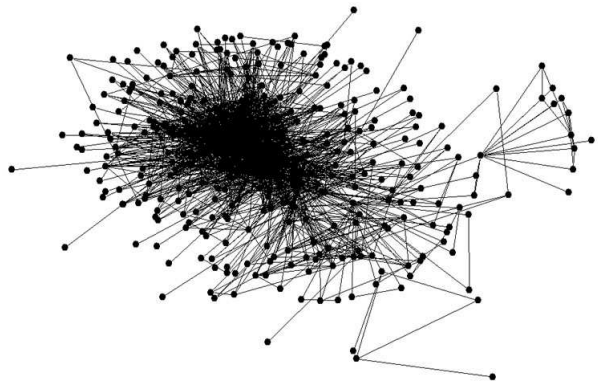


Figure 6.4: Topology of Network 2: the US airports network (generated with Pajek [12]).

Node	Min	Max	Node	Min	Max	Node	Min	Max	Node	Min	Max
1	17	32	14	17	35	27	15	39	40	11	38
2	16	36	15	12	37	28	10	40	41	13	32
3	11	33	16	20	40	29	19	35	42	14	34
4	12	38	17	12	31	30	17	37	43	20	39
5	19	38	18	15	31	31	14	37	44	18	32
6	19	33	19	18	34	32	17	31	45	15	34
7	12	30	20	13	30	33	10	38	46	20	33
8	14	31	21	15	33	34	17	40	47	19	36
9	19	40	22	10	31	35	18	39	48	16	38
10	15	38	23	18	39	36	11	33	49	11	31
11	13	39	24	16	32	37	11	37	50	16	37
12	13	39	25	14	40	38	13	35	51	18	36
13	14	34	26	15	36	39	16	34	52	13	35

Table 6.1: Parameters for the number of people affected by an attack to a single node in the network shown in fig. 6.3.

numbers of persons affected at the different nodes were also assumed to bear a uniformly distributed fashion. The ranges of variation for the first problem are shown in table 6.1, whereas for the second one the population at every node varies from 10 to 40 (U(10,40)). Five hundreds Monte Carlo evaluations are considered to assess the consequences of any particular attack.

The study is constrained to consider only one-off attacks, i.e. cascade attacks are not tackled in this study. Additionally, it is assumed that the resources are limited for attacking and protecting one single node of the network. On the other hand, we study only one type of protection that allows the node to be affected but curtails the propagation. Changes in the features just mentioned above render the problem more difficult and their study is left to future stages of the research.

In the following subsections, both problems are studied applying AUREO according to the principles introduced in chapter 4.

6.3.2 AUREO: Stage 1

Let us consider step by step the principles proposed in fig. 4.17 but in a different order:

1. Analyze the model

The propagation dynamic and its effects are assessed by means of a hybrid strategy that can be seen as a black box function. For a given network $G = (N, E)$ the defender or DM can only choose where to allocate a protection. Thus, the decision vector $\mathbf{x} = (x_1 = i)$, $i = \{1, 2, \ldots, |N|\}$ is equivalent to the index of the node n_i to protect due to the initial assumption of only one type of protection and only one node to protect. In general, the decision vector $\mathbf{x}$ is a matrix of assignment of protections to nodes that can be as complex as one wishes; for instance one can consider assigning several protections to the same node or even a protection shared by many nodes [120, 121, 122].

On the other hand, the initial assumptions reduced the number of points of attack to one, but the number of scenarios of attack can be very complex, so much so that in general the number of one-off attack scenarios is $|P_N|$. No matter what the assumptions about the number of attacks are, the node where the attack begins remains uncertain for the DM, although some information about the intentions of the antagonist could be collected by intelligence tasks. The point of attack is then an epistemic uncertain parameter to the model. In contrast, the propagation times through edges and the numbers of entities to be affected are aleatory quantities. Hence, the vector of parameters $\mathbf{p}$ for the problems under study comprises one point of attack, the propagations times and the numbers of entities to affect ($1 + |N| + |E|$ components).

Consider now the model of impact. As we discussed in sec. 6.2.1 the model of impact can be anyone meant to minimize risk. Some options are:

$$F(\mathbf{x}, \mathbf{p}) = (f_1 = \text{ANPA}(\mathbf{x}, \mathbf{p}), f_2 = \text{TTRAD}(\mathbf{x}, \mathbf{p}))^t \tag{6.3}$$

$$F(\mathbf{x}, \mathbf{p}) = p(n_i) \times (w_1 I_{\text{ANPA}} + w_2 I_{\text{TTRAD}}), \sum_j w_j = 1 \tag{6.4}$$

$$F(\mathbf{x}, \mathbf{p}) = p(n_i) \times (I_{\text{ANPA}}, I_{\text{TTRAD}})^t \tag{6.5}$$

Eq. 6.3 corresponds to the case of assessing impact using the raw measures ANPA and TTRAD directly. The ANPA is to be minimized whereas the TTRAD is to be maximized. Since the point of attack is unknown, the optimization should be performed over the whole range of scenarios. It means that for every protection there is a range of impacts and a criterion for comparison should be adopted. This model has the advantage that it does not require additional information like the probabilities of attacking every node, although the point of attack is still an epistemic parameter. In contrast, risk models like those in eq. 6.4 and eq. 6.5 require not to know the point of attack but the probability of performing it. If we adopt a model of such, the probabilities may be assumed to be epistemic uncertain parameters whose level of uncertainty is prone to be reduced with proper intelligence gathering.

Eq. 6.4 corresponds to a linear aggregation of the indexes of risk associated to the ANPA and the TTRAD. Transforming the latter measures to indexes is necessary to keep coherence with the notion of risk which is directly proportional

to impact. Hence, since the less TTRAD the greater the impact, a possible index should be $I_{\text{TTRAD}} = \max_N\{\text{TTRAD}\} - \text{TTRAD}$ where the maximum is calculated regarding the TTRAD of attacking a particular node n_i over the whole set nodes N. If we want to avoid aggregation we can simply adopt a multiple-objective model instead, as in eq. 6.5.

Notice that level of uncertainty regarding the probabilities of attack means that $p(n_i)$ varies within $[0 \leq \underline{p}(n_i), \overline{p}(n_i) \leq 1]$. Thus, at the maximal level of uncertainty the risk associated with a particular scenario ranges from 0 to the impact in terms of the TTRAD or the ANPA. Hence, a class 1 robustness model yields a comparison of intervals according to the criteria of sec. 4.1.3. Amongst the applicable criteria, the worst case analysis seems the best option for protection and defense purposes [189] since the initial premise of "a malevolent intelligence directed toward *maximum* social disruption"[4, pg. 361] (italics are ours) entails that the greatest impact the better for the antagonist. Now, observe that concerning the model, adopting the worst case analysis for eq. 6.5 is equivalent to solve eq. 6.3 using worst case comparison if $[0 = \underline{p}(n_i), \overline{p}(n_i) = 1]$. This is the case studied here ($p(n_i) \in [0, 1]$).

Moreover, one can prefer to establish level of acceptance of risk rather than making assertions about $p(n_i)$. In other words, we can consider d_p as undefinable due to the uncertainties associated to the point of attack (be it its location or its probability), leading us to a class 2 or class 3 robustness problem (see table 4.9), depending on our capacity to define $I(\cdot)$. Then, as the problem has a discrete domain, we can define robustness in the sense introduced by B. Roy in def. 17 to 19. Adapting the formulation making $\mathbf{x}$ the protection and the set of scenarios s_i equal to the set of possible attacks, in terms of minimization of impact the performance constraint $I(\cdot)$ is defined in terms of $b = \min_{\mathbf{x}} \max_s(F(\mathbf{x}, \mathbf{p}, s))$ and $w \geq b$. Then the absolute robustness is:

$$r_{bw}(\mathbf{x}) = \begin{cases} 0 & \text{if} \quad \exists s_i \ : \ F(\mathbf{x}, \mathbf{p}, s_i) > w \\ |\{s_i\}| \ : \ F(\mathbf{x}, \mathbf{p}, s_i) \leq b & \text{otherwise} \end{cases} \tag{6.6}$$

In words, r_{bw} is the number of nodes that can be attacked such that its associated impact is lower than the minimal maximum impact possible and no node produces an impact greater than w. Now if we simply say $w = b$, the most robust protection is that that generates the min-max impact, which is, once again, equivalent to solve eq. 6.3 using the worst case criterion.

As we can see, even when the DM/Analyst tackles the problem from different viewpoints, the min-max criterion seems the best option to confront the problem of optimal protection allocation in absence of more evidence according to the application of the AUREO.

2. Check for input uncertainties

This point was already analyzed in the previous section. In general, the vector of decision variables is composed by the protections assigned to the nodes and no uncertainty is considered to be attached to these variables. By contrast, the propagation times and the entities to affect are aleatory quantities assumed to be evenly distributed according to the parameters reported in sec. 6.3.1.

Finally, the strategy of attack remains uncertain, at least partially, thereby the features that matter, namely the place where the attack begins and/or the probabilities of nodes being attacked, are epistemic in nature.

3. Search for uncertain objective functions

The assessment of the ANPA and the TTRAD is done by means of the hybrid tool described in sec. 6.2.3. As the procedure is based on Monte Carlo simulation, the results are stochastic in nature and we get the expected value of the impact within a confident interval. This uncertainty should be considered during the ranking procedure.

The final robustness-seeking program

As we have seen, the model should aim at minimizing the maximal impact in order to achieve robustness. In terms of a multiple-objective formulation, the analyst can choose to solve eq. 6.3 or eq. 6.5. Due to the fact that the estimation of impact is done by two measures whose values vary as a function of the point of attack and the allocation of the protection, for both equations the assessment of the maximal impact produces a Pareto frontier given a decisional vector (a particular protection allocated) over the whole range of scenarios of attacks. Hence, if one wants to determine which allocation is better, one has to compare the Pareto frontiers associated to the different allocations and the node with minimal maximal impact should be the node whose protection is robust against the actual level of uncertainty.

For convenience let us split vector $\mathbf{p}$ into two vector, $\mathbf{p_a}$ and $\mathbf{p_G}$, the former is the point of attack and the latter the collection of propagations times and entities, with the correspondent δ_{p_G}. Now in mathematical terms the problem is to find the node n_i to protect ($\mathbf{x} = (i)$, $i = \{1, 2, \ldots, |N|\}$) such that

for eq. 6.3:

$$\mathbf{x} = \arg\{\min_{\mathbf{x}} \max_{\mathbf{p_a}} \text{ANPA}(\mathbf{x}, \mathbf{p_a}, \mathbf{p_G}) \wedge \max_{\mathbf{x}} \min_{\mathbf{p_a}} \text{TTRAD}(\mathbf{x}, \mathbf{p_a}, \mathbf{p_G})\}$$

for eq. 6.5:

$$\mathbf{x} = \arg\min_{\mathbf{x}} \max_{\mathbf{p_a}} \left(I_{\text{ANPA}}(\mathbf{x}, \mathbf{p_a}, \mathbf{p_G}), I_{\text{TTRAD}}(\mathbf{x}, \mathbf{p_a}, \mathbf{p_G})\right)^t$$

6.3.3 AUREO: Stage 2 and results

Once the mathematical program is defined, the second stage aims at determining the best possible implementations in terms of EA and particularly MOEA.

The preceding program comprises two tasks, the first one is to find the Pareto frontier for a particular protection over all the possible attacks, while the second one consists in determining the minimal frontier amongst the whole set of frontiers. Let us start by analyzing the first task.

If the antagonists would like to find the set of best target nodes in terms of impact, they are compelled to find the efficient or Pareto frontier by maximizing the ANPA and minimizing the TTRAD. For instant, consider the Pareto frontier for Network 1 shown in fig. 6.5; each point depicted represents a point of attack, i.e. the node where the attack begins. Hence, if the antagonists set on node 22 or node 24 they can get the optimal results in terms of damage. Likewise, if the defender wants to assess the quality of the protection implemented, by finding the Pareto frontier it is possible to estimate the reduction in the impact measures in terms of the displacement of the Pareto frontier towards the defender's ideal point, which is maximal TTRAD (infinite time) and minimal number of entities affected (no entity affected).

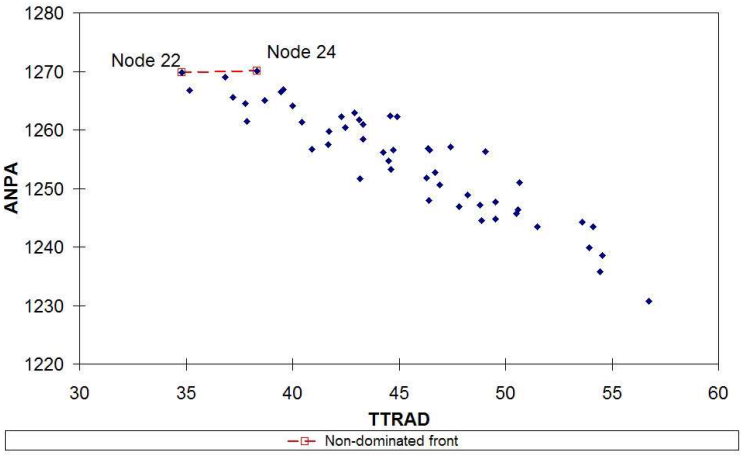


Figure 6.5: Attacker's efficient frontier for Network 1 (fig. 6.3) without protections [161, 234].

Concerning the algorithm, the efficient frontier can be found either by complete enumeration or using a MOEA. In both cases uncertain objective vectors are to be compared due to the use of simulation. Notice that different evaluations of the same argument produce different expected values and confidence intervals if the samples differs and the real mean value could lie outside the confidence interval. Some considerations to overcome these problems are:

- When using the mean as representative value for comparisons, make sure that two or more replications of the same decision vector are not being compared, otherwise one of them will dominate the others. If possible, avoid performing multiple evaluations of the same decision vector and implement cloning control during the recombination and ranking procedures.
- When comparing different decision vectors, especially if their images overlap, consider to verify that the differences amongst expected values are significant from both statistical and decisional viewpoints. In other words, even when two objective vectors can be numerically different, such a difference can have no statistical significance and moreover it can have no meaning for decision purposes, and that should have influence on the ranking and the archiving procedures. The use of ϵ-dominance to build thresholds of discernment (as in sec. 5.3.4) can be an option.

The second step aims at determining the minimal Pareto frontier. This may constitutes a major complication since we do not have a unique and universally accepted way to compare Pareto approximation sets (see sec. 3.3.5). On the other hand, the current performance measures are not designed to cope with uncertain outcomes (see sec. 3.4.6)). Nevertheless, even when none of the current metrics can capture all the features considered relevant to assess the quality of approximation sets, under some circumstances, like at least some of the frontiers considered do not intersect the others, they can suffice to reduce the group of Pareto frontiers to just a few ones and if one is lucky, to the min-max frontier.

The algorithm can be constructed to handle the problem through a double-loop configuration or a single one. Double-loop algorithms can be used to handle a variety of situations in EA (see [58] and references therein) by means of a outer and inner loop layout; the inner loop solves a particular instance of the problem using the variables of the outer loop as a fixed parameters, whereas the outer loop makes possible the exploration of the whole domain. In terms of the problem considered here, the double-loop will consists of an inner loop that finds the Pareto frontier for maximal ANPA and minimal TTRAD given that a particular node is protected, and an outer loop that varies the node to protect, calls the inner loop to get the efficient frontier for that node and then keeps the Pareto frontier with minimal impact amongst the nodes studied. The implementation options are:

1. The outer loop makes an exhaustive enumeration of protection allocations whereas the inner loop makes an exhaustive enumeration of the points of attack.

2. The outer loop makes an exhaustive enumeration whereas the inner loop finds the Pareto frontier using MOEA.

3. The MOEA explores the whole domain and stores the min-max solutions.

For the third option a chromosome with the point of attack and the node of defense was used, in such a way that the whole search space was explored. Additionally, the ranking procedure was modified in order to direct the search towards min-max results. The proposed procedure is as follows:

1. If both vectors to compare belong to the same protected node (maximization of impact case), apply ordinary ranking according to the objectives maximize the ANPA and minimize the TTRAD, else

2. vectors belong to different protected node (minimization of maximal impact case). Apply ordinary ranking according to the objective minimize the ANPA and maximize the TTRAD.

The goal consisted in finding the min-max frontier whose components may or not belong to the same protected node. In the latter case, additional effort would be necessary to find the robust protection amongst the nodes identified.

During this research, the three options were studied but only the two first came to good results, indicating the necessity of more research effort towards the development of the last alternative. In any case, it is important to remark that to the best of our knowledge, min-max problems with multiple objectives have been scarcely studied in EA (e.g. [79]) and we are not aware of any attempt within the EA or EMO community to solve problems of the type described above.

Robust protection for Network 1

Fig. 6.6 shows some selected Pareto frontiers for the analysis of Network 1. Due to the small dimension of Network 1, the efficient frontiers are easily found analyzing the whole network (52 nodes) and the use of MOEA does not seem justified in small networks of that such. However, if we relax the constraint

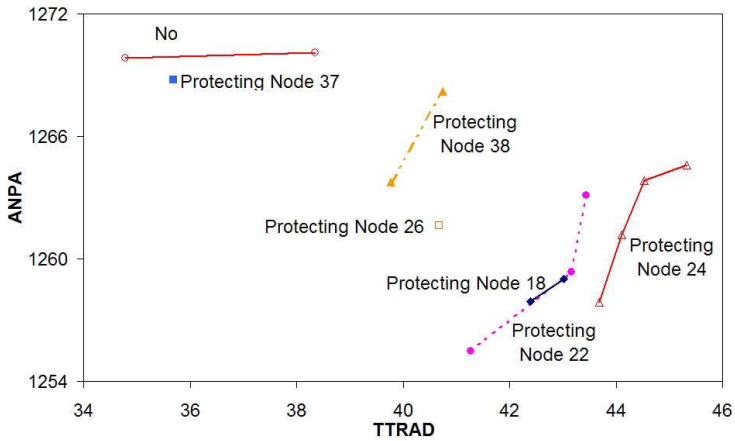


Figure 6.6: Selected antagonist's efficient frontiers in function of the protected node of Network 1 [161].

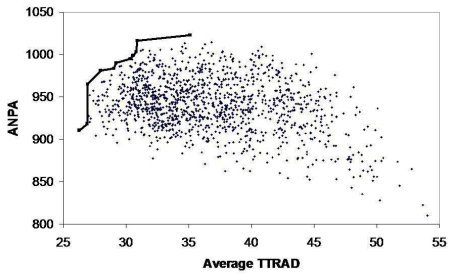


Figure 6.7: Antagonist's efficient frontier provided two simultaneous attacks on Network 1 [234].

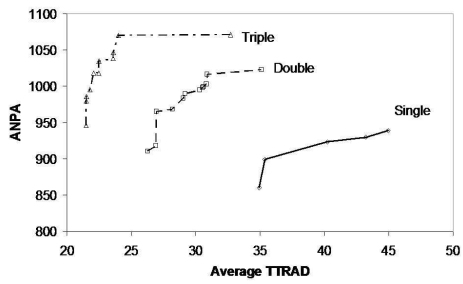


Figure 6.8: Antagonist's efficient frontiers up to three simultaneous attacks on Network 1 [234].

of one single attack and one single protection, as the size of the network increases (as is the case of Network 2) or the number of combinatorics considered augments, the combinatorial explosion makes the exhaustive enumeration not affordable in general. As an example of combinatorial explosion, fig. 6.7 brings the total number of cases and the Pareto frontier when two simultaneous attacks are performed on Networks 1. When simultaneous actions against three nodes are taken into account, the use of MOEA seems justified. Fig. 6.8 shows the efficient frontiers up to three simultaneous attacks, where the case of size three were analyzed using NSGA-II with 50 generations and sizes of population and archive equal to 20.

The results shown in fig. 6.6 can be analyzed either by visual inspection or using metrics. The first way is not applicable in many practical situations but the use of performance metrics cannot lead to definitive results in multitude of cases. Thus both approaches can be considered complementary. For example, the application of the ϵ-indicator leads conclude that the protection of nodes 22 and 24 minimize the maximal impact but does not allow to differentiate

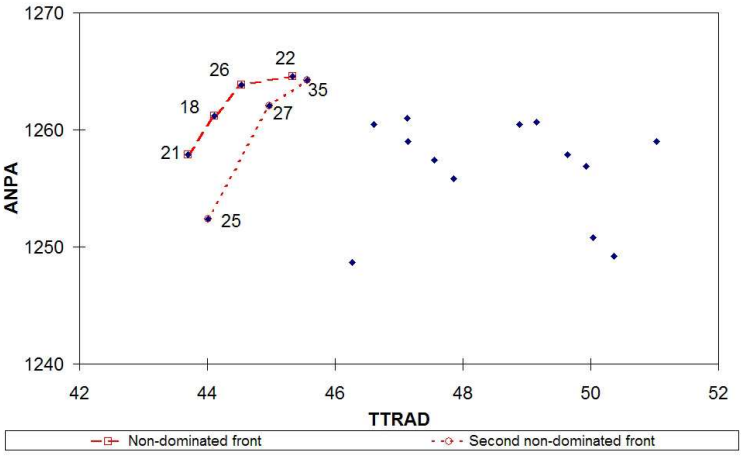


Figure 6.9: Antagonist's efficient frontiers when node 24 of Networks 1 is protected: labels represents nodes [161].

which one between these two nodes is better to protect. In consequence, the participation of the DM is crucial to identify the final alternative. For instance, if node 24 is identify as the node where the protection is more robust, an attack on nodes 21, 18, 26 or 22 produces the maximal impact (see fig. 6.9). The reader can verify that these nodes are topological neighbours in Network 1. This interesting result suggests a part of the network that could be under additional surveillance in order to reduce the risk of incurring in maximal damage. Thus, if the mentioned nodes were unable to be reached by an attack, the maximal impact is that produced by setting on node 25, 27 or 35.

Robust protection for Network 2

As the problem is conceptually the same, Network 2 and Network 1 share the same robustness-seeking program. However, the former has been only analyzed with exhaustive enumeration because the dominated nodes may offer worthwhile insights into the capabilities of the simulation tool employed to assess the impact.

For security reasons, the correspondence between nodes in fig. 6.4 and their labels are not given. Fig. 6.10 brings the antagonist's efficient frontiers that are to be minimized in order to find the more robust protection. Labels indicates the nodes that compose the front. Notice that the wholly of the frontiers are dominated by right point of the lower curve, which is generated when node 8 is protected. Hence, the use of the ϵ-indicator leads to identify the robust protection without problems.

The analysis of some dominated points indicates the possibility of generating islands or isolated subnetworks, as the result of protecting some special nodes called cut-nodes [189]. Indeed, by visual inspection, the reader can find some nodes whose removal disconnects the network generating islands. In consequence if such a node is protected, no matter what node is set on by the antagonist, the whole of the network would never be affected. Conversely, an

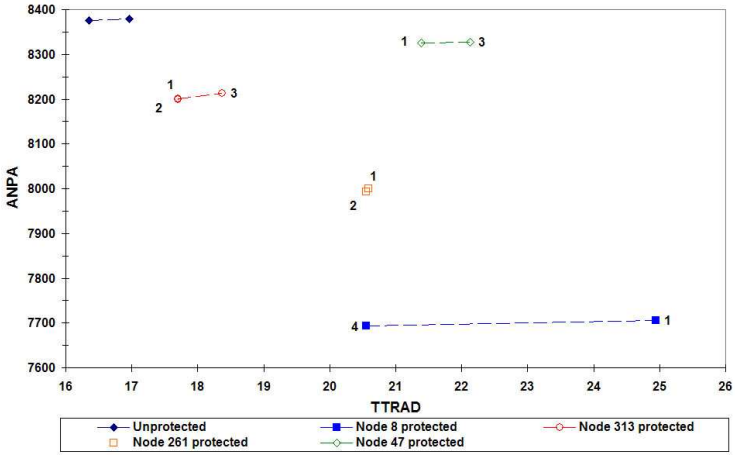


Figure 6.10: Selected antagonist's efficient frontiers if function the protected node of Network 2 [159].

attack on an unprotected cut-node will cause the immediate disconnection of the network, with the concomitant effects; something that could be appealing under certain circumstances. Sound interpretations from both the defensive and the antagonistic viewpoints of this issue and its incorporation into the impact assessment is a subject open to further research [159]. As the notion of impact evolves, the AUREO should be reapplied in order to guide the effort towards the robust solution.

6.3.4 What if the information changes?

Let us consider now the effect of having more information about the possible attacks and its influence in the defense strategy. Imagine that the intelligence department gathered information about the cost of performing the attack, understanding cost not necessarily in monetary terms but in terms of investment of resources, whatever resources means. Fig. 6.11 shows such costs (arbitrarily assigned) as a third dimension in the impact profile of attacking Network 1 nodes.

If the antagonist performs a multi-objective optimization and determines the efficient frontier, they will find that the same nodes considered as efficient in previous two dimensions analyses keep being non-dominates but in addition new nodes whose costs are efficient appear (all these solutions are highlighted with a ring). In general, the inclusion of more information in the form of objectives makes the identification of the robust protection more difficult since the number of non-dominated points increases concomitantly, although the final allocation could be the same if the new objectives do not modify the interpretation of 'impact'.

However, if the new information can influence the model, as were the case of having more precise probabilities, the robustness-seeking program could change and leading to different results. For example, assuming that the costs can be translated some how into likelihoods of attack, eq. 6.3 is no longer applicable

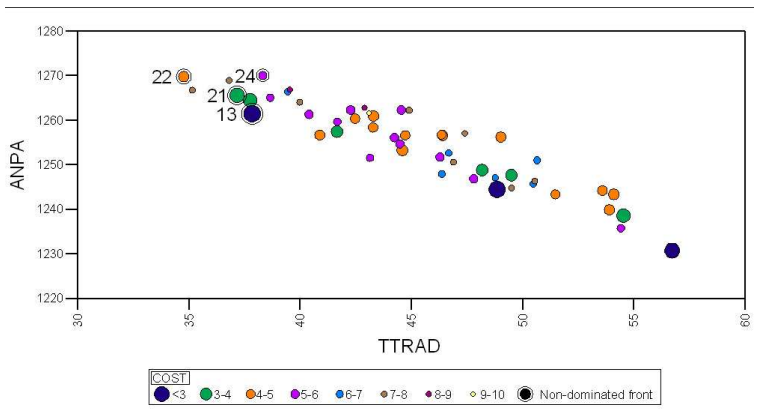


Figure 6.11: Cost and impact of attacking nodes of Network 1 from the antagonist perspective (diameters are inversely proportional to costs).

but eq. 6.5 should be used. For that matter, we directly transformed costs into probabilities by inverting and normalizing the values (thus diameters in 6.11 reflect probabilities). Besides, the impact measures were also transformed into risk indexes as suggested earlier on in this section. Finally the program was reformulated to yield:

$$\mathbf{x} = \arg\min_{\mathbf{x}} \max_{\mathbf{n_i}} p(n_i) \left(I_{\text{ANPA}}(\mathbf{x}, \mathbf{p_a} = n_i, \mathbf{p_G}), I_{\text{TTRAD}}(\mathbf{x}, \mathbf{p_a} = n_i, \mathbf{p_G}) \right)^t \quad (6.7)$$

The preceding program aims at minimizing the maximal risk. A version for the minimization of the expected risk is also possible, viz.:

$$\mathbf{x} = \arg\min_{\mathbf{x}} \left(\sum_{\mathbf{n_i}} p(n_i) I_{\text{ANPA}}(\mathbf{x}, \mathbf{p_a} = n_i, \mathbf{p_G}), \sum_{\mathbf{n_i}} p(n_i) I_{\text{TTRAD}}(\mathbf{x}, \mathbf{p_a} = n_i, \mathbf{p_G}) \right)^t$$

Alternative formulations can be proposed but its analysis lays on the risk theory realm.

Program of eq. 6.7 is solved with the same min-max procedure described before; the results are presented in fig. 6.12. Notice that the protection allocation changes from node 24 to node 22 and nodes 13 and 21 are now considered the most risky nodes in most of the cases. This result exemplifies how the amount of available information could influence the whole process of decision making for that the very model and the results are now different. However, in both cases the AUREO helped the process providing the analyst with a thorough methodology to cope with uncertainty in evolutionary optimization-based MCDM.

6.4 Summary of the chapter and concluding remarks

In this chapter we have reported the application of the AUREO to an optimal protection allocation problem with several possible models, epistemic uncertainty and aleatory in the input and aleatory uncertainty in the functional

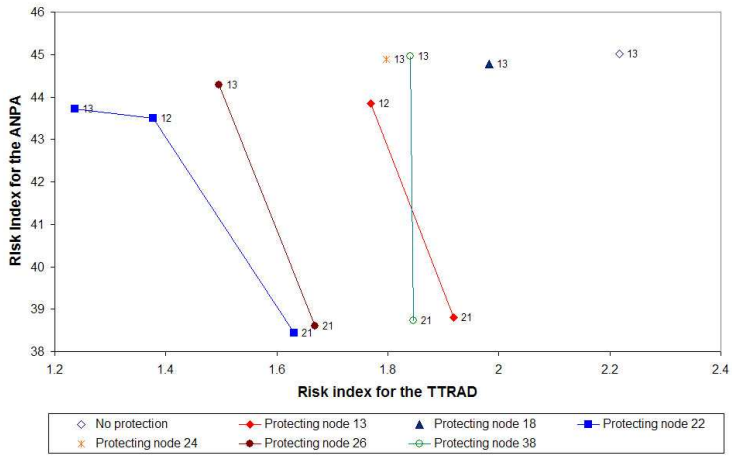


Figure 6.12: Selected min-max frontiers for risk reduction of Network 1 against single attacks (labels indicate the nodes of the efficient frontier).

evaluation. The problem is one of a salient research field, so its resolution is interesting for both the EMO and the risk analysis communities.

The AUREO made possible to define the robust-seeking problem for the allocation of the protections. Concurrently, the study of the solving implementations discovered some of the many complications of min-max problems with multiple-objectives, a problem that has not been studied previously to the best of our knowledge. Indeed, the problem of finding the protection with min-max impact resorts to solve many instances of maximization with a final search for the minimal Pareto frontier amongst the whole set of fronts obtained. Hence, the resolution requires the ability to find efficient frontiers under uncertainty and to compare approximation fronts, being both open issues in EMO.

The results obtained here show that it is possible to solve the problem although further research is needed in order to enhance the capabilities of the MOEA to avoid double-loop configurations. Likewise, from the network assessment viewpoint, it is necessary to refine the notion of impact and risk in order to reflect more situations, like the appearance of islands, into the indexes of damage.

Chapter 7

Conclusions and further research

Amongst the many facets of scientific and technological problems, decision-making is often present as one of the primary task for their resolutions. Indeed, many problems in engineering and related sciences are ultimately of decision-making. The complications often bore by real life problems in this realm along with constraints in different kinds of resources compel practitioners to find efficient resolution tools. As a result, metaheuristics have risen as one of the most appealing alternatives to solve problems where classical methods fail or their application seems unpractical. Moreover, the tremendous development of metaheuristics for solving multiple-criteria decision-making problems witnessed during the last decades, make them preferential tools in engineering practice nowadays.

However, real problems are subject to different kinds and sources of uncertainty whose presence can mislead the identification of the optimal alternatives, unless such uncertainty is properly addressed. In that sense, this thesis is concerned about the different aspects of uncertainty handling in decision-making processes based on evolutionary multiple-objective optimization.

On the contents and contributions

In the first part of this thesis we have offered some insights into the origin, sources and types of uncertainty along with a survey of the main theories for characterizing uncertainty. Afterwards, the main aspects of MOEA as well as the state of the art in EMO under uncertainty were presented. With this in mind, we have offered a thorough analysis of the effect of uncertainty in decision-making from the viewpoint of the algorithmic operation. In particular, we have addressed the alternatives for designing ranking and archiving procedures when the outcomes are characterized by non-probabilistic, probabilistic, fuzzy and imprecise probabilistic uncertainty. To the best of our knowledge this constitutes an original contribution to the field since previous efforts are scarce and have only focused on probabilistic outcomes.

A multiple-objective algorithm for non-probabilistic uncertain outcomes, named IP-MOEA has been proposed. This algorithm demonstrates that the construction of algorithms with this type of uncertainty is possible. Addition-

ally, the treatment of non-probabilistic uncertainty by means of the ϵ-dominance was also explored in connection with the concept of minimal regret. The analysis showed that, under some circumstances, the extension of minimal regret to continuous sets comparisons equals the application of the ϵ-indicator so that the use of the latter to handle non-probabilistic outcomes seems worthwhile.

In parallel, a procedure for designing a probabilistic archiving methods based on hypergrids was suggested. This proposal makes a probabilistic interpretation of the concept of hyperbox's density that could be useful to build archiving procedures that handle probabilistic uncertainty.

The concept of robustness was also explored theoretically in this thesis as a practical alternative to address uncertainty in decision-making. The state of the art in robust optimization and robust design was surveyed and the concept of robustness was investigated in terms of its possible formulations regarding the information available for the DM. As a result, a taxonomy of classes of robustness in terms of the available information was introduced. Additionally, we found that the problem of interpreting robustness with respect to possibilistic, fuzzy or imprecise probabilistic uncertainty seems to be not possed before. Some hints for studying the possible links between robustness and non-classical uncertainty frameworks were also offered.

The first part of this thesis concluded with the introduction of an innovative methodological framework for the Analysis of Uncertainty and Robustness in Evolutionary Optimization or AUREO. The AUREO constitutes the main contribution of this thesis regarding the use of EA in MCDM. The AUREO provides DM and analysts with a methodological framework that encompasses the theoretical and practical aspects of uncertainty and MOEA operation to be taken into account in order to: a) formulate the decision-making problem in terms of a uncertainty-handling program (stage 1) and b) to solve it with the help of MOEA (stage 2). The AUREO can be applied to select and to adapt existent MOEA as well to design new ones.

The second part of this thesis presented the application and validation of the AUREO to three dependability multiple-objective problems. In the first problem, the percentage representation introduced during the stage 2 of the AUREO allowed to enhance the efficiency of the algorithm noticeably. This representation scheme can be very promising in problems with double-loop configurations with dependencies between the variables controlled by each loop.

A real multiple-objective scheduling problem was also analyzed with the help of the AUREO, in order to identify robust solutions thereby reducing the set of optimal alternatives found in previous studies. As a result, some changes in the formulation of the problem as well as the EA used to solve it, yielded a set of robust optimal solutions with a suitable cardinality for decision-making.

Finally, the AUREO was also employed to extend a single-objective analysis of complex systems vulnerability into a multiple-objective analysis with uncertainty handling. Due to the novelty of this work, many aspects still remain unexplored, however, the AUREO helped to define the alternative robustness-seeking programs to study as well as to delineate the solving procedure using MOEA. Some challenging issues studied in this problem could not be settled during this research, viz. a) the development of efficient procedures to address min-max optimization with many objectives, which is related to b) the problem of ranking groups of approximation fronts and c) the problem of assess the quality of a uncertain approximation front. It must be remarked that these issues

have been not previously solved or even studied so far in EMO.

On the future work

Further research is well delineated by the issues not solved and the suggestions offered in this thesis.

On the one hand, new perspectives for the design of MOEA is provided by the probabilistic interpretation of the hyperbox's density suggested in this thesis. Indeed, density control under uncertainty is one of the least explored topics in the state of the art in MOEA under uncertainty. Thus, further contributions in the archiving procedures aimed at improving the convergence properties as well as producing more uniform frontiers from uncertainty outcomes would constitute a significant advance.

Likewise, the usefulness of the ϵ-indicator to handle non-probabilistic uncertainty should be assessed. As was pointed out in this thesis, non-probabilistic uncertainty can be tackled by several approaches, being the minimal regret criterion one possibility amongst them. The resemblance between the minimal regret and the ϵ-dominance remarked in this thesis suggests the plausibility of building ϵ-indicator compliant MOEA to handle non-probabilistic uncertain outcomes. This way a DM/analyst interested in the minimal regret criterion could take advantage of the existent MOEA to solve problems with such criterion.

The ϵ-dominance is also in connection with the lack of methods to assess the quality of uncertain approximation fronts, no matter the way the uncertainty is characterized. It would be interesting to explore metrics based on ϵ-dominance to assess the quality on such approximation frontiers.

Regarding the challenging issues pointed out in this thesis, the study of the efficiency of the min-max multiple-objective optimization seems interesting due to the existence of problems where the quality of an alternative is not given by a single solution but a frontier. Indeed, the difficulty in comparing frontiers when the intersect is a non-solved problem in quality assessment of MOEA. Besides, min-max optimization can be performed with double-loop algorithms, but at expense of dealing with the combinatorial explosion. Hence, contributions in min-max multiple-objective optimization would benefit not only analyst confronted to problems of the type studied in this thesis but a wide range of matters like quality assessment of MOEA and general min-max metaheuristics as well.

On the other hand, one important remark is that, due to the difference between types and amongst sources and frameworks for characterizing uncertainty, the design of all-purpose MOEA that handle uncertainty is not a proper goal. Instead, future efforts might be aimed at addressing classes of problems according to the taxonomy of uncertainty-handling and robustness-seeking programs introduced in this work. In that sense, new MOEA could be formulated by exploring the use of different theoretical frameworks to represent uncertainty, in the subclasses of robustness-seeking programs introduced in this thesis.

Finally, the questions risen about the possible interpretation of robustness when coping to possibilistic, fuzzy or imprecise probabilistic uncertainty require new theoretical efforts that could be worthwhile to the decision-making, the EMO and the approximate reasoning communities.

Appendix A

Robustness concept: A survey

A.1 Forewords

In this appendix we bring the whole information sent to and received from some prominent researchers that work in the MCDM, the EA and the uncertainty modelling fields.

A.2 Questionnaire

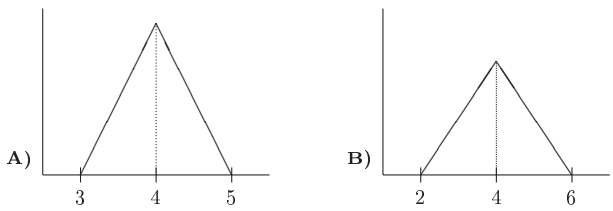
The body of the email sent to the researchers on April 25th 2007 is:

> *Dear Professor,*
>
> *As a part of my PhD thesis and for personal interests as well, I'm writing you in order to ask your kind contribution. Robustness is the core of my current research work which extends to multiple criteria decision making (MCDM) and evolutionary computation (EC). Special interest is put on engineering and risk assessment problems.*
>
> *The intention of the present experiment is to get a deeper insight into your meaning of robustness, what is it exactly and what is its worth as criterion within an EC+MCDM framework. To do so, I propose you a short list of cases to be classified. For that matter, I would like you to express your opinion about the robustness of the different set sketched below using the well know binary relations* ***strict preference*** *(P),* ***weak preference*** *(Q),* ***indifference*** *(I) and* ***incomparability*** *(J). Therefore, your judgements e.g. will take form of aPb (a strictly preferable to b). Then, please detail the reason of your selection.*
>
> *The results are to be included as part of my PhD dissertation. If you are willing to collaborate, please let me know if I can subscribe your opinions with your name or should I use* ***anonymous****.*
>
> *Any additional comment is highly appreciated. You are free to submit your answers in English, French or Spanish; please choose the language that makes you feel more comfortable.*
>
> *Thank you in advance,*
>
> *Daniel E. Salazar A.*

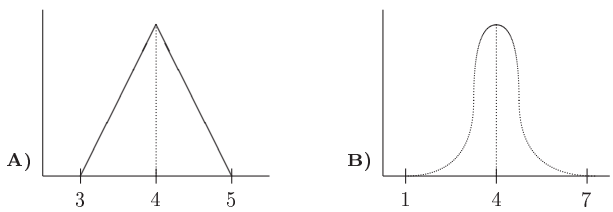
A *.doc file containing the questionnaire was attached to the above email. The content was as follows:

> *Imagine that each one of the following cases represents the output of a system, i.e. the same system under different operation conditions or two different systems under the same operation conditions. The goal is to classify each couple of outputs in terms of robustness. Notice that the figures are not necessarily scaled. If you need to make some assumptions in order to select the more robust, please indicate them. These are probability density functions; please indicate which one is more robust.*

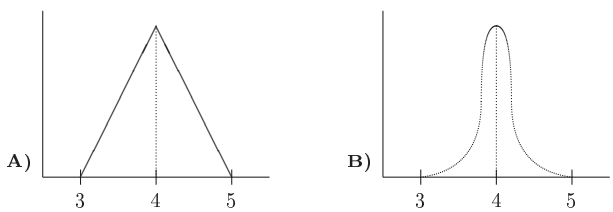
Problem 0:



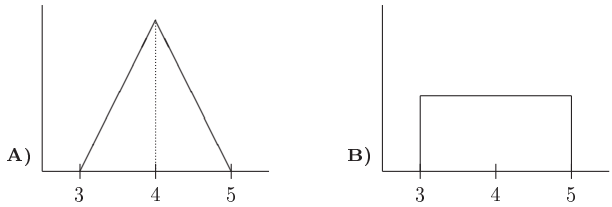
Problem 1:



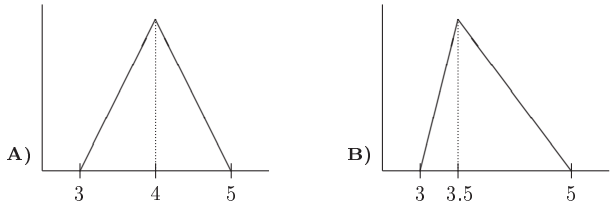
Problem 2:



Problem 3:



Problem 4:



These are classical sets:

Problem 5:

A = {1, 2, 3} B = {1, 3}

Problem 6:

A = {9, 10, 11} B = {0, 10, 20}

Problem 7:

A = {1, 2, 3} B = {0, 6}

These are sets of possible elements x such that each element has a probability $p(x)$ of occurrence:

Problem 8:

A = $\{p(1) = \frac{1}{3}, p(2) = \frac{1}{3}, p(3) = \frac{1}{3}\}$ B = $\{p(1) = 0.4, p(2) = 0.2, p(3) = 0.4\}$

Problem 9:

A = $\{p(1) = 0.5, p(2) = 0.5\}$ B = $\{p(1) = 0.2, p(2) = 0.8\}$

Problem 10:

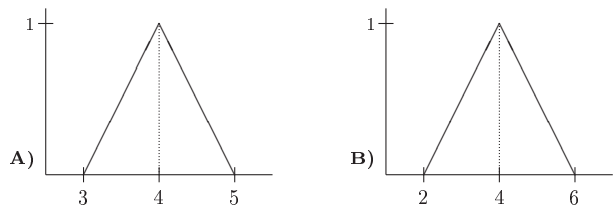
A = $\{p(1) = 0.5, p(2) = 0.5\}$ B = $\{p(1) = 0.5, p(25) = 0.5\}$

Problem 11:

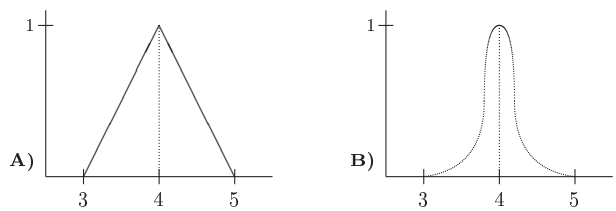
A = $\{p(1) = 0.5,\ p(2) = 0.5\}$ B = $\{p(1) = 0.2,\ p(25) = 0.8\}$

These are fuzzy sets:

Problem 12:



Problem 13:



A.3 General comments about the questionnaire

From the very beginning, this questionnaire was not meant to draw solid conclusions about robustness, but to get some insight into the features towards which analysts and DM centre their attention when consulted about robustness and the possible agreements or disagreements about in selecting such features.

The first group of questions (problems 0 to 4) aims at identifying what features amongst variance, range and mean are indicators of robustness. The second group (problems 5 to 11) poses the question of robustness of discrete sets where cardinality, mean, variance and range could serve as referential features. Finally, problems 12 and 13 formulate the problem when the uncertainty takes the form of fuzzy sets. Here the range of variation (support set) and the degree of possibility around the same centrepoint are the distinguishable features.

A.4 Answers

Amongst the fifteen researchers invited to participate, only three of them submitted their answers. As none of them claimed to appear as anonymous, their identities and their answers are given next.

A.4.1 B. M. Ayyub's answers

On April 28th, 2007, Prof. Bilal M. Ayyub submitted the following answer:

> *"Dear Daniel, My responses are attached.*
> *you may use my name or treat them as coming from an anonymous source – as you wish.*
> *Best wishes,*
> *Bilal"*

The email was attached with a *.doc file with the answers, viz.:

Case	Preferences
0:	AQB (Smaller coefficient of variation)
1:	AQB (Assuming smaller coefficient of variation and a smaller range)
2:	BQA (Assuming smaller coefficient of variation)
3:	AQB (Assuming smaller coefficient of variation)
4:	AIB (Skewness might have an importance depending on the application domain)
5:	BQA (Cardinality is smaller)
6:	AQB (Same cardinality, but smaller range)
7:	BQA (Cardinality is smaller)
8:	BQA (Greater entropy)
9:	BQA (Greater entropy)
10:	AQB (Greater range)
11:	AIB (Depending on the application domain)
12:	AQB (Smaller range)
13:	BQA (Tighter shape around central value)

Comments

When Prof. Ayyub mentions in his answer cardinality and entropy, he alludes to the Hartley ($H(\mathcal{A})$) and Shannon ($S(\mathcal{A})$) measures which are defined as [7]:

$$H(\mathcal{A} = \{x_i | i = 1, \ldots, n\}) = \log_2(|\mathcal{A}|) = \frac{\ln(n)}{\ln(2)} \qquad \text{Hartley measure}$$
$$S(\mathcal{A} = \{(x_i, p(x_i)) | i = 1, \ldots, n\}) = -\sum_i^n p(x_i) \log_2(p(x_i)) \qquad \text{Shannon measure}$$

According to these measures, for discrete sets for which no probability information is given, the larger the more nonspecific. Likewise, when a distribution of probability is associated, the sets are more entropic and therefore show a higher uncertainty when the mass associated with each element is more spread. These measures belong to a broader body of measures of uncertainty described in the Generalized Information Theory [104, 7].

The introduction of these kind of measures to classify robustness well worths a comment, since they produce some results that could seem counterintuitive. Notice e.g. that sets A and B in problem 7 the range of outcomes of set A is smaller than for set B, so are their variances. Thus, if robustness is associated with smaller variability, set A must be more robust than B; however the Hartley measure indicates the contrary.

Likewise, in problem 8 set B has lower Shannon entropy , thus it is preferred, although the range of variation and the mean is the same for both sets and the A has greater probability to hit the mean than B.

A.4.2 J. Branke's answers

On May 15th, 2007, Prof. Jürgen Branke submitted his answer:

> *"Dear Mr. Salazar,*
>
> *Robustness is an interesting topic. The term is used in many different ways, and it certainly makes sense to define it more closely. Some possible definitions are stated in my book (see attachment). However, I believe what definition is most suitable depends on the particular application, and I doubt it is possible to decide what robustness is by averaging over many different opinions.*
>
> [...]
>
> *The Santa Fe Institute has a working group on robustness which may be interesting to you.*
>
> *Having said this, here is my evaluation (indicating always the more robust one):*
> *0) A*
> *1) A*
> *2) B*
> *3) A*
> *4) A*
> *5) A*
> *6) A*
> *7) unclear*
> *8) A*
> *9) unclear*
> *10) B*
> *11) B*
> *12)/13) Not sure what the difference is between distributions and fuzzy sets.*
>
> *Overall, I guess with equal average I usually voted for smaller variance and better worst case (my assumption was maximization of the average)*
>
> *Hope that helps,*
>
> *Jürgen Branke"*

Comments

The researcher was consulted only once, so it is impossible to know what motivated the discrepancies between his preferences and Prof. Ayyub's. One can presume that the disagreement in problems 5 and 8 is somehow due to the reasons argued in the previous discussion about the conclusions drawn by the uncertainty measures. For example, in problem 5 set A has smaller variance then set B but larger cardinality. However, answers to problems 10 and 11 are seems incoherent with the rest of the answers. Notice that the range of B is significantly larger than A's and the probabilities are equal for both sets in problem 10. The researcher was not consulted again about these answers, but we feel that these answers do not express well the researcher's preferences.

A.4.3 B. Roy's answers

On April 27th, 2007, Prof. Bernard Roy answered:

> *"Bonjour,*
> *Ma charge de travail actuelle ne me laisse pas le temps de prendre connaissance en détail de votre pièce jointe. Peut-être trouverez-vous déjà des réponses dans le document joint ainsi que dans la bibliographie à laquelle il renvoie. Ce document paraîtra dans les prochaines semaines dans un numéro spécial des annales du LAMSADE consacré à la robustesse. Sans doute d'autres articles de ce numéro pourront vous aider.*
> *Sincères salutations,*
> *Bernard Roy"*

In his answer, Prof. Roy indicated that he was unable to review the questionnaire in detail due to his labour compromises. However, a draft version of an article titled '*La robustesse en recherche opérationnelle et aide à la décision: Une préoccupation multi facettes*' [170] was attached, suggesting the concepts introduced in there as possible solutions to the questions proposed (def. 17 to 19).

Comments

The definitions suggested by Prof. Roy are not applicable to the cases presented in our questionnaire since he conceptualize robustness in a very different fashion that is not compatible with our questions without additional information.

Appendix B

Acerca de la incertidumbre y la robustez en toma de decisiones multicriterio basada en algoritmos evolutivos

Este apéndice constituye un resumen extendido del contenido expuesto en inglés en los capítulos anteriores. Para mantener la homogeneidad en el documento, se conservan los acrónimos y algunas numeraciones introducidas anteriormente. En la sección siguiente se expone la problemática abordada y los objetivos de la investigación. Posteriormente se presentan los elementos básicos del marco teórico necesarios para la correcta comprensión del trabajo realizado. Seguidamente se resumen los análisis realizados, resaltando las contribuciones originales en materia metodológica y algorítmica, juntos con algunas propuestas de investigación surgidas de dichos análisis. Finalmente se exponen las conclusiones y las reflexiones finales relativas a las futuras líneas de investigación.

B.1 Motivación y objetivos de la tesis

Los problemas de toma de decisiones están caracterizados por la presencia de varias alternativas sobre las cuales debe realizarse una selección que, en la mayoría de los casos, está relacionada con la identificación de la mejor alternativa. El concepto de 'mejor', es decir la calidad de una alternativa, juega por tanto un papel fundamental en la resolución de los problemas de toma de decisiones.

Cuando la calidad de una alternativa está referida a un único aspecto de la realidad, se habla de un problema que tiende espontáneamente a un valor extremo de calidad [167]: el problema es *monocriterio*. En cambio, cuando distintos aspectos de la realidad son relevantes para evaluar la calidad de una alternativa, el problema se convierte en *multicriterio* [167]. En este tipo de problemas, es necesario balancear los diferentes criterios implicados, ya que los

mismos suelen estar en conflicto entre ellos. Como resultado, no se obtiene una única alternativa óptima o mejor solución, sino un conjunto de alternativas óptimas o *eficientes* que representan consensos entre los máximos niveles alcanzables en los diferentes criterios de calidad.

En la práctica, la resolución de problemas multiobjetivo requiere concluir con éxito estas tres tareas: modelar matemáticamente el problema para darle una forma adecuada en términos de la calidad del modelo y de cara a su resolución, construir un algoritmo o método para identificar el conjunto de alternativas eficientes y, finalmente, seleccionar una alternativa específica que será implementada como solución del problema. Esta tesis se enfoca en los procesos de resolución fundamentados en Algoritmos Evolutivos (EA) o metaheurísticas.

Actualmente existe una tendencia muy importante de utilizar EA para construir los métodos de identificación de alternativas eficientes, debido principalmente a versatilidad frente a problemas con dominios y/o funciones no lineales, discontinuas y/o discretas, además de su eficiencia en la resolución de problemas de alta complejidad computacional. La disciplina que estudia el diseño y aplicación de tales algoritmos de conoce como Optimización Evolutiva Multiobjetivo (EMO), y tiene gran impacto en ingeniería, ciencias básicas y ciencias económicas, donde dichos algoritmos son utilizados en todo tipo de actividades de I+D, lo que explica la relevancia de considerar dichos métodos en esta tesis.

El objetivo fundamental de este trabajo es el estudio del efecto de la incertidumbre sobre el proceso de resolución de problemas multiobjetivo. Es bien sabido que la incertidumbre puede afectar considerablemente la calidad de las soluciones halladas, pero se ha estudiado muy poco la forma en que dicha incertidumbre puede controlarse cuando se emplean Algoritmos Evolutivos de Optimización Multiobjetivo (MOEA)

Particularmente se ha propuesto:

1. Analizar las opciones para clasificar la calidad de las alternativas de solución de problemas multiobjetivo cuando las medidas de desempeño están sujetas a incertidumbre.

2. Estudiar las alternativas de adaptación y diseño de MOEA para manejar incertidumbre, incorporando las opciones de clasificación de calidad.

Como fruto de las reflexiones y resultados surgidos de las actividades mencionadas durante este trabajo, y con el objeto de crear un marco metodológico que permita la comprensión de los distintos elementos a considerar en un problema con incertidumbre, se ha propuesto finalmente la sistematización de conocimientos, conceptos y experiencias, a través del *Análisis de incertidumbre y robustez en Optimización Evolutiva* o AUREO. Esta metodología, está orientada a ayudar tanto el trabajo de los diseñadores de MOEA, como el de los analistas que pretenden resolver problemas de toma de decisiones multiobjetivo apoyados en metaheurísticas.

La descripción de AUREO se desarrolla a lo largo de este resumen, acompañándose de aplicaciones a problemas reales de ingeniería. A tal fin, primero se hará una revisión de los conceptos fundamentales necesarios para entender y aplicar AUREO, y seguidamente se presentará la metodología en cuestión y se ejemplificará su uso.

B.2 Conceptos fundamentales de incertidumbre

En esta sección se resumen las ideas y conceptos presentados en el capítulo 2.

B.2.1 Discusión inicial

La información es un elemento fundamental en la toma de decisiones. La misma es necesaria tanto para describir problemas y formular modelos asociados como para evaluar y predecir el desempeño de los sistemas y su interacción con el ambiente que le rodea, es decir para analizar la calidad de las alternativas de solución.

El término 'incertidumbre' está asociado a una imperfección o defecto en la información. No obstante, dado que la información puede estar afectada por defectos de cantidad (ausencia y carencia), de calidad (irrelevancia, ambigüedad, inconsistencia, etc.) o ambos, los autores manejan con frecuencia otras categorías de imperfecciones, entre las que figuran: ignorancia, incertidumbre objetiva, incertidumbre subjetiva, ambigüedad, inconsistencia, duda, error, imprecisión, etc. (ver [6, 7, 72, 95, 166])

En este trabajo, nuestra lectura del conocimiento es el de una combinación de experiencia e información, siendo la ignorancia un defecto en el conocimiento y la incertidumbre un defecto objetivo (real) en la información, entendida esta última como datos o evidencias. Desde un punto de vista práctico, la información es el producto de investigar '*¿qué se requiere?*' (ver tabla B.1) en término de datos o evidencias, para diseñar, analizar, conocer, construir u operar un sistema, sea cual fuere, mientras que la incertidumbre surge como resultado de las imperfecciones en las distintas operaciones realizadas sobre dicha información (ver tabla B.2). De este modo, los errores de medición producen una incertidumbre métrica, las limitaciones para poseer información plena sobre el futuro o el pasado genera una incertidumbre temporal, la dificultad práctica de considerar todos los datos y las interacciones entre variables cuando modelamos un sistema da paso a una incertidumbre estructural, mientras que finalmente la multitud y variedad de categorías conceptuales, aunado a la imperfección en los medios de transmisión y decodificación, da como resultado una incertidumbre interpretativa o comunicativa.

Es por tanto de vital importancia, en primera instancia y de cara a la toma de decisiones, realizar de la mejor manera posible las siguientes tareas:

- Recolectar la mayor cantidad posible de información (datos),
- identificar las clases de imperfecciones asociadas con dicha información,
- encontrar modos adecuados de representar la incertidumbre, y finalmente
- minimizar el efecto de dicha incertidumbre sobre el proceso de decisión, y si es el caso, sobre la alternativa implementada.

No obstante, la realidad impone ciertas regulaciones sobre la recolección de información, como se lee en la tabla B.1. Por ejemplo, no toda la información que se puede recolectar es útil, de allí que la pregunta '*¿para qué se requiere?*' imponga límites prácticos a la cantidad de datos que se maneje. En cambio, otras limitaciones son impuestas por la *naturaleza* misma de la incertidumbre, o dicho de otro modo, la posibilidad o no de *obtener* la información.

Palabras clave	**Preguntas**	**Objetivo**
¿Qué? Información	¿*Qué* se requiere?	Obtener *información*, o de forma equivalente, reducir la incertidumbre
¿Para qué? Propósito	¿*Para qué* se requiere?	*Propósito* o utilidad de la información
¿Obtenible? Naturaleza	¿Se puede *obtener* la información?	*Naturaleza* de la incertidumbre
¿Reducible? Precio	¿Es *reducible* la incertidumbre?	*Precio* de la información
¿De dónde? Fuentes	¿*De dónde viene* la incertidumbre?	*Fuentes* de incertidumbre
¿En dónde? Influencia	¿*En dónde afecta* al sistema la incertidumbre?	Relaciones sistémicas, *entrada–sistema–salida*

Table B.1: Palabras claves para analizar la incertidumbre.

Tipo de operación	**Actividad o método**	**Tipo de incertidumbre**	**Información fundamental**
Obtención de información (Recopilación de datos)	Medición Investigación Observación	Métrica	Nuevas mediciones Mediciones almacenadas Descripciones
Obtención de información (Minería de datos)	Predicción Retrodicción	Temporal	Futuro Pasado
Asociación de la información (Relacionar datos)	Modelado	Estructural	Complejidad
Transmisión de información (Mediante el lenguage)	Comunicación y descripción	Interpretativo	Conceptos, metas y valores

Table B.2: Operaciones relativas a la información (tomado parcialmente de [166]).

La incertidumbre de tipo I, también llamada aleatoriedad, incertidumbre sistémica o variabilidad, está caracterizada por no poder ser reducida mediante la obtención de nuevas evidencias. Es por ejemplo el caso de una variable aleatoria cuya distribución ya es conocida, por lo cual nuevos datos sobre los valores que ha tomado la variable no aportan nada para conocer el próximo valor que ha de tomar. Por otro parte, la incertidumbre de tipo II, también llamada reducible o epistémica, se caracteriza por poder ser reducida mediante la obtención de nueva información acerca del sistema de interés o el ambiente que le rodea [145, 146, 47, 50, 49]. La diferencia en la naturaleza de ambos tipos tiene implicaciones importantes en la práctica, que se reflejan tanto en la dificultad de procesar adecuadamente la incertidumbre epistémica como en el tipo de inferencias o conclusiones que pueden hacerse frente a la presencia de uno u otro tipo [145, 47].

Nótese que, independientemente de la naturaleza de la incertidumbre, la obtención de información tiene un coste que, en algunos casos, puede superar al coste en que podemos incurrir como consecuencia de no poseer más información. Por ello es necesario considerar el *precio* de la información para definir hasta qué punto conviene o no seguir recolectando datos.

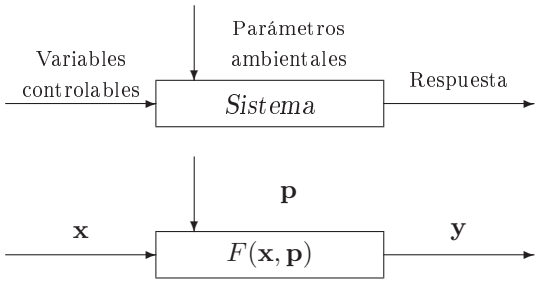


Figure B.1: Visión sistémica de las instancias y fuentes de incertidumbre: el sistema (arriba) y su representación funcional (abajo).

Finalmente, es de capital importancia analizar *en dónde* afecta la incertidumbre, para evaluar sus consecuencias. Considérese el sistema representado en la figura B.1, el cual es modelado mediante una función $F(\mathbf{x}, \mathbf{p})$ tal que $\mathbf{x}$ denota el el conjunto de variables y $\mathbf{p}$ el conjunto de parámetros ambientales. Los vectors $\mathbf{x}$ y $\mathbf{p}$ representan por tanto la totalidad de elementos, a ser considerados por quien modela al sistema, que alimentan o influyen sobre este último. La respuesta del sistema ante dichos elementos está representada por el vector de atributos $\mathbf{y}$.

La incertidumbre procede de muchas fuentes, siendo el poder descriptivo del modelo una de las más importantes. La complejidad de $F(\mathbf{x}, \mathbf{p})$ en términos del número de variables, parámetros y relaciones consideradas, es un punto fundamental en el análisis de incertidumbre. Otro factor de gran relevancia es la instancia afectada por la incertidumbre, la cual puede estar asociada a las variables ($\mathbf{x}$), al ambiente ($\mathbf{p}$), al propio sistema o al método que utilizamos para evaluarlo. Así, si el sistema está adecuadamente modelado, al menos todos las entradas importantes al sistema deberán estar contenidas en $\mathbf{x}$. Algunas variables pueden haber sido convenientemente evitadas por razones prácticas o ignoradas por errores de modelado. Por otra parte, siempre existen interacciones incontrolables dentro del sistema o con su medio ambiente, que han de ser consideradas durante la construcción del modelo.

La evaluación de la respuesta $\mathbf{y}$ está sujeta a dos instancias de incertidumbre, que son la calidad del modelo $F(\mathbf{x})$ y la idoneidad de los cálculos. Algunos errores en los algoritmos y/o códigos de computación o en los modelos de simulación son fuente de incertidumbre epistémica (ver [146]). Del mismo modo, las mediciones de $\mathbf{x}$ y $\mathbf{p}$ suelen estar sujetas a algún tipo de incertidumbre epistámica, debido a las limitaciones inherentes y a las incorrecciones en nuestros métodos y aparatos de medición. Adicionalmente, la aleatoriedad que procede del ambiente, falta de homogeneidad en los materiales, fluctuaciones en el tiempo y espacio [50, pg. 11], etc., debe ser considerada y correctamente modelada.

B.2.2 Marcos teóricos para la representación de la incertidumbre

A continuación, siguiendo [7, 105] se expondrán los conceptos matemáticos más relevantes sobre los que se fundamentan los marcos teóricos para representar

incertidumbre. Seguidamente se hará una exposición sucinta de dichas teorías.

Matemáticas de la incertidumbre: conjuntos y operaciones

Un *conjunto* es una colección de distintos *elementos* (también llamados *individuos* o *miembros*) definida a partir de un universo. Dicho universo,llamado conjunto universal y denotado por $\mathcal{S}$, contiene la totalidad de elementos pertinentes a un contexto particular.

Un elemento genérico x puede pertenecer ($x \in \mathcal{X}$) o no ($x \notin \mathcal{X}$) a un conjunto genérico $\mathcal{X}$. Si los elementos pertenecientes a $\mathcal{X}$ pueden ser etiquetados con enteros positivos, entonces el conjunto es *contable*, de lo contrario es *incontable*. Un conjunto sin elementos es un conjunto vacío ($\emptyset$).

El número total de elementos que componen a un conjunto es su *cardinalidad*. Los conjuntos contables pueden ser *finitos* o *contables infinitos*, mientras que los conjuntos incontables son siempre *infinitos*. La cardinalidad de un conjunto finito de n elementos es n ($|\mathcal{X}| = n$), en tanto que la de uno infinito es igual su tamaño ($|\{x|a \leq x \leq b\}| = b - a$). Un conjunto vacío tienen cardinalidad cero.

Las relaciones entre conjuntos se expresan mediante operadores. Dados los conjuntos $\mathcal{A}$ y $\mathcal{B}$, las siguientes relaciones son posibles: $\mathcal{A} = \mathcal{B}$ ($\mathcal{A}$ es *igual* a $\mathcal{B}$), $\mathcal{A} \subseteq \mathcal{B}$ ($\mathcal{A}$ es un *subconjunto* de $\mathcal{B}$), $\mathcal{A} \subset \mathcal{B}$ ($\mathcal{A}$ es un *subconjunto estricto* de $\mathcal{B}$. Los operadores anteriores admiten las negaciones $\neq, \nsubseteq, \not\subset$. De igual modo, existen operadores relacionados con operaciones entre conjuntos, a saber: $\mathcal{A} \cup \mathcal{B}$ ($\mathcal{A}$ *unión* $\mathcal{B} = \{x|x \in \mathcal{A} \vee x \in \mathcal{B}\}$), $\mathcal{A} \cap \mathcal{B}$ ($\mathcal{A}$ *intersection* $\mathcal{B} = \{x|x \in \mathcal{A} \wedge x \in \mathcal{B}\}$), $\mathcal{A} - \mathcal{B}$ ($\mathcal{A}$ *diferencia* $\mathcal{B} = \{x|x \in \mathcal{A} \wedge x \notin \mathcal{B}\}$), $\bar{\mathcal{A}}$ ($\mathcal{A}$ *complemento* $=\{x|x \in \mathcal{S} \wedge x \notin \mathcal{A}\}$).

El *conjunto potencia*, denotado $P_{\mathcal{X}}$, es el conjunto de todos los posibles subconjuntos que pueden ser definidos a partir de los elementos de un conjunto genérico $\mathcal{X}$, incluido el conjunto vacío $\emptyset$ y el mismo conjunto $\mathcal{X}$. Para conjuntos discretos finitos, la cardinalidad del conjunto potencia viene dada por $|P_{\mathcal{X}}| = 2^{|\mathcal{X}|}$.

Una forma de representar la pertenencia o no de un elemento x a un conjunto $\mathcal{X}$, es a través de la *función característica* $\mu : \mathcal{S} \rightarrow \mathcal{T}$. Nótese que si $\mathcal{T} = \{0, 1\}$ el conjunto es *clásico* y μ indica si x está contenido estrictamente (1) o no (0) en $\mathcal{X}$:

$$\mu(x) = \begin{cases} 1 & : \quad \forall x \in \mathcal{X} \\ 0 & : \quad \text{si no} \end{cases} \tag{B.1}$$

Aritmética y función de intervalo

La aritmética de intervalo es un marco teórico diseñado para manejar matemáticamente la imprecisión en computación. Su exponente más importante ha sido R. E. Moore [141], aunque las investigaciones en el área se remontan a cuarenta años antes [99, 74].

Dados dos intervalos $x = [\underline{x}, \overline{x}]$ e $y = [\underline{y}, \overline{y}]$, tales que $\underline{x} \leq \overline{x} \wedge \underline{y} \leq \overline{y} : x, y \in \mathbb{R}$, las operaciones básicas son:

$$x + y = [\underline{x} + \underline{y}, \overline{x} + \overline{y}] \tag{B.2}$$

$$x - y = [\underline{x} - \overline{y}, \overline{x} - \underline{y}] \tag{B.3}$$

$$x \times y = [min(\underline{x}\underline{y}, \underline{x}\overline{y}, \overline{x}\underline{y}, \overline{x}\overline{y}), max(\underline{x}\underline{y}, \underline{x}\overline{y}, \overline{x}\underline{y}, \overline{x}\overline{y})] \tag{B.4}$$

$$\frac{1}{x} = [1/\overline{x}, 1/\underline{x}] \quad (\text{if } \underline{x} > 0 \text{ o } \overline{x} < 0) \tag{B.5}$$

$$x \div y = x \times \frac{1}{y} \tag{B.6}$$

$$kx = [k\underline{x}, k\overline{x}] \quad (k \geq 0 \text{ es un escalar}) \tag{B.7}$$

La potenciación también es posible:

$$\begin{aligned} x^n &= [1,1] && (\text{si } n = 0) \\ &= [\underline{x}^n, \overline{x}^n] && (\text{si } \underline{x} \geq 0 \vee \underline{x} \leq 0 \leq \overline{x} \wedge n \text{ es impar}) \\ &= [\overline{x}^n, \underline{x}^n] && (\text{si } \overline{x} \leq 0) \\ &= [0, max\,(\underline{x}^n, \overline{x}^n)] && (\text{si } \underline{x} \geq 0 \vee \underline{x} \leq 0 \leq \overline{x} \wedge n \text{ es par}) \end{aligned} \tag{B.8}$$

Una aplicación importante de la aritmética de intervalos es la extensión de los conceptos anteriores a funciones. Una función intervalo (o *intervalar*) es una función que acepta como argumentos uno o más intervalos. Según el concepto de *extensión intervalar* [141], una función $F(\cdot)$ con argumentos $\mathcal{X}_1, \mathcal{X}_2, \ldots, \mathcal{X}_n$ es una extensión de una función real $f(\cdot)$ con argumentos $x_1, x_2, \ldots, x_n$ si $F(x_1, x_2, \ldots, x_n) = f(x_1, x_2, \ldots, x_n)$. En general la *propiedad de inclusión* establece que $f(x_1, x_2, \ldots, x_n) \subseteq F(\mathcal{X}_1, \mathcal{X}_2, \ldots, \mathcal{X}_n)$ para todo elemento $x_i \in \mathcal{X}_i$, $i \in \{1, \ldots, n\}$. Se desprende que el rango de $F(\cdot)$ puede ser calculado usando aritmética de intervalo [151]. Para más detalles sobre estos conceptos ver [141, 70].

B.2.3 Lógica difusa

La lógica difusa o borrosa, fue propuesta inicialmente por L. Zadeh en 1965 [226], teniendo un impacto considerable desde entonces. El postulado central de la lógica difusa es que la lógica bivalente es incapaz de representar adecuadamente muchas situaciones que incluyen incertidumbre. Por tanto, la lógica difusa propone una transición continua entre la aceptación y el rechazo, a través de grados de pertenencia. Del vasto marco teórico de la lógica difusa, solo se hace mención a los números e intervalos difusos y a su aritmética. Para obtener mayor información el lector puede consultar [105].

La lógica difusa se fundamenta en la afirmación de que la pertenencia de un elemento x a un conjunto cualquiera $\mathcal{A}$ puede graduarse desde 0 (exclusión perfecta) hasta 1 (inclusión o pertenencia perfecta). Esta idea se representa mediante la función de pertenencia $\mu : \mathcal{S} \rightarrow [0,1]$ cuya forma canónica es:

$$\mu_{\mathcal{A}}(x) = \begin{cases} f_{\mathcal{A}}(x) & : \quad \forall x \in [a,b) \\ 1 & : \quad \forall x \in [b,c] \\ g_{\mathcal{A}}(x) & : \quad \forall x \in (c,d] \\ 0 & : \quad \text{si no} \end{cases} \tag{B.9}$$

donde $a \leq b \leq c \leq d \in \mathcal{A}$, $f_{\mathcal{A}}(x) : [a,b) \rightarrow [0,1]$ y $g_{\mathcal{A}}(x) : (c,d] \rightarrow [0,1]$ son funciones reales crecientes y decrecientes respectivamente [7].

Todo conjunto difuso puede ser caracterizado por la distribución de sus elementos, mediante subconjuntos (clásicos) llamados α-cortes, los cuales están definidos para un conjunto difuso $\mathcal{A}$ como $\mathcal{A}_\alpha = \{x \in \mathcal{A} | \mu_{\mathcal{A}}(x) \geq \alpha\}$. Entre los α-cortes se destacan el *núcleo* o $\mathcal{A}_1$ y el *soporte* o $\mathcal{A}_0$. En la ec. B.9, el núcleo y el soporte son $\mathcal{A}_1 = \{x | x \in [b,c]\}$ y $\mathcal{A}_0 = \{x | x \in [a,d]\}$ respectivamente.

La cardinalidad de un conjunto difuso puede definirse de varias maneras, siendo la *cardinalidad escalar* igual a

$$|\mathcal{A}| = \sum_{x \in \mathcal{A}} \mu_{\mathcal{A}}(x) \tag{B.10}$$

o de forma relativa, respecto al conjunto universal $\mathcal{S}$ como

$$||\mathcal{A}|| = \frac{|\mathcal{A}|}{|\mathcal{S}|} \tag{B.11}$$

mientras que la *cardinalidad difusa* $C_{\mathcal{A}}(|\mathcal{A}_\alpha|)$ es el conjunto de pares $|\mathcal{A}_\alpha|$ (cardinalidad de los α-cortes) y α. Por ejemplo, si $C_{\mathcal{A}}(|\mathcal{A}_\alpha|) = \{1/1, 5/0.5, 10/0\}$, significa que $|\mathcal{A}_1| = 1$, $|\mathcal{A}_{0.5}| = 5$ y $|\mathcal{A}_0| = 10$.

Un número o un intervalo difuso viene entonces representado por una función de pertenencia tal que, para número difusos la cardinalidad del núcleo es la unidad, mientras que para los conjuntos es mayor. Tanto los números como los conjunto difusos se combinan en términos de sus α-cortes y las reglas de aritmética de intervalo, según las expresiones [101]:

$$\mathcal{A}_\alpha + \mathcal{B}_\alpha = [\underline{a}, \overline{a}]_\alpha + [\underline{b}, \overline{b}]_\alpha = [\underline{a} + \underline{b}, \overline{a} + \overline{b}]_\alpha \tag{B.12}$$

$$\mathcal{A}_\alpha - \mathcal{B}_\alpha = [\underline{a}, \overline{a}]_\alpha - [\underline{b}, \overline{b}]_\alpha = [\underline{a} - \overline{b}, \overline{a} - \underline{b}]_\alpha \tag{B.13}$$

$$\mathcal{A}_\alpha \times \mathcal{B}_\alpha = [\underline{a}, \overline{a}]_\alpha \times [\underline{b}, \overline{b}]_\alpha = [l(\mathcal{A}_\alpha, \mathcal{B}_\alpha), u(\mathcal{A}_\alpha, \mathcal{B}_\alpha)]_\alpha \tag{B.14}$$

$$\mathcal{A}_\alpha \div \mathcal{B}_\alpha = [\underline{a}, \overline{a}]_\alpha \div [\underline{b}, \overline{b}]_\alpha = [\underline{a}, \overline{a}]_\alpha \times [1/\underline{b}, 1/\overline{b}]_\alpha \tag{B.15}$$

donde

$$\begin{aligned} l(\mathcal{A}_\alpha, \mathcal{B}_\alpha) &= min(\underline{ab}, \underline{a}\overline{b}, \overline{a}\underline{b}, \overline{a}\overline{b}) \\ u(\mathcal{A}_\alpha, \mathcal{B}_\alpha) &= max(\underline{ab}, \underline{a}\overline{b}, \overline{a}\underline{b}, \overline{a}\overline{b}) \end{aligned}$$

y $\mathcal{A}_\alpha \div \mathcal{B}_\alpha$ aplica si $0 \notin [\underline{b}, \overline{b}]_\alpha$. Existen algunas restricciones para las expresiones anteriores [101]. En todo caso, mediante las mismas es posible propagar la incertidumbre [40, 63, 98].

B.2.4 Teoría de la posibilidad

La noción de transiciones graduales que sustenta la lógica difusa es aplicable más allá del contexto de cuantificación de pertenencias, siendo posible cualificar y cuantificar la posibilidad, mediante *medidas de posibilidad* o *distribuciones de posibilidad*.

Considérese el conjunto $\mathcal{E}$ y su función característica $\mu_{\mathcal{E}}(x)$, la afirmación $y \in \mathcal{E}$ tiene posibilidad $\mu_{\mathcal{E}}(y)$, en consecuencia posibilidad cero implica $y \notin \mathcal{E}$ mientras que posibilidad uno implica $y \in \mathcal{E}$. Esta visión binaria de la posibilidad es propia de la *Teoría clásica de posibilidad*. Por ejemplo, si x es un valor único tal que $x \in \mathcal{E}$, entonces [41]:

$$\Pi_{\mathcal{E}}(\mathcal{A}) = \begin{cases} 1 & \text{if} \quad \mathcal{A} \cap \mathcal{E} \neq \emptyset \\ 0 & \text{si no} \end{cases} \tag{B.16}$$

lo que significa que $\mathcal{A}$ tiene la posibilidad de contener a x si $\Pi_{\mathcal{E}}(\mathcal{A}) = 1$.

Alternativamente, si se adopta la lógica difusa, la función característica se convierte en una función de pertenencia, posibilitando la *Teoría de posibilidad gradual*. Siguiendo el ejemplo anterior, la posibilidad de $\mathcal{A}$ se calcula en términos del grado de pertenencia supremo de la intersección entre $\mathcal{A}$ y el núcleo de $\mathcal{X}$:

$$\Pi_{\mathcal{E}}(\mathcal{A}) = \sup_{x \in \mathcal{A}} \mu_{\mathcal{E}}(x) \tag{B.17}$$

De forma complementaria, una función llamada *necesidad* se define como la medida dual, respecto al complemento $\bar{\mathcal{A}}$ del conjunto de interés. Así, la necesidad $N_{\mathcal{E}}(\mathcal{A})$ es :

$$N_{\mathcal{E}}(\mathcal{A}) = 1 - \Pi_{\mathcal{E}}(\bar{\mathcal{A}}) = \begin{cases} 1 & \text{if} \quad \mathcal{E} \subseteq \mathcal{A} \\ 0 & \text{si no} \end{cases} \tag{B.18}$$

Si bien la posibilidad y la probabilidad se parecen, las afirmaciones posibilísticas son menos tajantes, ya que alguien puede conocer que algo es posible sin saber su probabilidad. Saber que algo es 'posible' admite varias interpretaciones que tienen eco en la teoría de probabilidad [41]. De aquí que la teoría de posibilidad pueda ser empleada para manejar tanto incertidumbre epistémica como aleatoria.

B.2.5 Teoría de la probabilidad

La teoría (clásica) de probabilidad es la más conocida y empleada de todas, razón por la cual en este apartado solo se mencionan los elementos básicos, siguiendo a [142, 7, 165].

La probabilidad es un índice que toma valores en $[0, 1]$ y que denota nuestro grado de creencia acerca de la ocurrencia de un suceso. Tal creencia puede recaer sobre un conocimiento previo de la frecuencia del suceso (*probabilidad objetiva*) o puede expresar el grado de confianza de un sujeto respecto a la ocurrencia de un evento (*probabilidad subjetiva*). Dicho de otro modo, la probabilidad es una proyección de lo que creemos que sucederá, alimentada por nuestro conocimiento del pasado y nuestra interpretación del mismo.

Las reglas que aseguran la coherencia de los postulados en la teoría clásica de probabilidad comprenden las asignaciones de probabilidad (normadas por los axiomas de Kolmogorov) y las relaciones entre eventos. Respecto al primer grupo tenemos:

Axioma 1: $0 \leq P(A) \leq 1$.

Axioma 2: $P(S) = 1$, donde S es el universo.

Axioma 3: Para una secuencia de eventos disjuntos $A_1, A_2, \ldots$ se verifica que $P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i), \quad n = 1, 2, \ldots, \infty$.

Los eventos se relacionan en términos de *dependencia* e *independencia*. En el primer caso la ocurrencia de un evento no influencia la probabilidad de ocurrencia del otro, mientras que en el segundo caso sí. Dado el universo S y su conjunto potencia P_S se tiene que las relaciones entre evento verifican las siguientes propiedades:

Propiedad 1: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

Propiedad 2: Si A y B son disjuntos o mutuamente excluyentes, entonces $P(A \cup B) = P(A) + P(B)$.

Propiedad 3: $P(\bar{A}) = 1 - P(A)$.

Propiedad 4: La probabilidad condicional de A dada la ocurrencia de B viene expresada por $P(A|B) = \frac{P(A \cap B)}{P(B)}$.

Propiedad 5: Si dos eventos son disjuntos, entonces $P(A \cap B) = P(A) \cdot P(B)$.

La actualización de conocimiento viene dada por el teorema de Bayes':

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \tag{B.19}$$

o de forma más general,

$$P(A_i|B) = \frac{P(B|A_i)P(A_i)}{\sum_j P(B|A_j)P(A_j)} \tag{B.20}$$

Dado un fenómeno, la abstracción fundamental es la de *variable aleatoria*, es decir una variable que adopta valores de forma aleatoria pero de acuerdo a algún patrón de comportamiento, el cual es representado mediante las funciones o leyes de probabilidad. Así, una *función de distribución acumulada* (CFD) del tipo $\mathtt{F}(x) \equiv P(\mathtt{X} \leq x)$ da la probabilidad de que la variable aleatoria $\mathtt{X}$ sea menor o igual a x. Esta relación también puede ser expresada en términos de las derivadas: para *variables aleatorias discretas* la derivada es llamada *función de masa de probabilidad* (PMF) y da la probabilidad de que $\mathtt{X}$ sea igual a x: $\mathtt{P}(x) = P(\mathtt{X} = x)$. De forma similar, para *variables aleatorias continuas* la *función densidad de probabilidad* (PDF) representa la densidad de probabilidad de x, la cual es cero para valores puntuales, y toma valores no nulos para intervalos.

Si bien existen diferencias en las repercusiones de modelar conjuntos o intervalos asociados o no a variables aleatorias [47], en ambos casos se suelen emplear valores representativos llamados *momentos*, los cuales se emplean para caracterizar distribuciones de probabilidad.

El primer grupo de momentos representa ejes de simetría:

Valor esperado o media Es el primer momento central, definido como

$$\mu = \int_{\mathtt{X}} x\mathtt{f}(x)dx \tag{B.21}$$

$$\mu = \sum_{i=1}^{n} x_i \mathtt{P}(x_i) \tag{B.22}$$

para PDF y PMF respectivamente. En la práctica puede estimarse con m muestras de x haciendo

$$\bar{x} = \frac{1}{m}\sum_{i=1}^{m} x_i \tag{B.23}$$

Mediana La mediana x_m de m muestras es el valor que divide al grupo de datos en dos conjuntos de la misma cardinalidad, uno de los cuales tiene sus elementos por debajo de x_m y el otro por encima de x_m.

El segundo grupo caracteriza la variabilidad:

Varianza Segundo momento central, definido como

$$\sigma^2 = \int_{\mathtt{X}} (x-\mu)^2 \mathtt{f}(x)dx \tag{B.24}$$

$$\sigma^2 = \sum_{\mathtt{X}} (x_i-\mu)^2 \mathtt{P}(x_i) \tag{B.25}$$

para PDF y PMF respectivamente. En ocasiones es expresada en términos de la *desviación estándar* σ, que es el cuadrado de la varianza. En la práctica es estimado como con m muestras de x como

$$S^2 = \frac{1}{m-1}\sum_{i=1}^{m}(x_i - \bar{x})^2 \tag{B.26}$$

Coeficiente de variación (COV) Está definido como COV $= \frac{S}{\bar{x}}$ y es una cantidad adimensional que puede ser interpretada como el tamaño de la desviación real respecto a $\bar{x}$ en términos porcentuales. Así, si se pretende reducir incertidumbre, mientas menor sea el COV, mejor.

La tabla 2.4 presenta las CDF, PDF y principales momentos de las distribuciones más empleadas en ingeniería.

Finalmente, las distribuciones también puede ser caracterizadas mediante parámetros llamados fractiles o cuantiles. El fractil x_p de una muestra es el valor tal que $p\%$ de los datos es menor o igual a x_p. En términos de probabilidad: $P(\mathtt{X} \leq x_p) = p\%/100$.

B.2.6 Teoría de la evidencia de Dempster-Shafer

Teoría de la evidencia de Dempster-Shafer (DST) fue introducida por Shafer [186] basándose en los desarrollos de Dempster [39]. Su estructura es considerada una generalización de la teoría clásica de probabilidad, donde la masa reside en conjuntos en lugar de puntos precisos, como en las PMF. En consecuencia, las asignaciones se hacen a conjuntos de eventos, llamados *puntos focales*, sin discriminar la probabilidad de los elementos que los constituyen, razón por la cual es considerada más apropiada para modelar incertidumbre epistémica.

Los tres elementos fundamentales en la DST son:

La *asignación básica*: Denotada como bpa o m, es la cantidad de evidencia que sustenta la afirmación de que un elemento particular del conjunto universal $\mathcal{S}$ forma parte del conjunto potencia $P_{\mathcal{S}}$ [183, pg. 13]. Matemáticamente se dice que los bpa son mapas $m : P_{\mathcal{S}} \rightarrow [0,1]$ tales que

$$m(\emptyset) = 0 \tag{B.27}$$

$$\sum_{\mathcal{A}\in P_{\mathcal{S}}} m(\mathcal{A}) = 1 \tag{B.28}$$

Del mismo modo, el bpa puede determinarse a partir de las medidas de creencia vía transformación de Möbius:

$$m(A) = \sum_{\mathcal{B}|\mathcal{B}\subseteq\mathcal{A}} (-1)^{|\mathcal{A}-\mathcal{B}|}\mathrm{Bel}(\mathcal{B}) \tag{B.29}$$

La *medida de creencia*: También llamada *soporte* y denotada Bel, es la suma de todas las masas de los subconjuntos del conjunto de interés:

$$\mathrm{Bel}(\mathcal{A}) = \sum_{\mathcal{B}|\mathcal{B}\subseteq\mathcal{A}} m(\mathcal{B}) \tag{B.30}$$

Las medidas de creencia cumplen con

$$\mathrm{Bel}(\emptyset) = 0 \tag{B.31}$$
$$\mathrm{Bel}(\mathcal{S}) = 1 \tag{B.32}$$

Finalmente, las creencias son superaditivas, es decir $\mathrm{Bel}(\mathcal{A}\cup\mathcal{B}) \geq \mathrm{Bel}(\mathcal{A}) + \mathrm{Bel}(\mathcal{B})$ para $\mathcal{A}\cap\mathcal{B} = \emptyset$.

La *medida de plausibilidad* Representada como Pls, es la suma de las masas de todas los conjuntos que intersecan al conjunto de interés:

$$\mathrm{Pls}(\mathcal{A}) = \sum_{\mathcal{B}|\mathcal{B}\cap\mathcal{A}\neq\emptyset} m(\mathcal{B}) \tag{B.33}$$

La plausibilidad cumple con

$$\mathrm{Pls}(\emptyset) = 0 \tag{B.34}$$
$$\mathrm{Pls}(\mathcal{S}) = 1 \tag{B.35}$$
$$\mathrm{Pls}(\mathcal{A}\cup\mathcal{B}) \geq \mathrm{Pls}(\mathcal{A}) + \mathrm{Pls}(\mathcal{B}) \tag{B.36}$$

Las relaciones entre Pls y Bel vienen dadas por:

$$\mathrm{Pls}(\mathcal{A}) \geq \mathrm{Bel}(\mathcal{B}) \tag{B.37}$$
$$\mathrm{Pls}(\mathcal{A}) = 1 - \mathrm{Bel}(\overline{\mathcal{A}}) \tag{B.38}$$
$$\mathrm{Pls}(\overline{\mathcal{A}}) = 1 - \mathrm{Bel}(\mathcal{A}) \tag{B.39}$$
$$\mathrm{Bel}(\mathcal{A}) = 1 - \mathrm{Pls}(\overline{\mathcal{A}}) \tag{B.40}$$
$$\mathrm{Bel}(\overline{\mathcal{A}}) = 1 - \mathrm{Pls}(\mathcal{A}) \tag{B.41}$$

B.2.7 Probabilidades imprecisas y p-box

Existen varias teorías sobre probabilidades imprecisas, siendo una de las más famosas la desarrollada por P. Walley [207], basado en las ideas de Keynes, Smith y Good [211, pg. 462]. Existe alrededor de esta teoría y otras afines una comunidad de investigación con cada vez mayor prestigio y relevancia (ver su sitio web `http://www.sipta.org`).

En la teoría clásica de probabilidad, la probabilidad de un evento x es un mapa desde el conjunto universal a un valor preciso en [0,1], mediante una distribución de probabilidad. En la teoría de probabilidades imprecisas, una relación tal se tiene por imposible en la práctica, admitiendo en cambio una clase $\mathfrak{M}$ de distribuciones de probabilidades plausibles para relacionar al evento x. Así, la *probabilidad inferior* $\underline{P}(x)$ y la *probabilidad superior* $\overline{P}(x)$ son definidas como el ínfimo y el supremo de las probabilidades $P(x)$ respectivamente, sobre toda distribución P contenida en $\mathfrak{M}$ [211, pg. 462]:

$$\underline{P}(x) = \inf\{P(x)|P \in \mathfrak{M}\} \tag{B.42}$$
$$\overline{P}(x) = \sup\{P(x)|P \in \mathfrak{M}\} \tag{B.43}$$

Un caso particular de probabilidades imprecisas son las cajas de probabilidad (*probability boxes*) o p-box. Están compuestas por dos funciones limitantes no decrecientes $\overline{F} : \mathbb{R} \to [0,1]$ y $\underline{F} : \mathbb{R} \to [0,1]$, tal que la p-box $[\overline{F}, \underline{F}]$ denota o

contiene toda CDF desconocida $F(x)$ definida para la variable aleatoria $\mathtt{X}$ (ver fig. 2.10). Debe recalcarse que las p-boxes permiten expresar diversos niveles de incertidumbre; por ejemplo si se conoce que la forma de la distribución es la de una normal, pero se desconocen sus parámetros, entonces la p-box adoptará la forma $N([\underline{\mu}, \overline{\mu}], [\underline{\sigma^2}, \overline{\sigma^2}])$. En consecuencia toda distribución normal con media en $[\underline{\mu}, \overline{\mu}]$ y varianza en $[\underline{\sigma^2}, \overline{\sigma^2}]$ cumple con la definición de la p-box.

B.2.8 Modelos Info-gap

Los modelos info-gap es una metodología no probabilísticas propuesta por Y. Ben-Haim que pretende cuantificar la disparidad entre lo que conoce la DM y lo que no [13]. Un modelo info-gap define una familia de conjuntos anidados $\mathcal{U}(\alpha, \tilde{\Phi})$, $\alpha \geq 0$ de vectores o funciones $\tilde{\Phi}(\mathbf{x})$ que aproximan o estiman al valor verdadero $\Phi(\mathbf{x})$ de un parámetro de interés. Entonces, a medida que el *parámetro de incertidumbre* α crece, el conjunto se vuelve más inclusivo [14]: $\alpha \leq \alpha' \Rightarrow \mathcal{U}(\alpha, \tilde{\Phi}) \subseteq \mathcal{U}(\alpha', \tilde{\Phi})$.

A partir de allí, Ben-Haim prescribe dos funciones que miden la *robustez* $\hat{\alpha}(x, r_c)$ y la *oportunidad* $\hat{\beta}(x, r_c)$ para el nivel actual de incertidumbre:

$$\hat{\alpha}(x, r_c) = \max\left\{\alpha : \left(\min_{\tilde{\Phi} \in \mathcal{U}(\alpha, \tilde{\Phi})} R(\mathbf{x}, \Phi)\right) \geq r_c\right\} \tag{B.44}$$

$$\hat{\beta}(x, r_w) = \min\left\{\alpha : \left(\max_{\tilde{\Phi} \in \mathcal{U}(\alpha, \tilde{\Phi})} R(\mathbf{x}, \Phi)\right) \geq r_w\right\} \tag{B.45}$$

La robustez de una alternativa $\mathbf{x}$ refleja el máximo nivel de incertidumbre α para el cual el beneficio R no es inferior a r_c. De modo similar, $\hat{\beta}(x, r_w)$ indica el nivel más bajo de deficiencia de información para el cual el beneficio del sistema puede alcanzar r_w [13]. Algunos ejemplos pueden hallarse en [14, 13].

B.3 Conceptos fundamentales de Optimización Evolutiva Multiobjetivo

En esta sección se resumen las ideas y conceptos presentados en el capítulo 3.

B.3.1 Fundamentos de toma de decisiones y optimización multiobjetivo

La toma de decisiones es el proceso en el cual la unidad de decisión (DM), con el apoyo de un analista, realiza una valoración de las alternativas disponible, con el fin de resolver alguna problemática de decisión, entre las que figuran: (1) *escoger* un subconjunto de mejores alternativas, (2) *ordenar* el conjunto de alternativas en subconjuntos de acuerdo a ciertas normas preestablecidas, y (3) *jerarquizar* las alternativas disponibles [81, 80]. Sea cual sea el tipo de problema a resolver, el principio rector establece que una DM racional pretende maximizar su nivel de beneficio o satisfacción [230, 229].

Cuando la toma de decisiones se fundamenta sobre métodos formales de ayuda a la decisión, las necesidades y aspiraciones de la DM, llamadas *preferencias*, deben ser modeladas matemáticamente. En este proceso se distinguen los

atributos, que se refieren a las características sobre las cuales se basan los juicios de la DM; los *objetivos*, que definen las direcciones de mejora de los atributos; las *metas*, que establecen niveles mínimos de beneficio para los objetivos, y finalmente los *criterios*, que son combinaciones de atributos, objetivos y metas [163]. La integración de dichos elementos permite formular un programa matemático como este:

$$\begin{aligned} & Opt\ F(\mathbf{x}) \\ \text{s.t.:} \quad & \\ & G(\mathbf{x}) \leq 0 \\ & H(\mathbf{x}) = 0 \end{aligned} \tag{B.46}$$

donde *Opt* indica la dirección de mejora (*maximización* o *minimización*) de la función $F(\mathbf{x})$. El dominio factible $\mathbf{X}$ está definido por $G(\mathbf{x})$ y $H(\mathbf{x})$ los cuales son respectivamente vectores de inecuaciones y ecuaciones.

En problemas con múltiples criterios, que son tema de estudio en este trabajo, la función $F(\mathbf{x})$ se convierte en un vector de funciones objetivo, y el programa matemático adquiere la forma:

$$\begin{aligned} & Opt\left(F(\mathbf{x}) = (f_1(\mathbf{x}), f_2(\mathbf{x}), \ldots, f_k(\mathbf{x}))^t\right) \\ \text{s.t.:} \quad & \\ & G(\mathbf{x}) \leq 0 \\ & H(\mathbf{x}) = 0 \end{aligned} \tag{B.47}$$

donde F es un vector de funciones $f_i : \mathbb{R}^n \to \mathbb{R}$ tal que $F : \mathbf{X} \to \mathbf{Y}$ relaciona a un vector de n variables de decisión $\mathbf{x} = (x_1, x_2, \ldots, x_n)^t$ (llamado *vector de decisión*, *vector de soluciones* o simplemente *alternativa*) en el *espacio de decisión* $\mathbf{X}$, definido por los vectores de restricciones $G(\mathbf{x})$ de g inecuaciones y $H(\mathbf{x})$ de h ecuaciones, con un *vector de objetivos* k-dimensional $\mathbf{y} = (y_1, y_2, \ldots, y_k)^t$ en el *espacio de objetivos* $\mathbf{Y} \subseteq \mathbb{R}^k, k \in \mathbb{N}$ [238].

Cuando los objetivos están en conflicto, esto es cuando la mejora de uno de ellos conlleva a la desmejora de algún otro, los métodos de resolución de problemas multicriterio conllevan a un conjunto de soluciones óptimas, llamado conjunto *eficiente*, *no dominado* o simplemente conjunto de *Pareto*. Tal conjunto representa los diferentes consensos posibles entre atributos. Por otro lado, la búsqueda del mismo es posible mediante distintos métodos, como la *optimización multiobjetivo*, que conduce al conjunto de Pareto, la *programación compromiso* que redirige la búsqueda hacia un punto ideal de referencia [228, 229], o la *programación por metas*, que propone una serie de mejoras secuenciales. La tabla B.3 resume las ideas expuesta en esta sección. Para más información el lector puede consultar [3, 100, 129, 163, 168, 171, 230].

Las definiciones fundamentales de calidad que relacionan optimización multiobjetivo y EA son:

Dominancia de Pareto: $\mathbf{x}_1$ domina a $\mathbf{x}_2$, denotado $\mathbf{x}_1 \succ \mathbf{x}_2$, si y solo si $f_i(\mathbf{x}_1) \leq f_i(\mathbf{x}_2) \wedge F(\mathbf{x}_1) \neq F(\mathbf{x}_2)$; $i \in \{1, 2, \ldots, k\}$. Si no existe un vector que domine a $\mathbf{x}_1$, entonces $\mathbf{x}_1$ es *no dominado*.

Dominancia débil de Pareto: $\mathbf{x}_1$ domina débilmente a $\mathbf{x}_2$, denotado $\mathbf{x}_1 \succeq \mathbf{x}_2$, si y solo si $f_i(\mathbf{x}_1) \leq f_i(\mathbf{x}_2)$; $i \in \{1, 2, \ldots, k\}$.

Óptimo de Pareto: Un vector solución $\mathbf{x}^* \in \mathbf{X}$ es óptimo de Pareto si y solo si $\neg\exists x \in X : \mathbf{x} \succeq \mathbf{x}^*$.

Actores
Unidad de decisión (DM)
Entidad (persona o grupo) con la autoridad para tomar la decisión final (selección) en el problema actual.
Analista
Entidad que ayuda al DM en modelar el problema, resolviendo el programa matemático e identificando la alternativa final.
Elementos
Atributos
Características valiosas del sistema que permiten al DM diferenciar las alternativas.
Objetivos
Direcciones de mejora de los atributos (*maximización* o *minimización*).
Metas
Niveles mínimos esperados en la mejora de los atributos.
Criterios
Principios de juicio construidos a partir de la combinación de atributos, objetivos y metas.
Tipos de problemas
Monocriterio
Optimización ordinaria.
Multicriterio
Programación multiobjetivo (búsqueda del conjunto de alternativas eficientes).
Programación compromiso (búsqueda de la alternativa más cercana al punto ideal).
Programación por metas (incremento paulatino de la satisfacción o beneficio).

Table B.3: Fundamentos de toma de decisiones.

Conjunto de aproximación de Pareto: $\{\mathbf{x}^*|\neg\exists\mathbf{x} : \mathbf{x} \succ \mathbf{x}^*;\ \mathbf{x}^*, \mathbf{x} \in \mathbf{D} \subseteq \mathbf{X}\}$

Frente o frontera de Pareto: $\{F(\mathbf{x}^*)|\neg\exists\mathbf{x} : \mathbf{x} \succ \mathbf{x}^*;\ \mathbf{x}^*, \mathbf{x} \in \mathbf{D} \subseteq \mathbf{X}\}$.

Frontera y conjunto real de Pareto Un conjunto de aproximación tal que $\mathbf{D} = \mathbf{X}$ es una conjunto real de Pareto. Su imagen es la frontera real de Pareto.

ϵ-dominancia [117, 116]: $\mathbf{x}^1$ domina $\mathbf{x}^2$, denotado $\mathbf{x}^1 \succ_\epsilon \mathbf{x}^2$, si y solo si para $i \in \{1, 2, \ldots, k\}$ y $\epsilon_i > 0$ se verifica que $f_i(\mathbf{x}^1) \leq (1 + \epsilon_i) f_i(\mathbf{x}^2)$ (forma multiplicativa) o $f_i(\mathbf{x}^1) \leq f_i(\mathbf{x}^2) + \epsilon_i$ (forma aditiva).

B.3.2 Algoritmo evolutivos de optimización multiobjetivo

Algoritmos evolutivos

En Inteligencia artificial, los Algoritmos evolutivos (EA) son propios de una corriente de métodos de ensayo y error no estadísticos. En particular, los EA designan una clase de procedimientos estocásticos de aprendizaje que *evolucionan* hacia un estado superior de conocimiento del objeto de estudio, mediante la aplicación iterativa de reglas *heurísticas* de aprendizaje, inspiradas en mecanismos naturales, sobre todo la dinámica de los seres vivos. La explotación de dicha capacidad de aprendizaje para solucionar problemas de optimización, hace posible identificar óptimos locales.

Naturaleza		EA
Selección natural: ¿cómo sobrevivir?	$\Rightarrow$	**Selección artificial:** ¿cómo aproximarse al óptimo?
Individuo: ser vivo.	$\Rightarrow$	**Individuo:** vector de decisión.
Origen poblacional: desconocido.	$\Rightarrow$	**Origen poblacional:** Muestra inicial.
Codificación natural: ADN.	$\Rightarrow$	**Codificación artificial:** esquema de representación.
Mecanismo natural de reproducción: cruce, mutación, clonación.	$\Rightarrow$	**Mecanismos de recombinación:** emulación del cruce y la mutación.
Unidad temporal: generaciones.	$\Rightarrow$	**Unidad temporal:** iteraciones.

Table B.4: Abstracciones de los postulados Darwinistas en EA: cómo la vida es emulada para solucionar problemas de optimización mediante EA (adaptado de [135]).

Para entender la filosofía subyacente en los EA, considérense los principios de la evolución natural postulados por Charles Darwin: La *supervivencia* es el problema fundamental de los seres vivos, el cual es abordado en la naturaleza tanto en un ámbito individual como grupal, mediante la *selección natural*. Así, una población aprende a sobrevivir (*adaptación*) mediante un proceso iterativo donde los mejores individuos poseen mayor oportunidad de reproducirse (*selección natural*), transfiriendo de este modo sus características ventajosas a las nuevas generaciones. En consecuencia, si la sobrevivencia es vista como optimización y los individuos como vectores de decisión, el proceso iterativo tenderá a localizar óptimos locales si los mecanismos de *selección* y *recombinación* genética son adecuadamente emulados [135]. La tabla B.4 explica cómo los postulados de Darwin son interpretados en las *metaheurísticas*.

Los EA trabajan sobre tres espacios, a saber: el espacio de representación Ω compuesto por *individuos* I, el espacio de decisión $\mathbf{X}$ y el de objetivos $\mathbf{Y}$. La función $\rho(I) : \Omega \rightarrow \mathbf{X}' \subseteq \mathbf{X}$ llamada *representación* o *codificación* de los individuos, emula los codificación genética hallada en el ADN. En la práctica, toda representación se basa en un alfabeto numérico o simbólico finito, por lo que toda dimensión definida a partir de dicho alfabeto es también finita y discreta en la práctica. Los métodos de representación más usados en EA son la representación *binaria* (los individuos son números binarios), la representación *Gray* (los individuos son números binarios con reglas adicionales), la representación *real* ($\rho(I) = \mathbf{x}$ y $\Omega = \mathbf{X}'$) y la *entera* (los individuos están representadas por una lista finita de estados).

La estructura algorítmica fundamental de un EA puede verse en la figura B.2. Otras versiones en pseudocódigos para EA generales pueden hallarse en [181, pg. 21][110, pgs. 28-29][203, pg. 2-19][235, pgs. 21-22]. La primera operación realizada es la inicialización ($\texttt{Initialize}(P^{(0)})$) que consiste en tomar una muestra aleatoria del espacio de representación Ω para completar la población inicial $P^{(0)}$ de M individuos. Luego los tres procedimientos principales de selección ($\texttt{Select}(P^{(t)}, F_a(I), M)$), generación ($\texttt{Generate}(M_p, M)$) y actualización ($\texttt{Update}(P^{(t+1)} + P^{(t)}, F_a(I), M)$) se repiten hasta que se alcanza la condición de parada. Obsérvese que `Select`, que llena al archivo de recombinación M_p, también llamado *matting pool*, y `Update` dependen de la función de adaptación $F_a(I)$, la cual es una expresión de la calidad de un individuo I en términos de

Algoritmo B.1 Algoritmo evolutivo: $\mathbf{EA}(F_a(I), M, t^{max})$

```
    /* Inicialización */
1:  Initialize(P^(0))
2:  t ← 0
    /* Exploración */
3:  while not Terminate(P^(t), t, t^max) do
4:    M_p := Select(P^(t), F_a(I), M)          /* Selecciona los individuos a recombinar */
5:    P^(t+1) := Generate(M_p, M)              /* Genera una nueva muestra (descendencia) */
6:    P^(t+1) := Update(P^(t+1) + P^(t), F_a(I), M)   /* Actualiza la población */
7:    t ← t + 1
8:  end while
9:  Output(P^(t))
```

Figure B.2: Algoritmo B.1: Algoritmo evolutivo.

su calidad (su distancia al óptimo). Dicha adaptación está relacionada con la función objetivo $F : \mathbf{X} \rightarrow \mathbf{Y}$ y sirve para jerarquizar a la población con el fin de designar a los individuos que habrán de recombinarse en `Generate` y los que serán conservados o eliminados mediante `Update`. `Generate` invoca algunos operadores tradicionales de recombinación llamados *cruce* y *mutación*, los cuales no son más que operadores que actúan respectivamente sobre varios o un único individuo. Por último, `Update` califica a los individuos de la unión $P^{(t+1)} + P^{(t)}$, reduciéndola a M individuos. Se debe recalcar que, aun cuando los últimos dos procedimientos dependen de la adaptación de los individuos, la manera en la que ellos manejan dichos valores no es necesariamente igual.

Algoritmos evolutivos de optimización multiobjetivo

En Optimización evolutiva multiobjetivo (EMO), las metaheurísticas para resolución de problemas multiobjetivo reciben el nombre de Algoritmo evolutivo de optimización multiobjetivo (MOEA). La figura B.3 esquematiza la estructura de un MOEA moderno, también llamado de 2ª generación [27]. En términos generales, todo MOEA intenta solucionar problemas multiobjetivo atendiendo a tres preocupaciones fundamentales: (1) cómo muestrear el espacio de decisión, (2) cómo jerarquizar las alternativas y (3) cómo almacenar las buenas soluciones halladas. La forma en la cual se afrontan las preocupaciones anteriores es la que distingue a un MOEA de otro. El primer punto se refiere a cómo explorar el espacio de decisión de forma eficiente, para lo cual se plantean distintos *motores de búsqueda*, entre los que se encuentran, además de los algoritmos genéticos, los algoritmos de hormigas [24], la evolución diferencial [154, 219, 78] y los algoritmos de partículas [153] entre los más relevantes, además de métodos no evolutivos, como el recocido simulado (ver fig. B.8). El segundo punto hace referencia a la calidad global (cercanía a la frontera real de Pareto) y local (densidad de la vecindad) de las soluciones. Finalmente, el tercer punto hace referencia a cómo almacenar las buenas soluciones para que puedan ser presentadas a la DM al finalizar la ejecución del MOEA.

El control de calidad en los MOEA moderno está íntimamente ligado con el concepto de dominancia, de forma tal que las soluciones no dominadas siempre son preferidas, y entre éstas, aquellas que se ubican en zonas poco densas o en

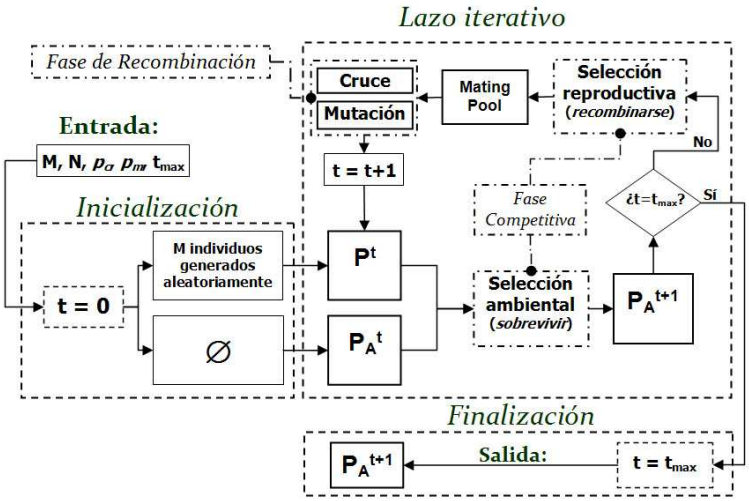


Figure B.3: Diagrama de un MOEA de 2ª generación basado en GA (tomado de [174, pg. 1060]).

Algoritmo B.2 Algoritmo de almacenamiento finito: **Archive₁**(P_A, I, N)

1: Si I no está dominado por los miembros de P_A entonces:
2: $P_A := P_A + I$ /* Añadir los nuevos individuos */
3: Si I domina a algún miembro I' de P_A entonces $P_A := P_A - I'$
4: Si `Size`$(P_A) > N$ entonces:
5: `Truncation`(P_A) /* Se reduce el archivo a N individuos */
6: Si no ir a 8
7: Si no `AddsDominated`(P_A, I, N) /* Decide si se añade el nuevo individuo */
8: `Output`(P_A)

Figure B.4: Algoritmo B.2: Algoritmo de almacenamiento en archivo finito.

puntos extremos. No obstante, si parte de la muestra o población es dominada, se debe implementar algún criterio, característico de cada MOEA, para manejar tales soluciones. Del mismo modo, en la mayoría de los casos, no es posible almacenar todas las soluciones no dominadas que se puedan conseguir por restricciones de recursos. En tal caso, si se tienen más soluciones no dominadas de las que se requieren o se puede manejar, es necesario eliminar algunas de ellas y almacenar el resto en el *archivo de no dominadas*.

La figura B.4 describe en pseudocódigo el procedimiento de almacenamiento en un archivo de tamaño finito. Nótese que el procedimiento `Truncation`(P_A), cuya implementación es característica de cada MOEA, reduce la población almacenada hasta un valor predeterminado de soluciones según algún criterio. Por otro lado, `AddsDominated`(P_A, I, N) incorpora en el archivo a una solución no dominada cuando el tamaño máximo del archivo no ha sido alcanzado, o cuando habiéndose alcanzado, es posible eliminar alguna solución dominada del mismo o una no dominada de inferior calidad en término de densidad (nunca de optimalidad).

Existen variantes de almacenamiento como el procedimiento llamado *Hy-*

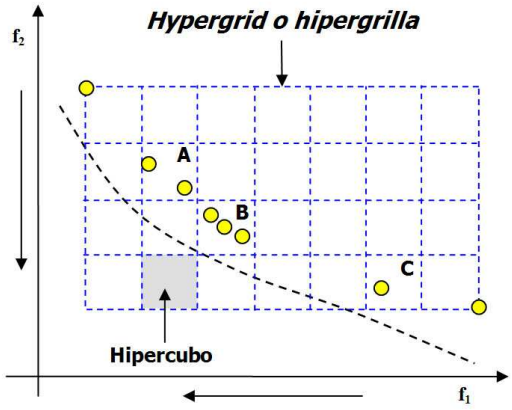


Figure B.5: Almacenamiento mediante *Hypergrid* según Knowles et al. [107] (tomado de [181, pg. 35]).

pergrid (ver B.5) y propuesto en [107]. En este método de almacenamiento, el archivo solo admite soluciones no dominadas. La densidad, por otro lado, es controlada dividiendo el espacio de objetivos en *hipercubos* de igual tamaño. Cuando es necesario truncar al archivo, se elimina aleatoriamente alguna solución perteneciente al hipercubo de mayor densidad, es decir con mayor número de soluciones.

NSGA-II y MOSA

De entre los MOEA más representativos en la actualidad, el NSGA-II y el MOSA sustentan los desarrollos propuestos en este trabajo.

El *Nondominated Sorting GA II* (NSGA-II) [36] emplea GA e implementa una clasificación por profundidad (fig. B.6a) y una subclasificación por densidad, basada en un operador llamado *crowding* (fig. 3.13b), que otorga mayor adaptación a los individuos ubicados en un rango más cercano al óptimo de Pareto, mientras que entre los del mismo rango, beneficia a aquellos para los cuales la distancia de Manhattan entre sus dos vecinos inmediatos es mayor. La figura B.7 presenta nuestra implementación del NSGA-II. Nótese que la función de adaptación $F_a(I)$ corresponde al operador *crowding*. El procedimiento `Rank`$(P^{(t)}, F_a(I))$ ordena su argumento de acuerdo a $F_a(I)$. Por su parte, `Update`$(P^{(t)}, F_a(I), N)$ asigna los N mejores individuos del argumento a un nuevo archivo, mientras que `Select`$(P_A^{(t+1)}, F_a(I), M)$ llena al *mating pool* con $M/2$ pares, cada uno de los cuales produce dos nuevos individuos. Esta implementación ha sido utilizada en [175, 174, 178, 176, 179, 58, 158, 234].

Por su parte el *Multiple-Objective Simulated Annealing* o recocido simulado multiobjetivo (MOSA) es un algoritmo específicamente diseñado para problemas de optimización combinatoria. Está basado en el recocido simulado monocriterio, pero introduce un archivo para soluciones no dominadas. La fig. B.8 presenta el pseudocódigo del MOSA propuesto por Ulungu et al. [201].

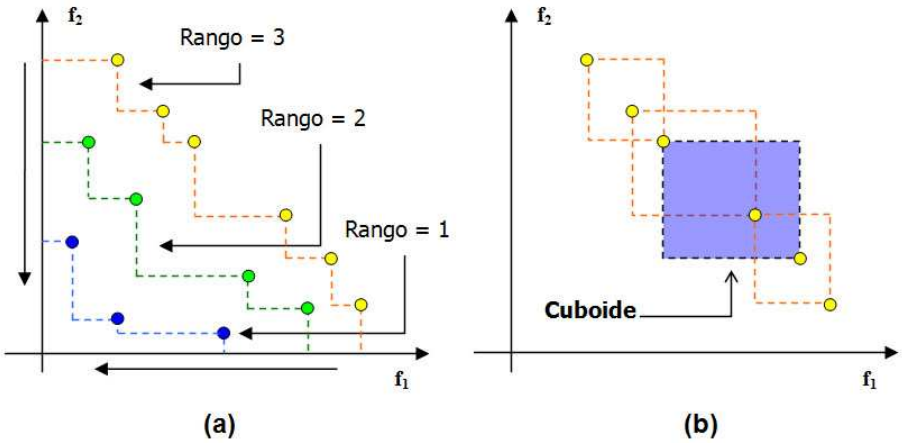


Figure B.6: Evaluación de la calidad en NSGA-II: (a) Jerarquización por fronteras de dominación y (b) control de la densidad basado en cuboides, caso minimización (tomado de [181, pg. 30][174, pg. 1061]).

Algoritmo B.3 Nondominated Sorting GA II: **NSGAII**$(F_a(I), M, N, t^{max})$

/* Inicialización */
1: `Initialize`$(P_A^{(0)})$ /* Inicializar la población aleatoriamente*/
2: $P^{(0)} := \emptyset$
3: $t \leftarrow 0$
/* Exploración */
4: **while** $(t < t^{max})$ **do**
5: $P^{(t)} := P^{(t)} + P_A^{(t)}$
6: `Rank`$(P^{(t)}, F_a(I))$ /* Aplicar el operador Crowding*/
7: $P_A^{(t+1)} :=$ `Update`$(P^{(t)}, F_a(I), N)$ /* Almacena los n mejores individuos */
8: $M_p :=$ `Select`$(P_A^{(t+1)}, F_a(I), M)$ /* Selecciona m individuos para recombinar */
9: $P^{(t+1)} :=$ `Generate`(M_p, M) /* Genera la nueva muestra (descendencia) */
10: $t \leftarrow t + 1$
11: **end while**
/* Finalización */
12: `Output`$(P_A^{(t)})$

Figure B.7: Algoritmo B.3: Nondominated Sorting Genetic Algorithm II (NSGA-II).

Algoritmo B.4 MOSA$(F_a(I), \alpha, T^{(0)}, T^{max}, t^{step}, t^{max})$

/* Inicialización */
$I_{ref}^{(0)} \leftarrow$`Initialize`$(I^{(0)})$ /* Inicializa al individuo de referencia aleatoriamente */
$P_A^{(0)} := I^{(0)}$
$t \leftarrow 0$
/* Exploración */
while $(t^{count} < t^{max}$ AND $T^{(t)} \geq T^{max})$ **do**
  $I^{(t)} \leftarrow$`GenerateNeighbour`$(I_{ref}^{(t)})$ /* Obtiene aleatoriamente un vecindad de $I_{ref}^{(t)}$ */
  If `Decide`$(I_{ref}^{(t)}, I^{(t)})$ then
    $I_{ref}^{(t)} \leftarrow I^{(t)}$
    $P_A^{(t+1)} :=$ `Archive`$(P_A^{(t)}, I^{(t)})$ /* Actualiza el archivo con $I^{(t)}$ */
    If $P_A^{(t+1)} \neq P_A^{(t)}$ then $t^{count} \leftarrow 0$
  Else
    $I_{ref}^{(t+1)} \leftarrow I_{ref}^{(t)}$
    $t^{count} \leftarrow t^{count} + 1$
  If $(t$ `mod` $t^{step})$=0 then $T^{(n+1)} = \alpha * T^{(n)}$ else $T^{(n+1)} = T^{(n)}$
  $t \leftarrow t + 1$
end while
/* Finalización */
`Output`$(P_A^{(t)})$

Decide$(I_{ref}^{(t)}, I^{(t)})$

Si $\Delta_S \leq 0$
  retornar `true`
Si no
  con probabilidad p retornar `true`
Si no retornar `false`

Donde:
$p = exp\left(-\frac{\Delta_S}{T^{(n)}}\right)$
$\Delta_S = \sum_{i=0}^{k} \lambda_i \left(f_i(I_{ref}^{(t)}) - f_i(I^{(t)})\right)$

Figure B.8: Algoritmo B.4: Recocido simulado multiobjetivo (MOSA).

Métricas: S y ϵ-indicador

La métrica S o hipervolumen es una medida de calidad para frentes de aproximación, definida como la métrica Lebesgue del politope definido por un punto de referencia y la frontera de aproximación.

El ϵ-indicador es una métrica para comparar pares de fronteras de aproximación, definida como:

$$I_\epsilon(A, B) = \inf_{\epsilon \in \mathbb{R}} \{\forall \mathbf{x}^2 \in B \exists \mathbf{x}^1 \in A : \mathbf{x}^1 \succ_\epsilon \mathbf{x}^2\} \tag{B.48}$$

para aproximaciones A y B, o en términos de un conjunto de referencia R como $I_\epsilon(A, R)$ [241].

B.3.3 Manejo de incertidumbre en EMO

A pesar de la gran importancia del manejo de incertidumbre en toma de decisiones, los esfuerzos realizados en EMO para incorporar el manejo de incertidumbre dentro del funcionamiento de los MOEA sigue siendo escaso. En general, se distinguen en EA cuatro categorías de problemas con incertidumbre [93], a saber:

Ruido: la función objetivo está afectada por ruido, que puede ser modelado como $F(\mathbf{x}) + \delta$, donde el primer sumando es la parte invariable, mientras que el segundo cambia cada vez que la función es evaluada con el mismo argumento.

Robustez: corresponde al caso de argumentos afectados por la incertidumbre, de modo que el valor real del vector nominal $\mathbf{x}$ puede variar. Este tipo de incertidumbre es modelado como $F(\mathbf{x} + \delta)$.

Funciones de aproximación: se refiere a los casos donde, debido a la dificultad de evaluar la función objetivo $F(\mathbf{x})$, se emplea una función de aproximación de F, introduciendo un error en la evaluación que se hace de la misma.

Funciones dinámicas: esta categoría encierra las funciones del tipo $F(\mathbf{x}, t)$, donde t se refiere al tiempo. En este caso la imagen de un argumento cambia con el tiempo, como es el caso de las acciones bursátiles por ejemplo.

Obsérvese que las categorías anteriores no diferencian entre incertidumbre epistémica y aleatoriedad. Por otro lado, los problemas reales no se encasillan en una única categoría, sino que son mixtos. De allí que se requieran más esfuerzos teórico-prácticos en EMO para incorporar el manejo de incertidumbre en los algoritmos.

La tabla B.5 presenta los principales esfuerzos en diseño o adaptación de MOEA para manejar incertidumbre. Las categorías mencionadas se revisan a continuación.

Procedimientos basados en dominancia probabilística

A esta categoría corresponden los primeros intentos de incorporar manejo de incertidumbre en MOEA, publicados en 2001 por E. Hughes [89, 88] y J. Teich [197]. Posteriormente J.E. Fieldsen & R.M. Everson [51] retomaron la idea en 2005.

El enfoque de E. Hughes [89, 88] se centra en funciones con ruido, y supone que cada vector de objetivos tiene una PDF asociada. Bajo la simplificación de que en la mayoría de los problemas de ingeniería el ruido sigue una distribución normal, el autor desarrolló fórmulas para calcular la probabilidad de cometer un error cuando se considera que un vector domina a otro, a partir de lo cual se puede construir un MOEA. Por su parte, J. Teich [197] considera vectores de objetivos contenidos en intervalos. Tras suponer una distribución uniforme dentro del intervalo, es posible calcular la probabilidad de que una alternativa domine a otra que la solapa en el espacio de atributos. Luego la decisión de aceptar la dominancia dependerá del umbral de aceptación que imponga la DM. Finalmente, J.E. Fieldsen & R.M. Everson [51] expanden los estudios de E.

Tema de estudio	Autores y trabajos
Dominancia probabilística:	
Optimización con valores de adaptación en intervalos	J. Teich [197]
Optimización de funciones objetivo con ruido	E. Hughes [89, 88], J.E. Fieldsen & R.M. Everson [51]
Procedimientos basados en indicadores de calidad:	
Optimización basada en un indicador	M. Basseur & E. Zitzler [11]
Optimización robusta:	
Optimización con propagación de incertidumbre	K. Sörensen [194], T. Ray [152], K. Deb & H. Gupta [34], C. Barrico and C.H. Antunes [10]
Diseño robusto basado en Info-gap	D. Lim et al. [124]
Optimización basada en fiabilidad	K. Deb et al. [35]
Limitación de probabilidades:	
Intervalo de tolerancia	S. Martorell et al. [136], J.F. Villanueva et al. [204]
Procedimientos no probabilísticos:	
Optimización con incertidumbre epistémica	P. Limbourg [125], P. Limbourg & D.E. Salazar A. [126, 173]

Table B.5: Estado del arte en diseño de MOEA para optimización con incertidumbre.

Hughes considerando varianza desconocida y la estimación de parámetros con ayuda de procedimientos Monte Carlo.

Procedimientos basados en indicadores de calidad

M. Basseur & E. Zitzler poponen en [11] un algoritmo basado en indicador que maneja la incertidumbre usando el valor esperado del indicador I_ϵ, suponiendo que cada solución tiene asociada una PDF. Así, tras estimar el promedio del indicador por simulación respecto al resto de la población, se puede determinar las alternativas a aceptar.

Procedimientos de optimización robusta

Los algoritmos de optimización robusta en EMO están inspirados en la propuesta de S. Tsuitsui y A. Ghosh [200], que consiste en suponer que la incertidumbre asociada a $\mathbf{x}$ es aleatoria y posee una PDF. De allí, la incertidumbre es propagada vía Monte Carlo, estimando el valor esperado como $\sum_i F(\mathbf{x} + \delta_i)$. Este enfoque es llamado *función efectiva* en [200] y es la base de las contribuciones de T. Ray [152], K. Sörensen [194, 193] y K. Deb & H. Gupta [34]. La propuesta de los últimos autores se diferencia en que restringe el tamaño de las desviaciones del valor esperado que se puede aceptar.

Un enfoque de propagación que también toma en cuenta el tamaño del dominio propagado fue propuesto por C. Barrico and C.H. Antunes [10]. En esta

propuesta el *grado de robustez* es el máximo coeficiente entero por el cual se puede multiplicar al dominio original, sin incurrir en una desviación superior a un valor especificado en el espacio de objetivos.

Limitación de probabilidades

Recientemente S. Martorell et al. [136] y J.F. Villanueva et al. [204] han propuesto el uso de intervalos de tolerancia para estimar el límite superior (correspondiente al peor caso) de un intervalo que limita o encierra una masa de probabilidad especificada con cierto nivel de confianza.

Procedimientos no probabilísticos

La desventaja principal de los procedimientos probabilísticos es que su aplicabilidad depende de la existencia y conocimiento de las PDF asociadas a las soluciones. En consecuencia, los casos de incertidumbre epistémica y ciertos casos de aleatoriedad no pueden ser tratados adecuadamente con dichos enfoques. Como respuesta, P. Limbourg en [125] y P. Limbourg & D.E. Salazar A. en [126, 173] han propuesto MOEA no probabilísticos para el tratamiento de la incertidumbre epistémica. El último enfoque se expone más adelante como una de las contribuciones de esta tesis.

Métricas y medidas de calidad bajo incertidumbre

Un problema abierto en EMO es la calidad de las fronteras de aproximación bajo incertidumbre. J.E. Fieldsen & R.M. Everson [51] representan las fronteras usando valores esperados y luego determinan la probabilidad de que una frontera domine a algún valor determinado tomando muestras. M. Basseur & E. Zitzler [11] propusieron clasificar la calidad de diversas fronteras (A) calculando el peor valor esperado estadísticamente significativo, respecto a una frontera de referencia, usando $E(I_\epsilon(A, R))$.

Más adelante en este trabajo, se propone el uso de la métrica S para fronteras con incertidumbre [126, 173].

B.4 AUREO: Contribuciones al análisis de incertidumbre y robustez en EMO

En esta sección se resumen los contenidos y contribuciones expuestos en el capítulo 4. Los temas son expuestos en cuatro bloques, siendo el primero el análisis del efecto de la incertidumbre en la toma de decisiones y particularmente en los MOEA. Seguidamente se presentan una serie de tres propuestas relacionadas con el diseño de MOEA para manejar incertidumbre. En tercer lugar se hace una revisión del concepto de robustez y se propone una taxonomía del mismo en relación al tipo de problemas multiobjetivo y los MOEA empleados para resolverlos. Finalmente se recopilan y expanden las ideas en una propuesta metodológica para el *Análisis de incertidumbre y robustez en algoritmos evolutivos*, abreviada AUREO.

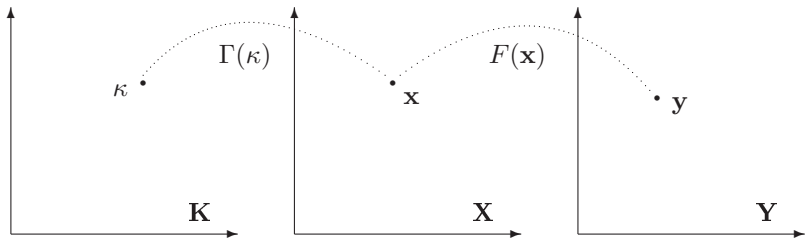


Figure B.9: Diferentes espacios del proceso de decisión.

B.4.1 Análisis de efectos de la incertidumbre en MOEA

Considérese un espacio $\mathbf{K}$ (*kappa*) que denota al cosmos ($\mathit{Κοσμος}$) y por tanto cubre todos los estados que una variable de decisión puede tomar. El espacio de decisión $\mathbf{X}$ de cualquier variable de decisión es por tanto un subespacio de $\mathbf{K}$. Sea Γ (de $\Gamma\mathit{νωσις}$) una función que relaciona instancias particulares de la realidad κ desde $\mathbf{K}$ en $\mathbf{X}$. Luego, el proceso de decisión comprende 3 espacios, como se muestra en la figura B.9, a saber: la realidad $\mathbf{K}$, el espacio de decisión $\mathbf{X}$ y el espacio de atributos $\mathbf{Y}$.

En un estado de información perfecta, Γ sería una función identidad, permitiendo que cada instancia de la realidad κ pueda ser representada por un único vector de decisión $\mathbf{X}$. Del mismo modo, $F(\mathbf{x})$ sería una función libre de incertidumbre, permitiendo que $\mathbf{X}$ sea relacionado perfectamente con $\mathbf{Y}$ y viceversa. No obstante, en la realidad no es posible una relación biyectiva de Γ, debido a nuestras limitaciones físicas y temporales para recopilar y representar la información, junto con nuestra incapacidad de diferenciar más allá de ciertos umbrales. Por tanto, $\mathbf{X}$ se convierte en un espacio de *vectores nominales*, tal que $\mathbf{x}$ es una representación nominal de un subconjunto de $\mathbf{K}$, según se muestra en la figura B.10, ya sea porque la DM no distingue entre los distintos estados de $\mathbf{K}$ representados en la figura, o porque la variabilidad inherente al sistema de interés hace imposible que ciertas variables o parámetros puedan ser controlados, siendo posible sin embargo, etiquetar a tal conjunto de estados con un vector nominal $\mathbf{x}$. Como consecuencia, en la práctica la incertidumbre debe ser propagada desde el espacio $\mathbf{K}$ a $\mathbf{X}$ y subsecuentemente a $\mathbf{Y}$ (figura B.11).

La incertidumbre epistémica surge como resultado de las imperfecciones en las operaciones sobre la información (tabla B.2), en particular la recolección de datos. Así, cualquier estado particular de la realidad κ es identificado como un

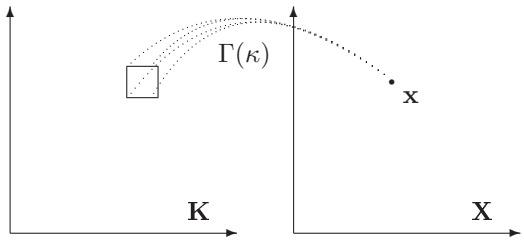


Figure B.10: Aleatoriedad, preferencias y espacio de decisión.

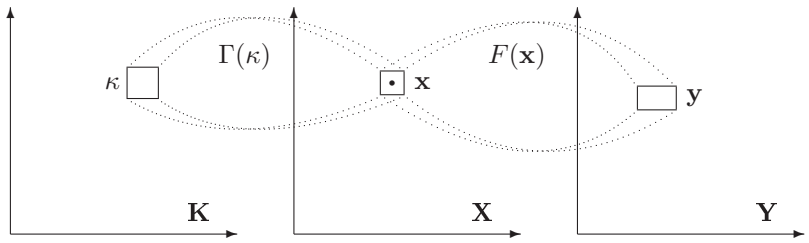


Figure B.11: Propagación de incertidumbre aleatoria desde la realidad al espacio de decisión y luego al espacio de objetivos. Nótese el subconjunto en $\mathbf{X}$ asociado al punto nominal $\mathbf{x}$.

subconjunto en $\mathbf{X}$ en lugar de un valor preciso. En consecuencia, es posible que los subconjuntos asociados a κ cambien con cada medición (generando $\mathcal{X}_1$ y $\mathcal{X}_2$ en la figura), y en todo caso al propagar la incertidumbre, un estado específico de $\mathbf{K}$ sea representado por un conjunto de vectores de objetivos en $\mathbf{Y}$ (figura B.12).

Tanto la incertidumbre epistémica como la aleatoriedad tiene como resultado la generación de subconjuntos en $\mathbf{Y}$, imponiendo una dificultad evidente sobre la toma de decisiones, que en consecuencia ya no se realizará con base en las comparaciones de puntos o vectores precisos, sino en la comparación de conjuntos, y que en algunos casos un mismo estado de la realidad puede estar asociado a distintos vectores de atributos ($\mathcal{Y}_1$ y $\mathcal{Y}_2$ en la figura B.12).

Criterios para comparar conjuntos

La dos cuestiones fundamentales que deben considerarse para comparar conjuntos usando MOEA son (1) cómo determinar la dominancia y (2) cómo estimar la densidad. A tal fin hemos de considerar dos casos, la comparación de intervalos sin información adicional y la comparación con información adicional.

Comparación de conjuntos sin información adicional

Cuando no se tiene más información que la definición de conjuntos bajo la forma de intervalos de la forma $\mathcal{A} = [\underline{a}, \overline{a}]$ y $\mathcal{B} = [\underline{b}, \overline{b}]$, la dominancia puede decidirse

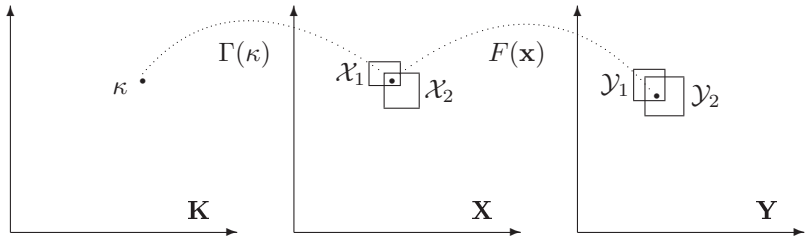


Figure B.12: Efecto de la incertidumbre epistémica en toma de decisiones: una instancia particular κ es asociada a $\mathcal{X}_1$ y $\mathcal{X}_2$ y evaluada en el espacio de objetivos como $\mathcal{Y}_1$ y $\mathcal{Y}_2$ respectivamente.

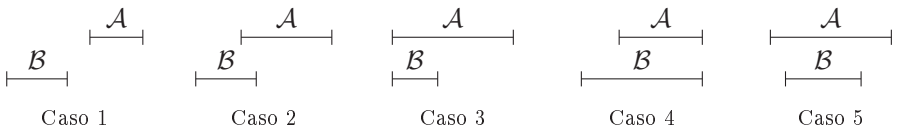


Figure B.13: Cinco posibles casos de comparación de intervalos en jerarquización de soluciones bajo incertidumbre usando MOEA.

a través de uno de los siguientes criterios:

Criterio de la razón insuficiente: también llamado principio de Laplace; establece que ante la falta de evidencias, todas los posibles valores que puede tomar una variable han de considerarse igualmente probables. En la práctica se reduce a suponer una distribución uniforme dentro del intervalo de estudio. Se considera, sin embargo, que ante incertidumbre epistémica este criterio debe ser asumido con cuidado [47, 13]) pues puede conducir a infravalorar riesgos.

Optimización del peor caso: como lo indica su nombre, consiste en fundamentar las valoraciones sobre el peor escenario posible, esto es el límite inferior del intervalo en maximización, y el superior en minimización.

Optimización del mejor caso: es el complemento del anterior y trabaja comparando el mejor escenario.

Criterio de Hurwicz: combina los criterios anteriores de forma convexa, haciendo $\alpha \cdot \overline{x} + (1-\alpha) \cdot \underline{x}$ para $\alpha \in [0,1]$, recayendo sobre la DM la decisión de fijar α.

Mínimo arrepentimiento: también llamado criterio de Savage; establece que ante diversos escenarios, se debe minimizar la mayor pérdida (que conduce a un arrepentimiento) debida a una decisión errada.

La aplicación de los criterios anteriores en comparaciones monocriterio se ejemplifican en la figura B.13. Obsérvese que el caso 1 no implica ningún riesgo de una decisión errada, mientras que en el resto de los casos está posibilidad está latente. La tabla B.6 resumen las conclusiones posibles a las que puede llegarse usando dichos criterios en MOEA para jerarquizar soluciones.

La aplicación de los criterios anteriores para problemas monocriterio puede implementarse en MOEA a través de la siguiente definición de *dominancia-SOI* (def. 8): *Un vector unidimensional* $\mathcal{A} = [\underline{a}, \overline{a}]$ *domina a* $\mathcal{B} = [\underline{b}, \overline{b}]$*, denotado* $\mathcal{A} \succ_{SOI} \mathcal{B}$*, si y solo si (caso maximización):*

$$\mathcal{A} \succ_{SOI} \mathcal{B} \Leftrightarrow \begin{cases} \overline{a} > \overline{b} & \textit{Mejor caso} \\ \underline{a} > \underline{b} & \textit{Peor caso} \\ (\overline{a} + \underline{a}) > (\overline{b} + \underline{b}) & \textit{Laplace} \\ \alpha \cdot (\overline{a} - \overline{b}) > (1-\alpha) \cdot (\underline{b} - \underline{a}),\ \alpha \in [0,1] & \textit{Hurwicz} \\ \overline{a} - \underline{b} > \overline{b} + \underline{a} & \textit{Mínimo arrepentimiento} \end{cases}$$

Si no se verifica $\mathcal{A} \succ_{SOI} \mathcal{B}$ *ni* $\mathcal{B} \succ_{SOI} \mathcal{A}$ *entonces los intervalos son incomparables:* $\mathcal{A} || \mathcal{B}$.

Casos (Fig. B.13)	Mejor caso	Peor caso	Razón Insuficiente	Criterio de Hurwicz	Arrepentimiento mínimo
Caso 1	$\mathcal{A} \succ \mathcal{B}$	$\mathcal{A} \succ \mathcal{B}$	$\mathcal{A} \succ \mathcal{B}$	$\mathcal{A} \succ \mathcal{B}$	$\mathcal{A} \succ \mathcal{B}$
Caso 2	$\mathcal{A} \succ \mathcal{B}$	$\mathcal{A} \succ \mathcal{B}$	$\mathcal{A} \succ \mathcal{B}$	$\mathcal{A} \succ \mathcal{B}$	$\mathcal{A} \succ \mathcal{B}$
Caso 3	$\mathcal{A} \succ \mathcal{B}$	$\mathcal{A}\|\|\mathcal{B}$	$\mathcal{A} \succ \mathcal{B}$	$\mathcal{A} \succ \mathcal{B}$	$\mathcal{A} \succ \mathcal{B}$
Caso 4	$\mathcal{A}\|\|\mathcal{B}$	$\mathcal{A} \succ \mathcal{B}$	$\mathcal{A} \succ \mathcal{B}$	$\mathcal{A} \succ \mathcal{B}$	$\mathcal{A} \succ \mathcal{B}$
Caso 5	$\mathcal{A} \succ \mathcal{B}$	$\mathcal{A} \succ \mathcal{B}$	$\mathcal{A}\{\succ, \|\|, \prec\}\mathcal{B}$	$\mathcal{A}\{\succ, \|\|, \prec\}\mathcal{B}$	$\mathcal{A}\{\succ, \|\|, \prec\}\mathcal{B}$

Table B.6: Conclusiones de la aplicación de cinco criterios no probabilísticos en jerarquización bajo incertidumbre usando MOEA (caso maximización): Dependiendo de la relación entre los intervalos (Mínimo arrepentimiento, Laplace) o del α (Hurwicz), se pueden obtener distintas soluciones ($\{\succ, ||, \prec\}$).

Ahora bien, la expansión de los criterios anteriores al caso multiobjetivo, permite un rango más amplio de definiciones de dominancia. Para explicarlo, considérense los distintos casos mostrados en la fig. B.14; para determinar dominancia es necesario establecer si se ha de verificar la dominancia SOI para cada objetivo, o si es posible aceptar criterios diferentes. Dado que en dominancia de Pareto se exige no ser peor en todos los objetivos y estrictamente mejor en un objetivo, en el caso de varios intervalos se puede hacer un requerimiento similar, como en [126, 173], que conduce a la dominancia MOI (def. 9): *Un vector multiobjetivo de intervalos* $\mathcal{A} = \{\mathcal{A}_i = [\underline{a}_i, \overline{a}_i] : i \in 1, \ldots, k\}$ *domina a* $\mathcal{B} = \{\mathcal{B}_i = [\underline{b}_i, \overline{b}_i] : i \in 1, \ldots, k\}$*, denotado* $\mathcal{A} \succ_{MOI} \mathcal{B}$*, si y solo si (caso maximización):*

$$\mathcal{A} \succ_{MOI} \mathcal{B} \Leftrightarrow \begin{cases} \mathcal{A}_i \succ_{SOI} \mathcal{B}_i \vee \mathcal{A}_i || \mathcal{B}_i : \forall i \in \{1, \ldots, k\} \\ \exists j \in \{1, \ldots, k\} : \mathcal{A}_j \succ_{SOI} \mathcal{B}_j \end{cases}$$

Si no se verifica $\mathcal{A} \succ_{MOI} \mathcal{B}$ *ni* $\mathcal{B} \succ_{MOI} \mathcal{A}$ *se dice que ambos vectores son incomparables:* $\mathcal{A}||\mathcal{B}$.

Nótese que la aplicación de los primeros 4 criterios expuestos anteriormente es directa con la dominancia MOI, mientras que el control de densidad puede realizarse con los métodos habituales de los MOEA, como el *hypergrid* o el operador *crowding*, puesto que los intervalos son representados por valores representativos.

La implementación del criterio de mínimo arrepentimiento es posible bien sea usando la definición 9 directamente, o ponderando el arrepentimiento mediante una métrica L_p de la forma $\left(\sum_i^k w_i |x_i|^p\right)^{\frac{1}{p}}$ donde x corresponde al valor máximo de la perdida debida a una mala decisión y los $w_i \geq 0$ son los pesos dados a cada objetivo. El uso de dicha métrica requiere que el analista o DM otorguen al exponente p y a los pesos w_i una semántica adecuada en el contexto del problema estudiado. Un caso particular es el uso de la norma $L_{\inf}$, cuyo planteamiento exhibe semejanzas interesantes con el ϵ-indicador aditivo $I_{\epsilon+}$(A,B) para dos intervalos A y B.

Comparación de conjuntos con información adicional

Cuando se dispone de información adicional sobre los intervalos, tal que para los intervalos $\mathcal{A} = F(\mathbf{x}_1)$ y $\mathcal{B} = F(\mathbf{x}_2)$ algunos valores dentro de dichos conjuntos son más probables o posibles que otros, los procedimientos de jerarquización de

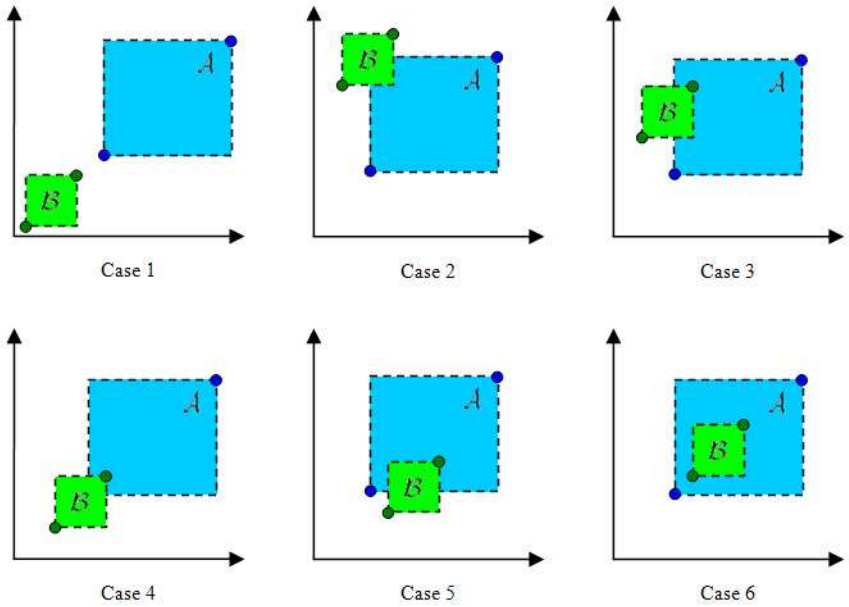


Figure B.14: Ejemplos de posibles casos de comparación de intervalos bidimensionales en jerarquización de soluciones usando MOEA.

soluciones en los MOEA pueden construirse, utilizando las siguientes definiciones según sea la caracterización de la incertidumbre:

Jerarquización con teoría clásica de probabilidad: Si se conocen las CDF de los vectores de objetivos y las funciones objetivo cumplen con ciertas condiciones, la dominancia puede verificarse con los siguiente conceptos de dominancia estocástica:

- *Dominancia estocástica de primer grado* [65, 80] (def. 10): para funciones objetivo $\mathbf{y} = F(\mathbf{x})$ crecientes ($F'(\mathbf{x}) > 0$), $\mathbf{x}_1$ *domina estocásticamente en primer grado* a $\mathbf{x}_2$ si $\mathtt{F}(\mathbf{y} = F(\mathbf{x}_1)) \leq \mathtt{F}(\mathbf{y} = F(\mathbf{x}_2))$, donde $\mathtt{F}(\mathbf{y} = F(\mathbf{x}))$ es la CDF de $\mathbf{y}$ y la inecuación es estricta para al menos un valor de $\mathbf{y}$. Si el test no es conclusivo, entonces se puede verificar la
- *Dominancia estocástica de segundo grado* [65, 80] (def. 11): para las condiciones anteriores y $F''(\mathbf{x}) < 0$, $\mathbf{x}_1$ *domina estocásticamente en segundo grado* a $\mathbf{x}_2$ si $\int_{-\infty}^{z} \mathtt{F}(\mathbf{y} = F(\mathbf{x}_1))d\mathbf{y} \leq \int_{-\infty}^{z} \mathtt{F}(\mathbf{y} = F(\mathbf{x}_2))d\mathbf{y}$ para todo z en $\mathbf{y}$ donde la desigualdad es estricta para al menos un valor de $\mathbf{y}$.

Si no se cumplen las condiciones ($F'(\mathbf{x}) > 0$) y ($F''(\mathbf{x}) < 0$), se puede emplear otros conceptos de dominancia probabilística como la *Semidominancia estocástica* o el *Análisis Media-Gini* descritos en [65]. No obstante, en general es complicado determinar la CDF y cumplir con el resto de condiciones de dichas definiciones, además del esfuerzo computacional que

entrañan. El uso de simulaciones y comparación de momentos es por tanto comprensible y necesario en muchos casos, pero cuando sea posible recurrir a los conceptos de dominancia, sería una buena práctica considerarlos junto con controles de densidad basados, por ejemplo, en algún momento de primer orden.

Jerarquización con lógica difusa: Aunque la lógica difusa ha sido empleada en EA y EMO para manejar la información de preferencias como 'preferible', 'aceptable', etc. cuando se comparan vectores solución sin incertidumbre, el estudio de vectores objetivo difusos no parece haber sido abordado hasta el momento en EMO.

Considérense la comparación de los conjuntos difusos $\mathcal{A} = F(\mathbf{x}_1)$ y $\mathcal{B} = F(\mathbf{x}_2)$, tal que los soportes $\mathcal{A}_0 = [\underline{\mathcal{A}_\alpha}, \overline{\mathcal{A}_\alpha}]$ y $\mathcal{B}_0 = [\underline{\mathcal{B}_\alpha}, \overline{\mathcal{B}_\alpha}]$ se solapan. La dominancia puede ser decidida a través de varios criterios (ver [77, 57] y sus referencias), entre los que destaca por su simplicidad la definición de dominancia de P. Fortemps y M. Roubens [57], que para el caso de maximización, establece que $\mathcal{A} \succ \mathcal{B}$ si y solo si $C(\mathcal{A} \geq \mathcal{B}) > 0$ y $\mathcal{A} \succeq \mathcal{B}$ si y solo si $C(\mathcal{A} \geq \mathcal{B}) \geq 0$, donde $C(\mathcal{A} \geq \mathcal{B}) = \mathcal{F}(\mathcal{A}) - \mathcal{F}(\mathcal{B})$ y $\mathcal{F}(\mathcal{A}) = \frac{1}{2}\int_0^1 (\underline{\mathcal{A}_\alpha} + \overline{\mathcal{A}_\alpha})d\alpha$. De forma similar, R. Groetschel and W. Voxman definen [25] (para maximización) el criterio $\mathcal{A} \succeq \mathcal{B} \Leftrightarrow \int_0^1 \alpha(\underline{\mathcal{A}_\alpha} + \overline{\mathcal{A}_\alpha})d\alpha \geq \int_0^1 \alpha(\underline{\mathcal{B}_\alpha} + \overline{\mathcal{B}_\alpha})d\alpha$, que se expresa preferencia estricta haciendo estricta la desigualdad.

Los criterios anteriores son sencillos de calcular y pueden integrarse en los MOEA haciendo uso de la dominancia MOI en la jerarquización, donde las relaciones anteriores equivaldrían a la dominancia SOI. Por otro lado, los diseñadores de MOEA pueden hacer uso de ciertos puntos representativos de los conjuntos difusos para el control de densidad. Así, la media posibilística de C. Carlsson and R. Fullér [25] (def. 14) de dos números difusos es definida como $\bar{M}(\mathcal{A}) = \frac{M_*(\mathcal{A})+M^*(\mathcal{A})}{2}$, donde $M_*(\mathcal{A}) = 2\int_0^1 \alpha\underline{\mathcal{A}_\alpha}d\alpha$ y $M^*(\mathcal{A}) = 2\int_0^1 \alpha\overline{\mathcal{A}_\alpha}d\alpha$. Esta definición promedia las medias de los α-cortes, ponderándolos con el valor de α. Nótese que no solo el control de densidad sino todo el MOEA puede emplear esta definición para jerarquizar y almacenar soluciones.

Jerarquización con teoría de probabilidades imprecisas: Cuando la incertidumbre está representada por probabilidades imprecisa, las PDF y CDF poseen límites inferiores y superiores que no permiten calcular valores representativos precisos, sino intervalos para dichos valores. En consecuencia, no es posible adaptar directamente un problema que tenga esta estructura a la forma actual de los MOEA.

Hasta donde se tiene noticia, no existe en la literatura ningún estudio práctico o teórico del uso de probabilidades imprecisas en EMO, y más aun, ni siquiera ha sido planteada la posibilidad de hacerlo.

Dado que la teoría de probabilidades imprecisas contempla la existencia de reglas para toma de decisiones monocriterio, es posible pensar en emplearlas como dominancia SOI, y construir luego una regla de dominancia MOI. En [209, 199, 127] se abordan los criterios de dominancia usando probabilidades imprecisas, mientras que [23, 22] reporta aplicaciones en optimización monocriterio de caja negra usando p-boxes. Otros enfoques

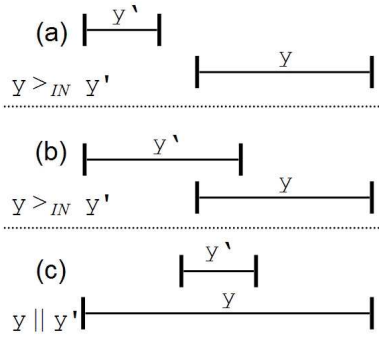


Figure B.15: Ejemplo de comparación de intervalos (maximización) [126, 173].

generalizan la idea de valor esperado para probabilidades imprecisas [224]. El control de densidad, sin embargo, requeriría un trabajo adicional para extender el razonamiento probabilístico a las reglas de almacenamiento.

B.4.2 Propuestas para incorporar manejo de incertidumbre en MOEA

En esta sección se presentan tres propuestas relativas a la incorporación del manejo de incertidumbre en el diseño de MOEA, fundamentadas en los temas discutidos y analizados hasta el momento. La primera propuesta consiste en un MOEA que opera con intervalos en el espacio de objetivos sobre premisas no probabilíticas. La segunda y tercera propuestas está relacionadas con el uso del criterio de arrepentimiento mínimo y en el problema del control de densidad de vectores objetivo inciertos.

Algoritmo IP-MOEA

En secciones anteriores se han analizado distintas políticas aplicables a la comparación de intervalos sin información adicional, para construir MOEA. En este apartado presentamos un enfoque llamado IP-MOEA, propuesto poe P. Limbourg & el autor en [126, 173]. Algunos tópicos son similares a los ya expuestos anteriormente, pero se han respetado para reflejar el razonamiento de los autores en su momento.

La figura B.15 muestra los tres posibles casos de comparaciones de intervalos unidimensionales junto con las conclusiones correspondientes en términos del operador $>_{IN}$, para el cual se establece que (def. 15) *un intervalo* $y = [\underline{y}, \overline{y}]$ *domina a* $y' = [\underline{y'}, \overline{y'}]$, *denotado* $y >_{IN} y'$, *si y solo si* $\underline{y} \geq \underline{y'} \wedge \overline{y} \geq \overline{y'} \wedge y \neq y'$. Se desprende entonces que el caso (a) es un caso de dominancia segura, luego $y >_{IN} y'$. El caso (b) por su parte es mucho más crítico. Se puede suponer que en algunos casos existan PDF asociadas a los intervalos, pero las mismas pueden variar entre y y y'. Por otro lado, la única información que se asume como cierta es que el valor verdadero de cada vector de objetivos yace dentro de los intervalos correspondiente. Si bien es posible que y' sea mayor que y, se acepta que y es preferible a y' pues implica una mayor posibilidad

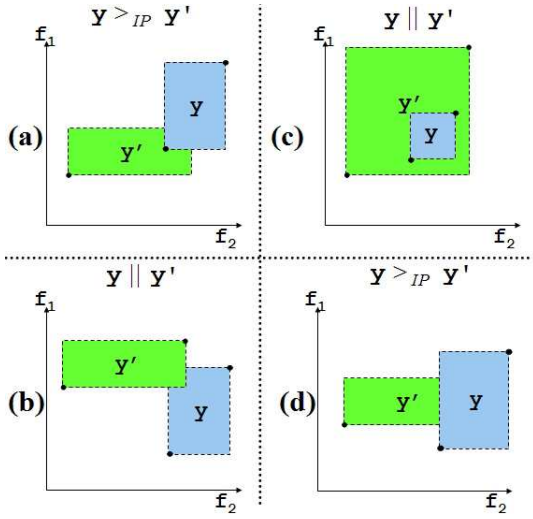


Figure B.16: Ejemplos de comparación de intervalos multiobjetivo (caso maximización) [126, 173].

de poseer un valor verdadero mejor (caso maximización). La incomparabilidad $y||y'$ también es posible, pero generaría un gran número de conclusiones neutras; además tanto el criterio de peor caso como el de mejor caso coinciden en que y domina a y'. Finalmente, el caso (c) es claramente un caso de incomparabilidad en ausencia de información adicional. La extensión a múltiples objetivos del análisis anterior toma forma en la dominancia IP (def. 16) que establece que *un vector k-dimensional de objetivos inciertos* $y = \{y_i = [\underline{y}_i, \overline{y}_i] : i \in 1, \ldots, k\}$ *domina a* $y' = \{y'_i = [\underline{y'}_i, \overline{y'}_i] : i \in 1, \ldots, k\}$, *denotado* $y >_{IP} y'$, *si y solo si* $\forall i \in \{1, \ldots, k\} : y_i >_{IN} y'_i \vee y_i || y'_i$ *y* $\exists j \in \{1, \ldots, k\} : y_j >_{IN} y'_j$.

La figura B.16 muestra como funciona la dominancia IP: en el caso (a) tanto el peor como el mejor caso de y' están dominados por sus equivalentes en y, por eso es un caso claro de dominancia. Por el contrario, en el caso (b) ni los mejores ni los peores casos se dominan, lo que corresponde a un caso de incomparabilidad o indiferencia en la toma de decisiones. Por otra parte, el caso (c) es también de incomparabilidad pero por razones distintas, puesto que no hay una dominancia clara en ningún objetivo. Por último, esta situación se simplifica en el caso (d) donde al menos se cumple $y_i >_{IN} y'$ en una dimensión, permitiendo aceptar la dominancia, si bien para distintos DM podría corresponder a un caso de incomparabilidad.

El control de densidad se planteó usando la métrica S en un (indicador de hipervolumen I_H) como en [45], pero tomando en cuenta la contribución segura ($\underline{I_H}(y, P) = \underline{I_H}(P) - \underline{I_H}(P - y)$) y la plausible ($\overline{I_H}(y, P) = \overline{I_H}(P) - \overline{I_H}(P - y)$) para cada intervalo y respecto a toda la población P. Luego, se adaptó el NSGA-II según los siguientes criterios: (1) comparar los individuos usando $>_{IP}$, si no se puede decidir entonces (2) comparar la contribución en hipervolumen $[\underline{I_H}(y, P), \overline{I_H}(y, P)]$ usando el criterio $>_{IN}$; si no se puede decidir, entonces (3) escoger aleatoriamente.

Para evaluar el algoritmo se comparó con el NSGA-II operando con los centroides de los intervalos. La calidad de los resultados se midió con la métrica S,

Función	IP-MOEA		NSGA-II		
ZDT_I1	*Media*	*RIC*	*Media*	*RIC*	$p(0)$
Peor caso	0.7139	1.15E-02	0.7064	1.57E-02	9.60E-07
Mejor caso	0.8991	1.59E-02	0.8899	1.97E-02	3.23E-05
ZDT_I2	*Media*	*RIC*	*Media*	*RIC*	$p(0)$
Peor caso	0.6528	1.91E-02	0.6324	2.32E-02	7.25E-12
Mejor caso	0.8219	2.50E-02	0.7954	2.74E-02	1.04E-11
ZDT_I4	*Media*	*RIC*	*Media*	*RIC*	$p(0)$
Peor caso	0.7311	3.96E-02	0.7017	4.45E-02	1.04E-09
Mejor caso	0.9062	4.37E-02	0.8728	5.58E-02	1.74E-09
ZDT_I6	*Media*	*RIC*	*Media*	*RIC*	$p(0)$
Peor caso	0.4759	3.87E-02	0.4531	3.99E-02	7.12E-07
Mejor caso	0.7591	5.26E-02	0.7248	4.88E-02	2.92E-06

Table B.7: Hipervolumen luego de 100 generaciones para los valores inciertos de las funciones test ZDT_I [126, 173].

adaptada para el caso de valores inciertos (calculando $\underline{I_H}$) y $\overline{I_H}$), y usando cuatro funciones test bien conocidas (ZDT1, ZDT2, ZDT4, ZDT6 [236]) adaptadas para producir intervalos (ver tabla 4.4) de forma determinística.

Ambos algoritmos se evaluaron con 100 generaciones, tamaño de población 20, probabilidad de cruce 0.9 y de mutación 0.1. Se fijaron 30 variables de decisión para todas las funciones, y se compararon resultados calculando la métrica S para el peor y mejor caso del frente de aproximación. Se fijó como punto de referencia el menor valor posible de objetivos y se escalaron los resultados en [0,1]. La tabla B.7 reporta los resultados obtenidos, junto al rango intercuartil (RIC) y la probabilidad de la hipótesis nula, $p(0)$ acerca de la igualdad de medias. El test Kruskal-Wallis se empleó para comprobar la significación estadística de la igualdad de medias.

Los resultados señalan que el procedimiento propuesto mejora el comportamiento del algoritmo. Nótese que en todos los casos test tanto el en peor como el mejor caso $(\underline{I_H}(y,P), \overline{I_H}(y,P))$ el hipervolumen es mayor, lo que indica mejor convergencia. Además la hipótesis nula puede ser rechazada ($p(0) < 0.001$ en todas las pruebas). A partir de los resultados, los autores concluyen que el algoritmo IP-MOEA ha mostrado ser al menos tan bueno como el NSGA-II trabajando con valores esperados, por lo que el IP-MOEA luce como una propuesta interesante para enfrentar incertidumbre epistémica en forma de intervalos.

El ϵ-indicador visto desde la óptica del concepto de mínimo arrepentimiento

El concepto de ϵ-dominancia ha recibido una acogida positiva de parte de la comunidad que trabaja en EMO, debido a que su incorporación mejora notablemente las propiedades de convergencia de los MOEA [117, 116], y al mismo tiempo permite construir indicadores calidad para los frentes de aproximación. En particular, el ϵ-indicador ha sido utilizado por M. Basseur & E. Zitzler en [11], para construir un algoritmo que favorece a aquellos candidatos cuyo valor esperado del ϵ-indicador, calculado mediante simulación y suponiendo informa-

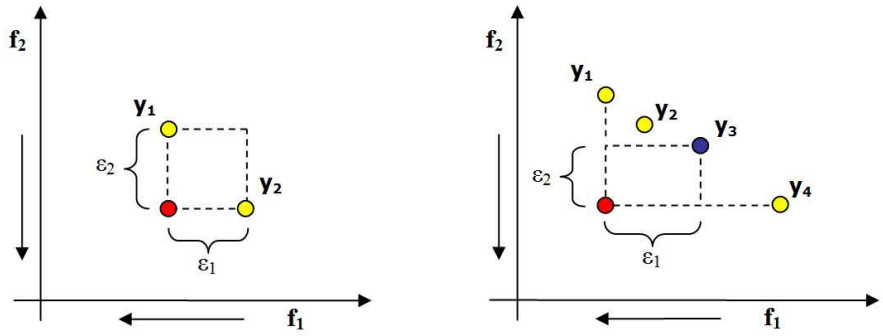


Figure B.17: Comparición de la ϵ-dominancia y el arrepentimiento mínimo (caso minimización): entre dos puntos (izquierda) y entre un punto y un conjunto (derecha).

ción probabilística, es mejor respecto al resto de la población (ver sec. B.3.3 y 3.4.2).

Sin embargo, una lectura no probabilística del ϵ-indicador para manejar incertidumbre es también posible, observando las coincidencias entre la ϵ-dominancia aditiva y el criterio de mínimo arrepentimiento usado toma de decisiones con información no probabilística. Considérense los casos mostrados en la figura B.17. El ϵ-indicador $I_\epsilon(A, B)$ '*es igual al mínimo factor ϵ tal que cada vector objetivo en B es ϵ-dominado por al menos un vector objetivo de A*' [241, pg. 122]. En consecuencia, si aplicamos el ϵ-indicador entre dos puntos (fig. 4.9 izquierda) para determinar cuál es mejor, como en [11], el indicador ha de favorecer aquel punto más cercano al punto ideal del conjunto (punto rojo), que es el que tiene el menor ϵ. Esta comparación es equivalente a la del monto de la pérdida en que se incurre cuando se toma una decisión errada, como se explicó en la sec. B.4.1 (ver también la sec. 4.1.3), de allí que la alternativa que produzca el mínimo arrepentimiento es también la de menor ϵ, indicada por el ϵ-indicador. Al extender el análisis para comparar un vector con un conjunto (fig. B.17 derecha), el efecto es el mismo. No obstante, cuando se comparan dos conjuntos, los resultados de emplear el ϵ-indicador y el criterio de mínimo arrepentimiento no son iguales, ya que el último equivale a comparar la distancia de los puntos ideales de cada conjunto, mientras que el ϵ-indicador no.

Si consideramos ahora vectores objetivo con incertidumbre, la pérdida continúa siendo una expresión de la distancia entre el peor y mejor caso, aunque en este caso dicha distancia sería multidimensional. Si consideramos calcularla con la norma L_∞, el arrepentimiento entre dos alternativas (ver fig. B.18 izquierda) y la aplicación del ϵ-indicador aditivo producen los mismos resultados. Del mismo modo, la pérdida de escoger un vector en lugar de un conjunto, es equivalente a la distancia máxima entre el peor caso de dicho vector y el punto ideal del conjunto (ver fig. B.18 derecha). Se desprende que la construcción de un MOEA basado en el ϵ-indicador aditivo, similar al algoritmo de M. Basseur & E. Zitzler [11] pero para manejar intervalos no probabilísticos, parezca una opción atractiva.

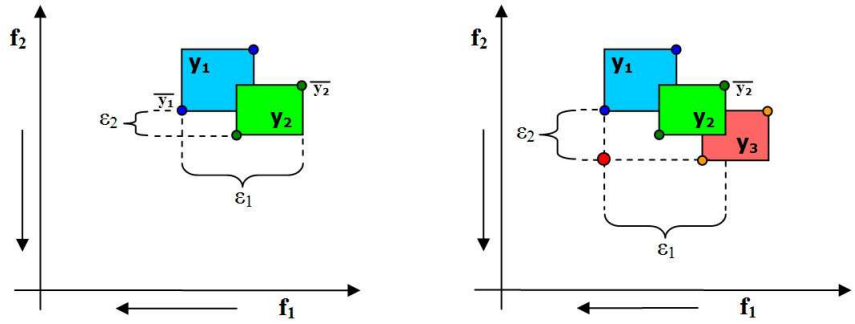


Figure B.18: Comparición de la ϵ-dominancia y el arrepentimiento mínimo para vectores objetivo inciertos (caso minimización): entre dos puntos (izquierda) y entre un punto y un conjunto (derecha).

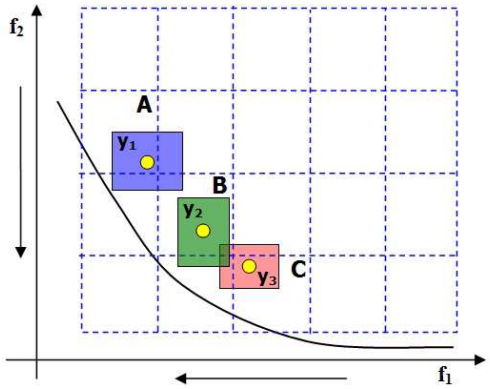


Figure B.19: Ejemplo de almacenamiento mediante *hypergrid* con vectores objetivo inciertos.

Interpretación probabilística de la densidad de un hipercubo

En los análisis de la sección B.4.1 se sugirió la posibilidad de dar una interpretación probabilística al control de densidad en los procedimientos de almacenamiento de soluciones de los MOEA. En este apartado se desarrolla y sugiere la idea de un almacenamiento mediante *hypergrid* con información probabilística. Para ello, considérese la figura B.19 que muestra una sección de un *hypergrid* con tres vectores objetivos con incertidumbre $\mathbf{y_1}, \mathbf{y_2}, \mathbf{y_3}$ y sus correspondientes centroides, que coinciden con el valor esperado si se supone una distribución uniforme o normal, por ejemplo.

Considérese el caso de almacenamiento usando los centroides. El procedimiento `Truncation`(P_A) ejecutado por el *hypergrid* consistiría en seleccionar el hipercubo con mayor número de soluciones y eliminar de allí una solución escogida al azar; pero dado que los hipercubos **A**, **B** y **C** tienen la misma densidad, si se pretende reducir el archivo en un elemento, cualquier alternativa

sería igualmente válida (suponiendo que $\mathbf{y_1}$ y $\mathbf{y_3}$ no están siendo favorecidos por ser puntos extremos). Por tanto, considerar los valores esperados como patrón para manejar el almacenamiento se traduciría en tratar a las alternativas mencionadas como iguales. No obstante, el hipercubo $\mathbf{B}$ tiene la posibilidad de contener hasta tres soluciones simultáneamente, mientras que los demás no, por lo que parece lógico que $\mathbf{B}$ posea mayor densidad.

Una posible estrategia es hacer que la densidad d_p se convierta en un vector N_p-dimensional, donde N_p es el número de candidatos a permanecer en el archivo. Por ejemplo, la densidad del hipercubo $\mathbf{A}$ en la fig. B.19 sería

$$d_p(\mathbf{A}) = (P(|\mathbf{A}| = 1), P(|\mathbf{A}| = 2), \ldots, P(|\mathbf{A}| = N_p))$$

con $N_p = 3$. $P(|\mathbf{A}| = 1)$ es la probabilidad de que el hipercubo $\mathbf{A}$ contenga un punto, la cual coincide con $P(\mathbf{y_1} \in \mathbf{A})$. Dado que solo $\mathbf{y_1}$ se solapa a $\mathbf{A}$, $P(|\mathbf{A}| = 2) = 0$ y $P(|\mathbf{A}| = 3) = 0$. Del mismo modo, el vector densidad $d_p(\mathbf{C})$ del hipercubo $\mathbf{C}$ sería $(P(\mathbf{y_3} \in \mathbf{B}), 0, 0)$.

Considérese ahora al hipercubo $\mathbf{B}$, cuyo vector de densidad $d_p(\mathbf{B})$ contiene componentes no nulos:

$$\begin{aligned}
P(|\mathbf{B}| = 1) &= \sum_{i=1}^{N_p} \left(P(\mathbf{y_i} \in \mathbf{B}) \prod_{j=1;j\neq i}^{N_p} P(\mathbf{y_j} \notin \mathbf{B}) \right) \\
P(|\mathbf{B}| = 2) &= \sum_{i=1}^{N_p} \left(P(\mathbf{y_i} \notin \mathbf{B}) \prod_{j=1;j\neq i}^{N_p} P(\mathbf{y_j} \in \mathbf{B}) \right) \\
P(|\mathbf{B}| = 3) &= \prod_{i=1}^{N_p} P(\mathbf{y_i} \notin \mathbf{B})
\end{aligned}$$

Obsérvese que no se consideran probabilidades condicionales pues los eventos se suponen independientes.

Para calcular las probabilidades anteriores, la CDF de $\mathbf{y_1}$,$\mathbf{y_2}$,$\mathbf{y_3}$ debe ser determinada analíticamente o estimada mediante simulación. El uso de CDF permitiría también emplear probabilidades imprecisas, considerando los límites inferior y superior. Por otro lado, el cálculo de todos los componentes del vector puede ser muy costoso, dado que implica 2^{N_p} términos. De allí que se pueda reducir el número de elementos N_p, por ejemplo al máximo de centroides contenido en un hipercubo.

El procedimiento de reducción del archivo (*truncation*) consistiría en eliminar la alternativa que contribuye más a la densidad. Una posibilidad sería emular el procedimiento ordinario del *hypergrid*, seleccionando primero el hipercubo con mayor probabilidad no nula de poseer la mayor cantidad de puntos menos uno. En el ejemplo se refiere al hipercubo $\mathbf{B}$, que tras eliminar un solución, es el único que tendría una probabilidad no nula de poseer dos soluciones simultáneamente ($P(|\mathbf{B}| = 2) > 0$). Posteriormente, el punto a eliminar sería aquel que contribuye más a reducir la probabilidad considerada, esto es:

$$\mathbf{y_i} = \arg\min_i \{P(|\mathbf{B}| = 2 | \mathbf{y_i} \notin \mathbf{B})\}$$

La implementación de la interpretación probabilística de la densidad en el *hypergrid* no es abordada en este trabajo, quedando abierta para ser ensayada,

evaluada y modificada en el futuro. No obstante, conviene recalcar que, debido a la carencia de estudios y propuestas similares en EMO, cualquier esfuerzo adicional, ya sea teórico o simplemente especulativo, resultará valioso a la vez que necesario.

B.4.3 Contribuciones acerca del concepto de robustez

El concepto de robustez ha sido estudiado por varios autores, los cuales han propuesto definiciones distintas, aunque más o menos convergentes, sobre la materia. Para B. Roy, la robustez es una aptitud para resistir a aquello que es aproximado o mal definido. R. Hites por su parte, enfatiza que la percepción de dicha aptitud es bastante subjetiva, [80, pg. 18], y depende de los criterios de la DM.

Por su parte, P. Vincke [205] señala cuatro tipos de problemas relativos a la robustez: 1) decisiones en contextos dinámicos, 2) solución de problemas de optimización, 3) conclusiones o 4) métodos. Obviamente todos los tipos mencionados se ven afectados por la incertidumbre en el espacio de decisión y en los parámetros ambientales, además de otras fuentes relacionadas con cada tipología. De todas ellas, interesa especialmente a esta tesis la robustez en optimización.

Para S. Sayin, "*la robustez se refiere a la habilidad de un sujeto para enfrentar bien la incertidumbre*"[182, pg. 6]. En optimización, dicha habilidad es perseguida de diferentes formas (ver [16]). Por ejemplo, para Kouvelis y Yu ([182, pg. 6]) la robustez se alcanza minimizando la máxima desviación del mejor caso. Otros enfoques considerados menos pesimistas o conservadores, emplean momentos centrales, como en los trabajos de S. Tsutsui & A. Ghosh [200] y M. Sevaux & K. Sörensen [185] donde se relaciona la robustez con el valor esperado.

En otros casos, el valor esperado no es suficiente para caracterizar la robustez, y se ha recurrido a la varianza, siguiendo la línea propuesta por G. Taguchi [38]. Robustez equivale entonces a insensibilidad a la incertidumbre de entrada, que en la práctica es buscada optimizando la media y restringiendo o minimizando la varianza en formulaciones monocriterio y multicriterio, apoyadas en métodos clásicos o evolutivos (ver [149, 137, 29, 123, 143, 162, 66, 132, 93, 15]).

La totalidad de las formulaciones referidas anteriormente tienen en común que la robustez es referida al espacio de objetivos $\mathbf{Y}$. Una visión diferente viene de considerar la robustez en el espacio $\mathbf{X}$, relacionándola con el vector nominal $\mathbf{x}$ en términos de las desviaciones de los valores nominales que se pueden aceptar sin afectar los requerimientos de la DM. Este enfoque busca definir un dominio efectivo de trabajo. Por ejemplo E. Hendrix et al. [76] estudian distintas formulaciones de dicho dominio mediante métricas L_p con $p \in \{1, 2, \infty\}$ alrededor de un centroide $\mathbf{x}$. En el área de EA, esta idea ha sido estudiada con éxito en formulaciones monocriterio [155, 156] y en EMO por el presente autor [158, 178, 179, 176, 180]. B. Roy también ha propuesto conceptos similares para casos discretos [170]. Nótese también la convergencia entre esta visión y aquella dada a la robustez en la teoría de info-gap.

A continuación se presenta una propuesta de taxonomía original del concepto de robustez a través de una clasificación según la información disponible para la DM. Esta clasificación surge del análisis de las propuestas citadas anteriormente

y en respuesta a una encuesta realizada entre investigadores del área (anexo A), donde se evidencia la diversidad de acepciones del concepto de robustez.

Formulación general de un programa de robustez

Considérese un modelo de desempeño de un sistema, $F(\mathbf{x})$, donde $F : \mathbb{R}^n \to \mathbb{R}^k$ relaciona el espacio de decisión $\mathbf{X} \subseteq \mathbb{R}^n, n \geq 1$ con el espacio de atributos $\mathbf{Y} \subseteq \mathbb{R}^k$ compuesto por $k \geq 1$ medidas de calidad de dicho desempeño. Para este caso, la evaluación de calidad correspondería a evaluar $F(\mathbf{x})$, mientras que la optimización del desempeño conduce a un programa del tipo

$$\begin{aligned} & Opt\,(F(\mathbf{x})) \\ \text{s.a.:} \quad & \\ & G(\mathbf{x}) \leq 0 \end{aligned}$$

donde k determina el carácter mono o multicriterio de las soluciones. $G(\mathbf{x})$ es un vector de restricciones de igualdad o desigualdad, establecidos por la DM en función de restricciones físicas o de metas.

Un modelo más informativo, que considere los parámetros ambientales $\mathbf{p}$, puede ser expresado como $F(\mathbf{x}, \mathbf{p})$, tal que el desempeño del sistema es también función de tales parámetros incontrolables. La preocupación por la robustez surge entonces cuando la respuesta o comportamiento del sistema es incierto, causado por incertidumbre aleatoria o epistémica en $\mathbf{x}$ y/o $\mathbf{p}$. Si intentamos enfocar la robustez a partir de los conceptos mencionados, la formulación más general sería:

$$\begin{aligned} & R\,(F(\mathbf{x}), \mathbf{x}, \mathbf{p}, \delta_x, \delta_p, \gamma) \\ \text{s.a.:} \quad & \\ & \mathbf{x} \in \mathbf{X}\ ,\ \mathbf{p} \in \mathbf{P} \\ & \underline{\delta_x} \leq \delta_x \leq \overline{\delta_x} \\ & \underline{\delta_p} \leq \delta_p \leq \overline{\delta_p} \end{aligned}$$

que puede leerse como *la robustez de un sistema, determinado por sus vector nominal de decisión* $\mathbf{x}$ *en el espacio de decisión* $\mathbf{X}$*, es un criterio definido en términos de la variación de su desempeño* $F(\mathbf{x})$*, respecto a una cantidad de referencia* γ*, y a la incertidumbre asociada con* $\mathbf{x}$*, denotada* δ_x*, y su equivalente* δ_p *en el espacio de parámetro ambientales* $\mathbf{P}$.

La función $R(\cdot)$ es central en esta visión, y puede ser definida de diversas formas según sean las necesidades de la DM. En todo caso, es entendida como un tipo especial de criterio cuya maximización implica una mejora en el sistema ($R(\mathbf{a}) > R(\mathbf{b}) \Leftrightarrow \mathbf{a} \succ \mathbf{b}$) cuando el objetivo perseguido es aumentar la robustez del sistema. Para efectos de este trabajo, la robustez es definida como un objetivo a ser maximizado.

La definición anterior es suficientemente general para manejar dominios continuos o discretos. Por otro lado, la función $R(\cdot)$ puede ser definida de varias formas, como un clasificador de robustez (*no robusto* y *robusto*) con $R : \mathbb{R}^k \to \{0, 1\}$ o como una función continua $R : \mathbb{R}^k \to [0, 1]$. En la práctica, sin embargo, el problema real está en determinar la mejor manera de definir dicha función.

Por simplicidad, la incertidumbre asociada a $\mathbf{x}$ es expresada como $\mathbf{x}+\delta_x$ para ambos tipos de incertidumbre. Así, en el caso epistémico, $\mathbf{x} + \delta_x$ indica que el valor real se desvía del nominal $\mathbf{x}$ en δ_x unidades, resultando $[\mathbf{x} + \underline{\delta_x}, \mathbf{x} + \overline{\delta_x}]$. De forma similar, para el caso aleatorio la formulación es la misma, pero ahora

existe una PDF que gobierna el comportamiento de $\mathbf{x}$. El mismo razonamiento aplica para $\mathbf{p}$.

Del análisis anterior se desprende que la búsqueda de robustez puede expresarse, de la forma más general posible, a través del siguiente *programa de búsqueda de robustez* (def. 20): *Sea $F(\mathbf{x})$ un medida del desempeño de un sistema determinado por el vector $\mathbf{x}$ e influenciado por el vector de parámetros ambientales $\mathbf{p}$, cada uno de los cuales está sujeto a las incertidumbres δ_x y δ_p respectivamente. Sea $G(\cdot)$ un vector de restricciones (de inecuaciones y ecuaciones) definido para la optimización de $F(\mathbf{x})$, y finalmente sea $I(\cdot)$ un vector de restricciones de robustez, impuestas sobre el desempeño. La optimización de la robustez consiste en resolver el siguiente programa*

$$
\begin{aligned}
& Opt\left(R\left(F(\mathbf{x}), \mathbf{x}, \mathbf{p}, \delta_x, \delta_p, \gamma\right)\right) \\
\text{s.a.:} \quad & \\
& \mathbf{x} \in \mathbf{X}\ ,\ \mathbf{p} \in \mathbf{P} \\
& \underline{\delta_x} \leq \delta_x \leq \overline{\delta_x} \\
& \underline{\delta_p} \leq \delta_p \leq \overline{\delta_p} \\
& G(\mathbf{x}, \mathbf{p}, \delta_x, \delta_p) \leq 0 \\
& I(F(\mathbf{x}), \mathbf{x}, \mathbf{p}, \delta_x, \delta_p, \gamma) \leq 0
\end{aligned}
$$

La cantidad de información disponible determina la formulación particular de $R(\cdot)$ y $I(\cdot)$, y por tanto la clase de programa a resolver. Esencialmente existen dos clases de programas, con procedimientos bien definidos para resolverlos, y un tercer programa que surge de la mezcla de los anteriores [177].

Clase 1: programas de propagación de incertidumbre

Esta clase está caracterizada por una descripción adecuada de $\underline{\delta_x} \leq \delta_x \leq \overline{\delta_x}$ y $\underline{\delta_p} \leq \delta_p \leq \overline{\delta_p}$, de tal modo que la incertidumbre puede ser propagada a través de $F(\mathbf{x})$. Si la misma es aleatoria, δ_x y δ_p tienen PDF asociadas. Por ejemplo, δ_x puede estar distribuida normalmente ($N(0, \sigma)$) tal que $\underline{\delta_x} = -\infty$ y $\overline{\delta_x} = \infty$. Por el contrario, si la incertidumbre es epistémica, $\underline{\delta_x}$ y $\overline{\delta_x}$ son escalares finitos.

La robustez suele asociarse con la media y la varianza en esta clase. La tabla B.8 muestra algunas de las formulaciones presentes en la literatura, estando todas asociadas a razonamientos probabilísticos o no probabilísticos.

No se conocen extensiones del programa de búsqueda de robustez Clase 1 a lógica difusa o probabilidades imprecisas. No obstante, este tipo de aplicaciones es posible en la práctica mediante la dominancia de intervalos (SOI y MOI). Existen en la literatura algunos acercamientos a la estimación de la media y varianza en lógica difusa y probabilidades imprecisas, que podrían servir de base para comparar conjuntos y construir MOEA. La tabla B.9 presenta algunas contribuciones en la materia.

Clase 2: programas para determinar dominios efectivos

En algunos casos no hay suficiente información para definir adecuadamente a $\underline{\delta_x} \leq \delta_x \leq \overline{\delta_x}$ y/o $\underline{\delta_p} \leq \delta_p \leq \overline{\delta_p}$ sin hacer suposiciones adicionales. En consecuencia, no es posible propagar la incertidumbre usando una función Clase 1 $R(\cdot)$, sin correr el riesgo de incurrir en error. En estos casos, en lugar de minimizar la variabilidad en $\mathbf{Y}$ se intenta maximizar el número de estados reales aceptables que el vector nominal $\mathbf{x}$ puede tomar.

Funciones	Autores y trabajos
Esperanza como valor representativo:	
• $R(F(\mathbf{x}),\mathbf{x},\delta_x) = \frac{1}{n}\sum_{i=1}^{n} F(\mathbf{x}+\delta_{x,i})$	S. Tsuitsui & A. Ghosh [200], M. Sevaux & K. Sörensen [185, 193], Y-S. Ong et al [147], I. Paenke et al [148].
• $R(F(\mathbf{x}),\mathbf{x},\delta_x) = \left(f_1^{\text{eff}}(\mathbf{x},\delta_x),\ldots,f_k^{\text{eff}}(\mathbf{x},\delta_x)\right)^t$ s.t.: $I(F(\mathbf{x}),\mathbf{x},\delta_x) \equiv \frac{\lvert f_i^{\text{eff}}(\mathbf{x}+\delta_{x,i})-f_i(\mathbf{x})\rvert}{\lvert f_i(\mathbf{x})\rvert} \leq \eta$, where $f_i^{\text{eff}}(\mathbf{x},\delta_x) = \frac{1}{n}\sum_{i=1}^{n} f_i(\mathbf{x}+\delta_{x,i})$ and $F(\mathbf{x}) = (f_1(\mathbf{x}), f_2(\mathbf{x}),\ldots,f_k(\mathbf{x}))^t$	K. Deb & H. Gupta [34].
Esperanza y varianza como valores representativos:	
• $R(F(\mathbf{x}),\mathbf{x},\delta_x) = \frac{\sigma[F(\mathbf{x})]}{\overline{\sigma}[\mathbf{x}]}$ donde $\sigma[\cdot]$ es la desviación estándar de los argumentos y $\overline{\sigma}[\mathbf{x}]$ es el promedio de $\sigma[x_i]$.	Y. Jin & B. Sendhoff [93].
• $R(F(\mathbf{x}),\mathbf{x},\delta_x) = w_1 E[F(\mathbf{x})] - w_2\sigma^2[F(\mathbf{x})]$ donde $E[\cdot]$ es la función esperanza matemática.	D.W. Coit, T. Jin & N. Wattanapongsakorn [29].
• $R(F(\mathbf{x}),\mathbf{x},\delta_x) = \left(E[F(\mathbf{x})], -\sigma^2[F(\mathbf{x})]\right)^t$ donde $\sigma^2[\cdot]$ es la función varianza de los argumentos.	D. Greiner et al [66], M. Marseguerra et al [132], D.W. Coit et al [29], I. Paenke et al [148].
Valores extremos como valores representativos:	
• $R(F(\mathbf{x}),\mathbf{x},\delta_x) \equiv$ peor caso$\{F(\mathbf{x}+\delta_{x,i}) : \forall i\}$	Y-S. Ong et al [147].
• $R(F(\mathbf{x}),\mathbf{x},\delta_p) \equiv$ peor caso$\{F(\mathbf{x}+\delta_{p,i}) : \forall i\}$ donde los escenarios están definidos por los valores que toma el vector $\mathbf{p}$.	P. Kouvelis & G. Yu (cf. [80]).
• $R(F(\mathbf{x}),\mathbf{x},\delta_p) \equiv$ Peor caso estimado para ciertos niveles de confianza y contenido de probabilidad.	S. Martorell et al. [136], J.F. Villanueva et al. [204].

Table B.8: Algunos ejemplos de funciones $R(\cdot)$ Clase 1 en la literatura.

La Clase 2 se caracteriza por la existencia de restricciones de desempeño del tipo $I(F(\mathbf{x}),\mathbf{x},\mathbf{p},\delta_x,\delta_p,\gamma) \leq 0$ que constituyen metas o requerimientos de calidad, junto con una función de robustez $R(\cdot)$, que puede definirse de muchas maneras, pero que siempre orientada a maximizar el dominio aceptable de las variables.

La expresión funcional de $R(\cdot)$ está abierta a investigación. Una posibilidad es definirla en función de la métrica L_p, maximizando esta distancia respecto al centroide de valores nominales $\mathbf{x}$ [76, 140, 139]. Otros enfoque determinan la cardinalidad de la desviación aceptada, como en la definiciones propuestas por B. Roy para casos discretos [170], o el uso de métricas de Lebesgue para el caso continuo, como en [158, 155, 156] o de forma indirecta como plantean C. Barrico y C.H. Antunes [10].

Considérese en programa de búsqueda de robustez; *la hipercaja interior (def. 21)* $\mathbf{IB}$ *de un vector nominal* $\mathbf{x}$*, denotada* $\mathbf{IB}(\mathbf{x})$ *es un subconjunto definido por:*

$$\mathbf{IB}(\mathbf{x}) = \{\mathbf{z} \in \mathbb{R}^n | z_i \in [x_i + \underline{\delta_{x,i}}, x_i + \overline{\delta_{x,i}}]\}$$

tal que $G(\cdot) \leq 0$ *y* $I(\cdot) \leq 0$ *se cumplen.* Tal definición coincide con la máxima cápsula convexa definible para el producto cartesiano $[x_1+\underline{\delta_{x,1}}, x_1+\overline{\delta_{x,1}}] \times \ldots \times [x_i + \underline{\delta_{x,i}}, x_i + \overline{\delta_{x,i}}] \times \ldots \times [x_n + \underline{\delta_{x,n}}, x_n + \overline{\delta_{x,n}}]$, un hipercubo definido a partir del conjunto de n intervalos $[x_i + \underline{\delta_{x,i}}, x_i + \overline{\delta_{x,i}}]$, cada uno de los cuales está

Tema de estudio	Autores y trabajos
Lógica difusa:	
Estimación de media y varianza.	R. Kröger [113], D. Dubois and H. Prade [42], A.G. Bronevich [21], C. Carlsson & R. Fullér [25].
Dempster-Shafer:	
Cálculo de media y varianza.	V. Kreinovich, G. Xiang & S. Ferson [112]
p-boxes:	
Estimación del valor esperado.	M. Bruns, C.J.J. Paredis & S. Ferson [22, 23].
Intervalos:	
Cálculo rápido de estadísticos para intervalos.	G. Xiang [218], S. Ferson, L. Ginzburg, V. Kreinovich & J. Lopez [46].
Simulación Monte Carlo:	
Técnicas para procesar intervalos de incertidumbre en ingeniería.	V. Kreinovich, J. Beck, C. Ferregut, A. Sanchez, G. R. Keller, M. Averill & S. A. Starks [111].

Table B.9: Algunas contribuciones al cálculo de media y varianza de cantidades difusas o probabilísticas imprecisas, que podrían ayudar a resolver funciones $R(\cdot)$ Clase 1.

definido para la componente i-ésima de $\mathbf{x}$, tal que $G(\cdot) \leq 0$, $I(\cdot) \leq 0$ y $\mathbf{x} \in \mathbf{X}$ se cumplen.

La definición anterior es intuitiva en la práctica, puesto que permite la construcción de intervalos simétricos alrededor del punto de interés. Luego, la cardinalidad de una caja interior viene dada por su métrica de Lebesgue:

$$|\mathbf{IB}(\mathbf{x})| = \prod_{i=1}^{n} |\overline{\delta_{x,i}} - \underline{\delta_{x,i}}|$$

y si adicionalmente se requiere la simetría, tal que $x_i = (\overline{\delta_{x,i}} - \underline{\delta_{x,i}})/2$, entonces la expresión anterior puede ser reescrita como

$$|\mathbf{IB}(\mathbf{x})| = 2^n \prod_{i=1}^{n} |\overline{\delta_{x,i}} - x_i| = 2^n \prod_{i=1}^{n} |\underline{\delta_{x,i}} - x_i| \equiv \prod_{i=1}^{n} |\overline{\delta_{x,i}} - x_i|$$

que corresponde a una definición en términos del centroide $\mathbf{x}$. De aquí que, para maximizar la robustez Clase 2, y bajo la condición de simetría dado un centroide (fig. B.20 izquierda), el problema equivalga a resolver un *programa de maximización de caja interior para centroide fijo* (def. 22): *Encontrar el límite inferior del vector nominal* $\underline{\delta_x}$ *tal que:*

$$\underline{\delta_x} = \arg\max_{\underline{\delta_x}} R\left(F(\mathbf{x}), \mathbf{x}, \delta_x\right)$$

s.a.:

$$\begin{gathered} R\left(F(\mathbf{x}), \mathbf{x}, \delta_x\right) = |\mathbf{IB}(\mathbf{x})| \equiv \textstyle\prod_{i=1}^{n} |\underline{\delta_{x,i}} - x_i| \\ \mathbf{x} \in \mathbf{X} \\ G(\mathbf{x}, \delta_x) \leq 0 \\ I(F(\mathbf{x}), \mathbf{x}, \delta_x, \gamma) \leq 0 \end{gathered}$$

De forma similar, si el centroide es una variable de decisión (fig. B.20 izquierda) se tiene el *programa de maximización de caja interior para centroide*

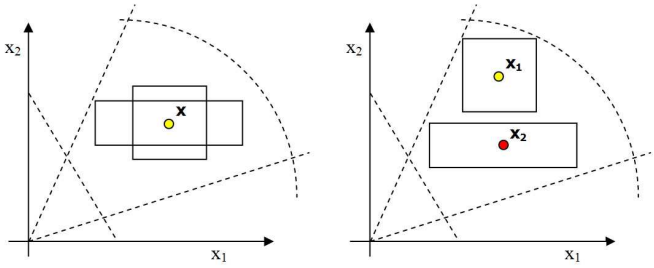


Figure B.20: Ejemplos de distintas cajas internas para centroides especificados (izquierda) y no especificados (derecha) (las restricciones $G(\cdot)$ e $I(\cdot)$ están representadas por líneas discontinuas).

fijo (def. 23): *Encontrar el límite inferior del vector nominal* $\underline{\delta_x}$ *tal que:*

$$\mathbf{x} = \arg\max_{\mathbf{x},\underline{\delta_x}} R\left(F(\mathbf{x}), \mathbf{x}, \delta_x\right)$$

s.a.:

$$R\left(F(\mathbf{x}), \mathbf{x}, \delta_x\right) = |\mathbf{IB}(\mathbf{x})| \equiv \prod_{i=1}^{n} |\underline{\delta_{x,i}} - x_i|$$
$$\mathbf{x} \in \mathbf{X}$$
$$G(\mathbf{x}, \delta_x) \leq 0$$
$$I(F(\mathbf{x}), \mathbf{x}, \delta_x, \gamma) \leq 0$$

La expansión de la Clase 2 a otras teorías de incertidumbre parece posible en algunos casos (ver sec. 4.3.3) aunque no se aborda en este resumen. Cabe mencionar, sin embargo, que interpretaciones posibilísticas, probabilísticas con información incompleta y de lógica difusa han sido consideradas. En particular para esta última, se propone la ecuación 4.19, que establece como medida de robustez a $\overline{\mathcal{A}_\alpha} = \int_\alpha \alpha \mathcal{A}_\alpha d\alpha$, la cual puede ser estimada como $\sum_{\alpha_i} \alpha_i \mathcal{A}_{\alpha_i}$, constituyendo así una posible función $R(\cdot)$ de Clase 2 para variables con incertidumbre difusa.

Clase 3: procedimiento mixto para la búsqueda de robustez

Para finalizar, considérese una vez más el programa de búsqueda de robustez. Las dos clases estudiadas hasta ahora surgen de un conocimiento parcial, ya sea de la incertidumbre en la entrada (Clase 1) o de las especificaciones de la DM para la salida (Clase 2). Sin embargo, si la DM no posee información sobre ninguno de dichos aspectos, no es posible establecer δ_x, δ_p ni $I(F(\mathbf{x}), \mathbf{x}, \delta_x, \gamma)$, resultando imposible definir $R(\cdot)$.

En ese caso, es obligatorio generar algún tipo de información que ayude al analista/DM a establecer δ_x, δ_p o $I(\cdot)$, de forma que el problema decante en uno de Clase 1 o Clase 2. Tal procedimiento es ejemplificado más adelante.

B.4.4 AUREO: Análisis de robustez desde una perspectiva basada en la información disponible

En este apartado se sistematizan los conceptos y propuestas presentadas hasta el momento en una metodología dirigida a asistir al analista/DM en la definición

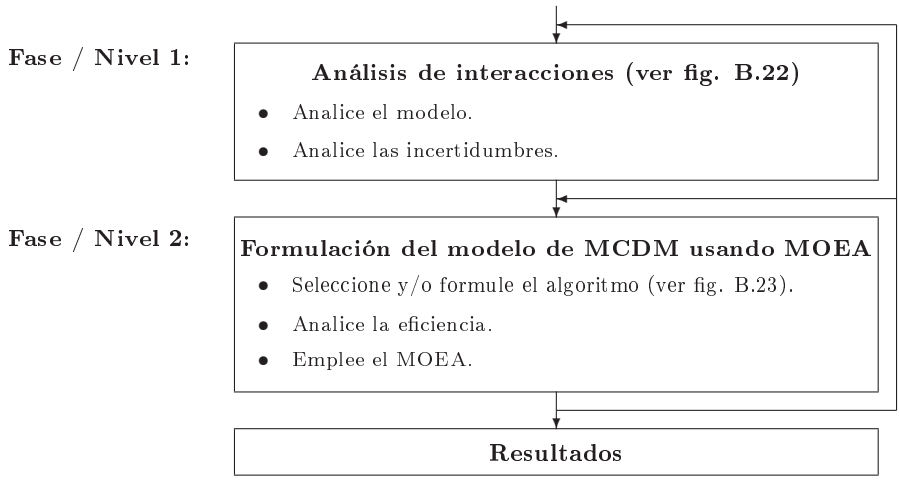


Figure B.21: AUREO: Dos niveles de análisis.

del problema de decisión tomando en cuenta la incertidumbre, y en la escogencia y/o adaptación de la metaheurística a emplear para solucionar el problema.

La metodología propuesta llamada *Análisis de incertidumbre y robustez en optimización evolutiva* (AUREO), se fundamenta en la premisa de que la información disponible es la que determina el tipo de formulación del programa matemático a resolver, y que a su vez tiene influencia en la formulación, escogencia y adaptación del MOEA a emplear en la resolución de dicho programa.

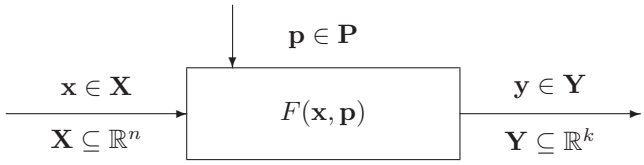
El método AUREO se compone de dos etapas, como se muestra en la figura B.21: la primera se enfoca en los tipos, fuentes y efectos de la incertidumbre respecto a los atributos a optimizar o controlar. Adicionalmente, esta etapa se ocupa de transformar el problema original en un programa de búsqueda de robustez. Por otra parte, en la segunda etapa, el esfuerzo se centra en la implementación del MOEA, las alternativas para propagar la incertidumbre, la eficiencia algorítmica y la comparación de resultados. Las etapas no se desarrollan necesariamente en una secuencia lineal, sino más bien en un lazo o espiral.

AUREO: Etapa 1

El objetivo principal durante la primera etapa es determinar las alternativas para reformular el problema original para manejar su incertidumbre. En consecuencia, el análisis está dirigido a identificar las fuentes y las áreas de influencia de la incertidumbre asociada al problema, así como sus efectos. En la figura B.22 se muestran las tres instancias sobre las cuales de realiza este análisis.

En primer lugar, la atención se centra en el modelo, ya que el mismo debe ser adecuado para representar los atributos de interés que se pretende mejorar. Si las funciones objetivo no describen correctamente dichos atributos, el control de incertidumbre se vuelve inútil, puesto que las funciones no representan lo que deberían. Por otro lado, las características del dominio (continuo o discreto) y de las funciones objetivo determinan no solamente la aplicabilidad de uno u otro método de resolución, sino la forma de reformular el problema.

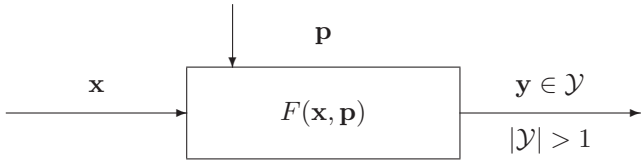
1. Analice el modelo:



1.1. Verifique que el modelo es adecuado.

1.2. Considere las características del dominio y las funciones objetivo.

2. Constate si hay funciones objetivo con incertidumbre:



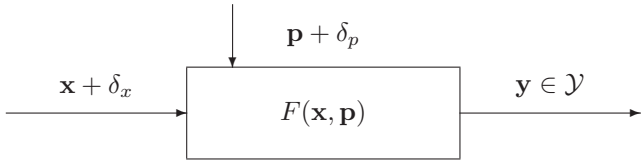
2.1. ¿Distintas evaluaciones del mismo argumento producen distintos resultados?

2.2. ¿Es $F(\mathbf{x}, \mathbf{p})$ una función dinámica o estocástica?

2.3. ¿Cómo se evaluará $F(\mathbf{x}, \mathbf{p})$ (emulación, aproximación, simulación)?

2.4. ¿La cardinalidad de $\mathcal{Y} \subseteq \mathbf{Y}$ puede reducirse a la unidad?

3. Verifique la incertidumbre en la entrada:



3.1. ¿Está $\mathbf{x}$ afectado por incertidumbre (δ_x)? En caso afirmativo, ¿de qué tipo?

3.2. ¿Pueden variar los parámetros ambientales (δ_p) ?

3.3. ¿Es la función objetivo sensible a argumentos inciertos ($|\mathcal{Y}| > 1$)?

Figure B.22: AUREO Fase 1: Análisis de interacciones modelo-incertidumbre.

Argumento: $\mathbf{x}+\delta_x, \mathbf{p}+\delta_p$		**Respuesta:** $\mathbf{y} \in \mathcal{Y}, \lvert\mathcal{Y}\rvert > 1$	
δ_x	δ_p	$I(\cdot)$ definible	$I(\cdot)$ indefinible
No hay	No hay	**Comparación de conjuntos ≡ Clase 1**	
No hay	Definible	**Clase 1**	**Clase 1**
No hay	Indefinible	**Clase 2**	**Clase 3**
Definible	No hay	**Clase 1**	**Clase 1**
Definible	Definible	**Clase 1**	**Clase 1**
Definible	Indefinible	**Clase 2**	**Clase 3**
Indefinible	No hay	**Clase 2**	**Clase 3**
Indefinible	Definible	**Clase 2**	**Clase 3**
Indefinible	Indefinible	**Clase 2**	**Clase 3**

Table B.10: AUREO Fase 1: Clases de formulaciones para manejar incertidumbre en función de la información disponible.

En segundo lugar, cuando la idoneidad del modelo se ha comprobado, el análisis se centra en verificar si existe incertidumbre asociada naturalmente a las funciones objetivo (función dinámica, función con ruido) o si esta puede surgir de la forma en que las mismas han de ser evaluadas (aproximación, simulación, etc). Puede suceder que en la práctica sea necesario aceptar cierto nivel de incertidumbre para disminuir el esfuerzo computacional.

Por último, el tercer elemento a considerar es la presencia de incertidumbre asociada a las variables del problema. Si dicha incertidumbre existe, esto indica un programa de búsqueda de robustez. Sin embargo, el tipo de incertidumbre y su impacto son factores cruciales, por tanto un análisis preliminar de sensibilidad puede ayudar a determinar si tal incertidumbre puede ser despreciada o no.

Una vez que las incertidumbres han sido identificadas, el analista debe determinar la formulación final del programa de manejo de incertidumbre que se desprende del análisis y de la información disponible. Así, con relación a la tabla B.10, ante un escenario de una función incierta con un dominio perfectamente definido y sin incertidumbre (tercera fila), es un caso de comparación de conjuntos, que en la práctica, es equivalente en su tratamiento con MOEA a un problema de robustez Clase 1, definido para optimizar algunos valores representativos. No obstante, los procedimientos para determinar tales valores difieren entre ambos casos, dado que el programa de robustez implica la propagación de incertidumbre, (ver fig. B.23) mientras que para él escenario descrito tal propagación no tiene lugar. En cambio, sea o no $F(\mathbf{x}, \mathbf{p})$ una función incierta, las variables inciertas siempre conllevan a un programa de búsqueda de robustez, a pesar de que el mismo pueda ser evitado por razones prácticas tras el análisis de sensibilidad. Por ello, el resto de los casos de la tabla B.10 requieren una reformulación del problema en uno de robustez.

AUREO Etapa 2: del programa a la implementación

El nexo entre el ambas etapas de AUREO viene dado por el análisis de alternativas de implementación de la metaheurística multiobjetivo. La figura B.23 muestra tres preguntas fundamentales que resumen los objetivos de análisis de esta etapa. La secuencia de respuesta puede variar de un problema a otro.

Para cada problema específico el analista debe identificar el programa corres-

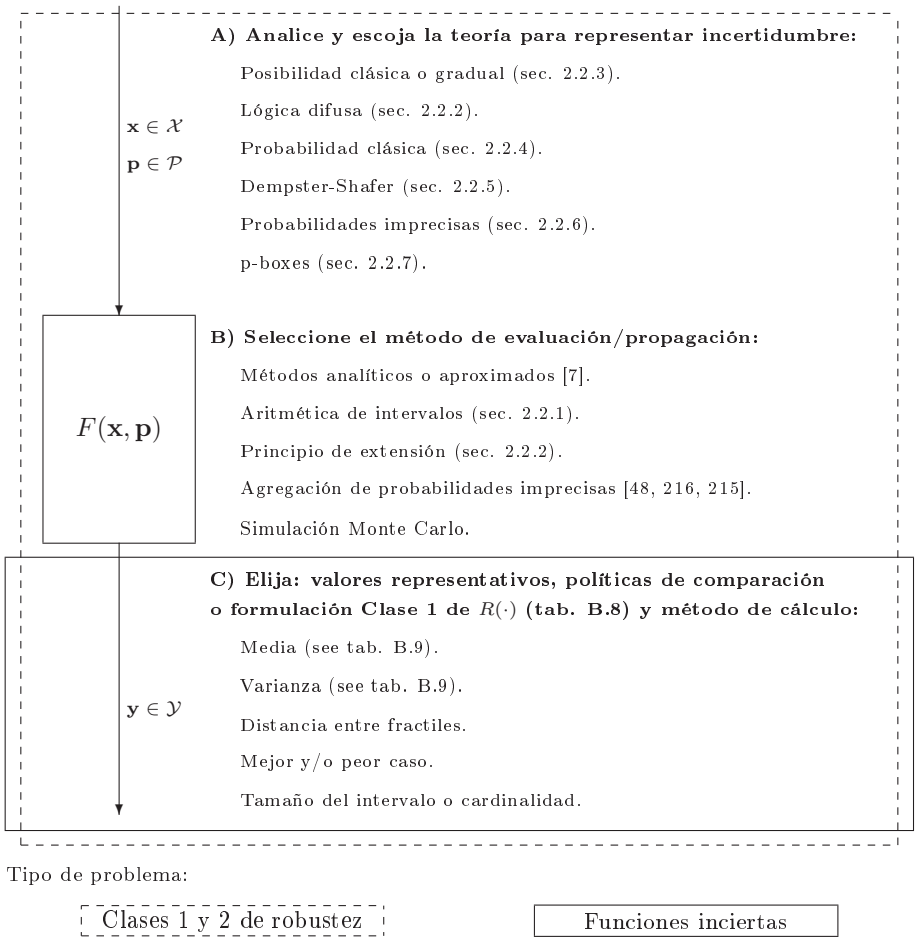


Figure B.23: AUREO Fase 2: Elementos de la formulación de modelos para el manejo de incertidumbre.

pondiente según se prescribe en la tabla B.10. Posteriormente, el análisis debe guiar a identificar el MOEA existente a emplear, con las adaptaciones que sean necesarias, o la herramienta que ha de ser diseñada. La tabla B.11 presenta algunas alternativas existentes y posibles extensiones de las mismas para resolver los tipos de programas prescritos en la tabla B.10.

Si el problema consiste en comparar conjuntos o en uno de robustez Clase 1, el MOEA deberá optimizar una función del tipo $R_1(\cdot)$ cumpliendo las restricciones del tipo $G(\cdot)$ y posiblemente también del tipo $I(\cdot)$. La robustez de Clase 2 requerirá un tratamiento similar pero para funciones del tipo $R_2(\cdot)$. En la práctica, en los problemas de Clase 1 el proceso de propagación se aplica una sola vez, si bien dicho proceso puede implicar evaluar la función objetivo en repetidas ocasiones, para determinar el vector de atributos asociado. En cambio, en la Clase 2 es necesario verificar que las restricciones $I(\cdot)$ se cumplan para cada dominio candidato, es decir para cada $\mathbf{IB}(\mathbf{x})$ evaluado, hasta hallar la caja máxima para un $\mathbf{x}$ particular; luego para un mismo $\mathbf{x}$ se pueden propagar

Formulación para manejo de incertidumbre	**Algoritmos y referencias**	**Posibles ampliaciones**
Comparación de conjuntos:		
• Sin información adicional:		
Optimización con incertidumbre epistémica:	• P. Limbourg [125], • P. Limbourg & D.E. Salazar A. IP-MOEA [126, 173]	• Adaptación de MOEA existentes para trabajar con valores representativos. • Construcción de algoritmos basados en $I_{\epsilon+}$.
• Con información adicional:		
Optimización con intervalos de adaptación	• J. Teich [197]	
Optimización de funciones objetivo con ruido	• E. Hughes [89, 88], • J.E. Fieldsen & R.M. Everson [51]	• Control de la significación estadística para evitar alternativas repetidas y mejorar la convergencia y diversidad.
Optimización basada en un indicator	• M. Basseur & E. Zitzler [11]	
Programas busca robustez:		
• Programas Clase 1:		
Optimización con propagación de incertidumbre:	•Ver comparación de conjuntos en esta tabla y las referencias, • K. Sörensen [194], • T. Ray [152], • K. Deb & H. Gupta [34]	• Adaptación de MOEA existentes (see e.g. sec. 3.3.4) para trabajar con las funciones $R_1(\cdot)$ de la tabla B.8. • Construcción de algoritmos basados en $I_{\epsilon+}$. • Extensiones de los algoritmos existentes para propagar diferentes tipos de incertidumbre usando los procedimientos de la tabla B.9. • Control de la significación estadística para evitar alternativas repetidas y mejorar la convergencia y diversidad.
• Programas Clase 2:		
Búsqueda de dominios:	• C. Barrico and C.H. Antunes [10]	• Adaptación de MOEA existentes para manejar dominios discretos (ver def. 17 to 19).

Table B.11: AUREO Fase 2: Alternativas existentes y posibles ampliaciones en EMO para resolver las formulaciones de manejo de incertidumbre prescritas en la tabla B.10.

los distintos $\mathbf{IB}(\mathbf{x})$ posibles. Esto significa que el esfuerzo de propagación es crítico para ambos tipos de problema, pero puede serlo mucho más para la robustez Clase 2.

Por otro lado, la significación estadística de los valores representativos determinados para funciones tipo $R_1(\cdot)$ podría afectar la eficiencia del MOEA. Considérese, por ejemplo, que dos alternativas clonadas o repetidas pueden tener diferentes valores de adaptación en el EA debido a (1) una función objetivo con ruido, o (2) el uso de métodos aproximados o simulación para propagar incertidumbre. En ambos casos dos alternativas iguales tendrán imágenes distintas. De forma similar, la dominancia matemática no implica dominancia en el sentido estadístico. Así, dos valores representativos pueden ser diferentes, pero dicha diferencia estadísticamente no significativa. Este fenómeno no parece haber sido estudiado en EMO hasta el momento, aunque podría tener un impacto importante sobre la eficiencia del algoritmo.

La implementación eficiente de un MOEA implica varios elementos de su estructura y del problema estudiado. Algunas pautas generales para mejorar la eficiencia algorítmica son:

Representación: Encuentre la mejor forma de codificar el problema. Considere el tipo de dominio (discreto o continuo). ¿La representación encaja bien con el tipo de representación considerada?

Propagación de la incertidumbre: ¿Es necesario propagar? En caso afirmativo, ¿qué características tiene la función objetivo? Considere la aplicabilidad de los métodos analíticos de propagación. En caso de usar simulación, escoja parámetros adecuados.

Asignación de adaptación: ¿La jerarquización maneja individuos repetidos? ¿considera la significación estadística? ¿Cómo se estima la densidad?

Número de objetivos: ¿Cuántos valores representativos están siendo considerados? ¿Se han de considerar todos simultáneamente o de forma lexicográfica?

En la sección siguiente se resumen algunas aplicaciones de AUREO que incluyen ambas fases. Entre las categorías mostradas en la tabla B.11, la Clase 2 es la que ha recibido menos atención en EMO, por tanto, los esfuerzos reportados pertenecen a esta categoría de robustez.

B.5 Aplicación y validación de AUREO en problemas de Seguridad global

En esta sección se resumen las tres aplicaciones de AUREO referidas en los capítulos 5 y 6, relacionadas con distintas aplicaciones de seguridad de funcionamiento.

B.5.1 Aplicación de AUREO a un problema de fiabilidad

El problema considerado corresponde al diseño robusto de un sistema de remoción de calor residual (RHR) de baja presión, que forma parte del sistema de seguridad de una planta de energía nuclear. Para tal sistema, la no fiabilidad

Q_S, es decir la probabilidad de que el sistema falle en un período de tiempo determinado, es estimada mediante la expresión [131]:

$$\begin{aligned} Q_S \approx \; & Q_1Q_3Q_5 + Q_1Q_3Q_6R_5 + Q_1Q_3Q_7R_5R_6 \\ + \; & Q_1Q_3Q_8R_5R_6R_7 + Q_1Q_4Q_5R_3 + Q_1Q_4Q_6R_3R_5 \\ + \; & Q_1Q_4Q_7R_3R_5R_6 + Q_1Q_4Q_8R_3R_5R_6R_7 + Q_2Q_3Q_5R_1 \\ + \; & Q_2Q_3Q_6R_1R_5 + Q_2Q_3Q_7R_1R_5R_6 + Q_2Q_3Q_8R_1R_5R_6R_7 \\ + \; & Q_2Q_4Q_5R_1R_3 + Q_2Q_4Q_6R_1R_3R_5 + Q_2Q_4Q_7R_1R_3R_5R_6 \\ + \; & Q_2Q_4Q_8R_1R_3R_5R_6R_7 \end{aligned}$$

donde Q_i y R_i son la no fiabilidad y la fiabilidad del evento básico R_i, y se consideran 8 eventos básicos que pueden inducir un fallo. La fiabilidad del sistema se estima como $R_S = 1 - Q_S$, y se busca que la misma se mantenga entre $0.99 \leq R_S \leq 1$. Para este sistema, la fiabilidad de los componentes se restringe a $0.80 \leq R_i \leq 1$.

El problema de *diseño robusto*, consiste en determinar la máxima variación aceptable para cada componente R_i, tal que la meta de calidad $0.99 \leq R_S \leq 1$ se cumpla. Como valor nominal, inicialmente se proponen los centroides $c_i = 0.90, i = 1, 2, \ldots, 8$. Adicionalmente, interesa restringir el coste del sistema, calculándolo como $C_S = 2\sum_i K_i R_i^{\alpha_i}$, donde el vector de coeficientes es $K = \{100, 100, 100, 150, 100, 100, 100, 150\}$ y los exponentes asociados a la fiabilidad de los componentes es $\alpha_i = 0.6, \forall i$.

AUREO Etapa 1

El problema planteado de diseño robusto es conceptualmente afín con los programas de búsqueda de robustez, de modo que el análisis en la etapa 1 no puede aportar demasiado a la formulación del problema.

El cálculo de R_S se realiza mediante una función bien definida, de modo que el mismo no es fuente de incertidumbre para el problema. Por otro lado, no existen parámetros ambientales a considerar. La única fuente de incertidumbre que se acepta es la posible variación de los R_i, sin mencionar tampoco si la misma es epistémica o aleatoria.

La formulación final puede hacerse considerando un centroide fijo (que sería $R_i = 0.90$) o uno variable. Para el primer caso, el programa establece encontrar la caja interna máxima (MIB) para $\mathbf{x}$ con costo mínimo C_S, donde $\mathbf{x}$ es un vector de fiabilidades de componentes tal que $\mathbf{x} = (R_1, \ldots, R_8)^t$, sujeto a las restricciones de desempeño $I(R_S) \equiv 0.99 \leq R_S \leq 1$ y a las de diseño $0.80 \leq R_i \leq 1$. Los R_i están centrado en $c_i = c, \forall i$ con c constante, esto es, la fiabilidad real del componente i, R_i, se supone centrada en c_i (valor nominal) y se quiere investigar cuánto puede desviarse de dicho valor. Para el caso de centroide no especificado, basta con conservar el programa anterior eliminando la última restricción. En ambos casos se aceptan solo desviaciones simétricas.

AUREO Etapa 2

La implementación del MOEA en el primer caso es directa, siendo el manejo de restricciones el único factor relevante. Cada $\mathbf{IB}(\mathbf{x})$ analizado debe cumplir las restricciones de diseño $0.80 \leq R_i \leq 1$, que se logran restringiendo el espacio de búsqueda a $R_i = [0.8, 1]$, y las de desempeño $I(R_S) \equiv 0.99 \leq R_S \leq 1$, que requieren un método para propagar $\mathbf{IB}(\mathbf{x})$ de manera que $R_S(\mathbf{IB}(\mathbf{x})) \in$

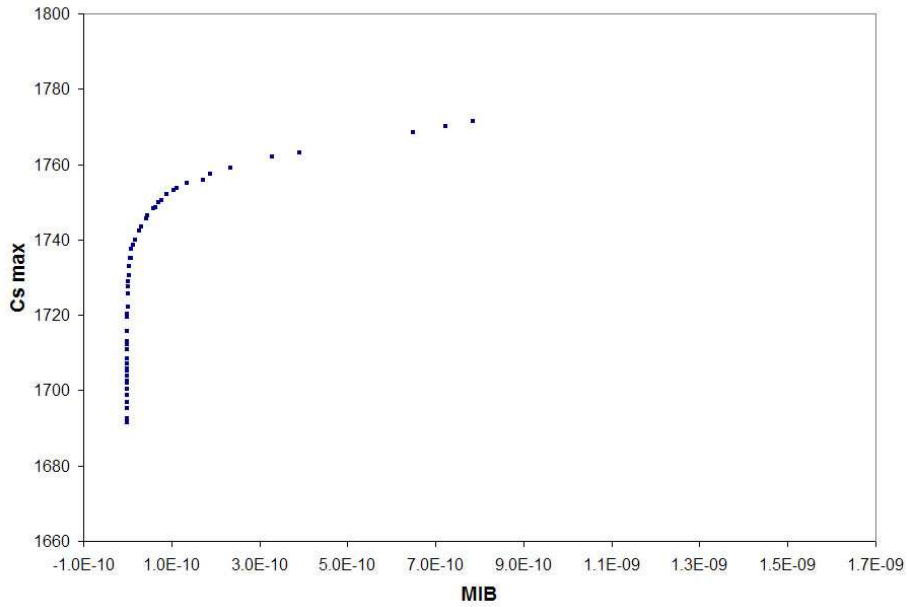


Figure B.24: Frente de aproximación de Pareto robustez-costo con centroide $c_i = 0.90$ [179].

$[0.99, 1]$ pueda ser verificada. A tal fin, R_S puede ser evaluada como una función intervalo, respetando las reglas de aritmética de intervalo. Por otra parte, el MOEA debe implementar un mecanismo de control de restricciones.

La figura B.24 muestra lo resultados de usar NSGA-II con las siguientes características:

- Implementación algorítmica: ver algoritmo en fig. B.7.
- Representación: real.
- Recombinación: cruce y mutación.
- Manejo de restricciones: técnica de Deb [36].
- Tamaño de población $M = 50$, tamaño de archivo $N = 50$, número de generaciones $t^{max} = 250$
- Tasa de cruce $\rho_c = 0.9$, peso del cruce real $\alpha = 0.8$, tasa de mutación $\rho_m = 0.3$.
- Número de evaluaciones funcionales: 12550.

Desde el punto de vista algorítmico, el primer problema está solucionado exitosamente, estableciéndose distintos juegos de variaciones aceptables de los R_i, tal que la DM puede escoger el nivel de robustez que más le convenga respecto a la máxima inversión en la que puede incurrir por ello. No obstante, si el valor del centroide no se fija a priori, el procedimiento descrito no es suficiente para resolver el problema. Se requiere en su lugar un procedimiento iterativo

de doble lazo o lazos anidados, donde el lazo externo fija el valor del centroide, y el lazo interno resuelve el problema para el centroide fijado, como ya se hizo antes.

La figura B.25 presenta el resultado de resolver el problema con un procedimiento de doble lazo. Por simplicidad la restricción $R_i = c_i, \forall i$ se conservó, de manera que un único lazo para variar c_i fuese necesario. El resultado para cada $\mathbf{x}$ se ve en la figura B.25 como puntos pequeños, mientras que la frontera de Pareto obtenida tras explorar distintos c_i se muestra con anillos y líneas punteadas.

Obsérvese que el frente de aproximación no se encuentra en ningún valor particular de c_i, sino que diferentes fracciones corresponden a distintos valores de $\mathbf{x}$, lo que sugiere que la condición inicial de centrar $\mathbf{x}$ en $c_i = c, \forall i$ debe ser eliminada. Tómese en cuenta también la baja eficiencia del doble lazo, que requiere una exploración completa con NSGA-II para cada valor de $\mathbf{x}$. Así, si se desea por ejemplo explorar distintos centroides, como en la figura B.25, es necesario realizar $12500 * 4 = 50000$ evaluaciones de la función objetivo. El número de evaluaciones (12500) por cada centroide puede reducirse afinando parámetros, no obstante la complicación debida a la estructura de lazos anidados permanece. De hecho, dado que el espacio de búsqueda es continuo, si se pretende variar $\mathbf{x}$ en todas sus componentes, se requieren nueve lazos (ocho para fijar las componentes de R_i y uno adicional para encontrar el MIB) por lo que, por ejemplo para variaciones de tamaño 0.01, el número total de evaluaciones es $(0.2/0.01)^9 * t^{max} * M * N$, lo que corresponde a 5.12E11 veces el número de evaluaciones que necesita el MOEA para encontrar el MIB. Evidentemente, la estrategia de lazos anidados no es práctica en la mayoría de los casos.

Representación porcentual

En la fase 2 de AUREO, los esfuerzos deben concentrarse en la eficiencia algorítmica del MOEA. La estructura de lazos anidados presentada anteriormente es claramente ineficiente, por tanto en esta etapa, es un objetivo del analista el buscar formas alternativas de aumentar la eficiencia sin sacrificar la calidad de las soluciones.

En este sentido, considérese que el uso de lazos anidados no permite el intercambio directo de información entre los MIB halladas para cada valor de $\mathbf{x}$. En el caso general de diseño con centroide no especificado y espacio de decisión n-dimensional, las componentes de $\mathbf{x}$ son los centroides c_i y las coordenadas de la caja interna $\mathbf{IB}(\mathbf{x})$, la cual puede expresarse en función de los vértices inferiores $\underline{x_i}$, donde $\underline{x_i} \in [\underline{c_i}, \overline{c_i}], 0 \leq \underline{c_i} \wedge \overline{c_i} \leq 1, i = 1, 2, \ldots, n$. Para cada c_i se tienen las siguientes restricciones: el vértice inferior debe verificar $\underline{c_i} \leq \underline{x_i} \leq c_i$ mientras que el superior, que puede calcularse como $\overline{x_i} = 2c_i - \underline{x_i}$ debido a la condición de simetría, está restringido a $c_i \leq \underline{x_i} \leq \overline{c_i}$. Existe por tanto una dependencia entre los límites de $\underline{x_i}$ y el valor de c_i. Dicha dependencia evita que el algoritmo pueda recombinar individuos óptimos para un valor particular de c_i con cromosomas definidos para c_i diferentes, dado que se pueden generar cromosomas inválidos. En términos sencillos, no se pueden mezclar cajas con diferentes centroides sin correr el riesgo de violar restricciones.

Si en lugar de emplear un esquema ordinario de representación se adopta una representación porcentual [176], las cajas interiores pueden expresarse en términos relativos respecto a $\mathbf{x}$, embebiendo los espacios de búsqueda en uno

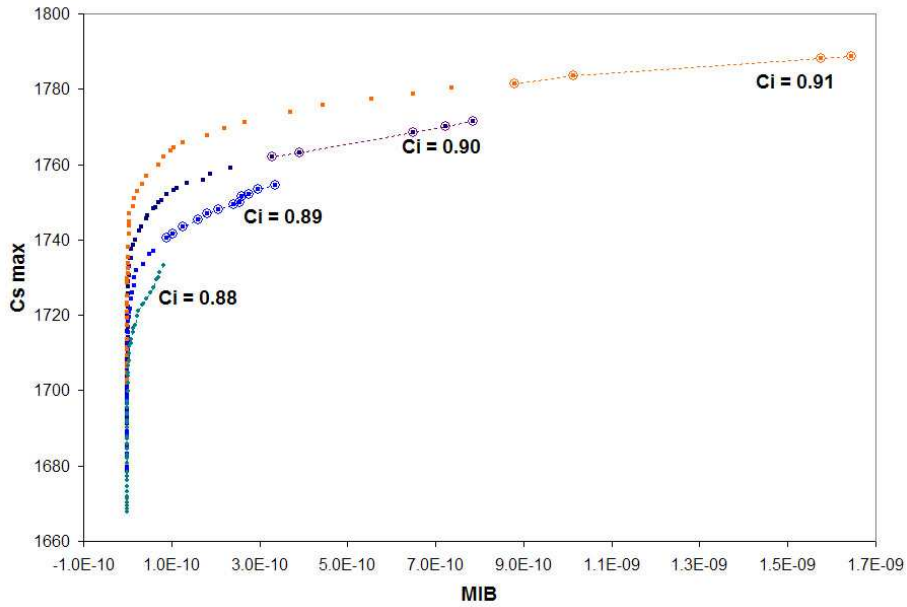


Figure B.25: Frentes de aproximación de Pareto robustez-coste para centroides no especificado, usando el método de doble lazo (los puntos no dominados están resaltados con anillos) [179, 178].

de mayor tamaño, pero que permite la aplicación de los mecanismos de recombinación sin problemas. Para el caso particular descrito antes, cada individuo es codificado como un grupo de 8 pares$(c_i, \%x_i^{\max})$ donde $\%x_i^{\max}$ representa el porcentaje de la distancia o desviación máxima aceptables $x_i^{\max}$ entre c_i y sus límites. Así, la relación matemática para determinar el vérticex_i que garantiza que solo se producirán individuos factibles mediante recombinación es [176]:

$$\%\overline{x_i} = \frac{|c_i - x_i|}{\%x_i^{\max}} = \frac{|c_i - x_i|}{\min\{\overline{c_i} - c_i, c_i - \underline{c_i}\}}$$

Obsérvese que x_i pudiera ser tanto el vértice superior como el inferior. Por tanto es necesario aplicar una regal de decisión en la decodificación, obteniendo $\underline{x_i}$ a partir de [176]:

$$\underline{x_i} = \begin{cases} c_i - \%x_i^{\max} \cdot (c_i - \underline{c_i}) & \text{if} \quad c_i - \underline{c_i} < \overline{c_i} - c_i \\ c_i - \%x_i^{\max} \cdot (\overline{c_i} - c_i) & \text{otherwise} \end{cases}$$

La figura B.26 exhibe el resultado de emplear la representación porcentual con el NSGA-II para el problema de centroide no especificado. Las soluciones obtenidas claramente superan en calidad a las obtenidas con el doble lazo, dado que es posible obtener valores de MIB mucho más grandes a menor coste máximo. Adicionalmente, la frontera de aproximación de Pareto fue extendida alrededor de ocho veces el máximo valor de MIB originalmente hallado con una sola corrida del MOEA. De esta forma, los resultados ejemplifican bien la esencia de AUREO y cómo, una vez solucionado el problema de formular el

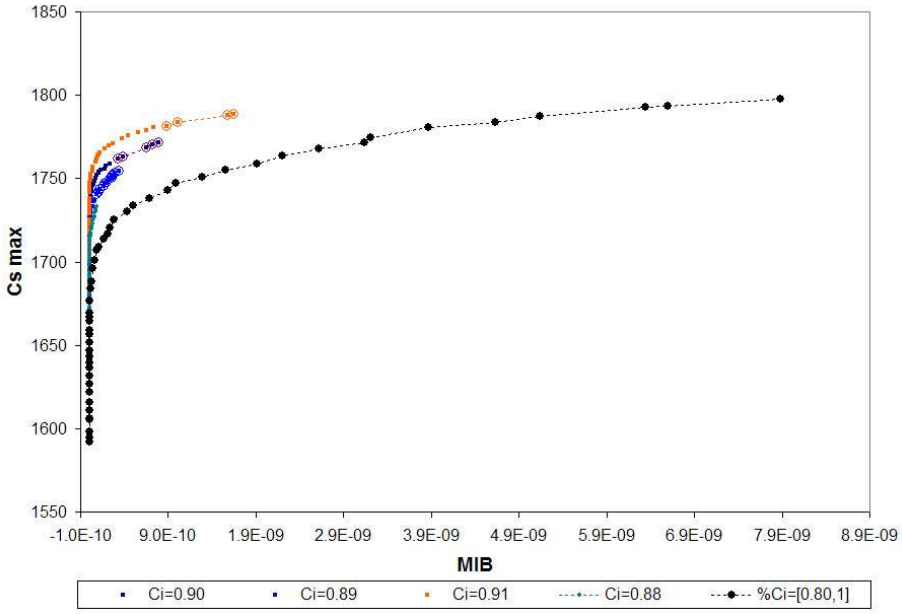


Figure B.26: Fronteras de aproximación de Pareto robustez-coste para centroide no especificado usando representación porcentual. [179, 178]

problema de búsqueda de robustez, los esfuerzos se centran en mejorar la calidad algorítmica, como ocurre con la inserción de la representación porcentual en este problema.

B.5.2 Aplicación de AUREO a un problema de planificación

El problema de planificación estudiado fue abordado por primera vez en [59] y se origina en una planta de tratamiento de residuos. Es un problema de planificación de tipo '*flow-shop*' que consiste en una primera etapa compuesta por un conjunto de $m \geq 1$ silos que pueden ser considerados máquinas paralelas idénticas, en los cuales un grupo de camiones descargan los residuos (1ª operación), y una segunda etapa con un único mezclador (máquina crítica) que procesa los residuos y los prepara para finalmente deshacerse de ellos (2ª operación). No hay capacidad de almacenamiento, de modo que una vez que un silo acepta una carga, éste queda ocupado hasta que la misma es traspasada y procesada totalmente en el mezclador. La secuencia empieza estrictamente en la primera etapa y acaba en la segunda y una vez iniciado un proceso no puede detenerse. Un ejemplo de la secuencia descrita está en la figura 5.6).

Desde el punto de vista de la optimización del procedimiento, se pretende minimizar el máximo tiempo (*Makespan*) para completar el procesamiento. Por otra parte, la entidad de transporte penaliza las esperas, por lo que el tiempo total de espera (TWT) entre la llegada de los camiones y el inicio de la descarga debe ser minimizado. El primer intento de resolución de este problema biobjetivo se planteó en [59] (ver también [133]) empleando una versión del MOSA

(ver fig. B.8) usando instancias reales de llegadas de camiones. Contrariamente a lo esperado, se obtuvieron múltiples secuencias de descarga cuyos valores de objetivo (*Makespan* y TWT) son iguales, y distinguiéndose solamente dos o tres puntos no dominados en el espacio de objetivos con valores muy similares en sus magnitudes.

Claramente, la toma de decisiones ante tal escenario choca con muchas dificultades, puesto que no es posible diferenciar entre una gran cantidad de secuencias de descarga alternativas. Se planteó por tanto estudiar el efecto de la incertidumbre en este problema y ponderar su utilidad para distinguir entre soluciones.

AUREO Etapa 1

Para estudiar la incorporación de incertidumbre en este problema, primero se investigó qué información podía brindar la DM y qué características posee el problema, obteniéndose los siguiente resultados:

- La información disponible acerca del problema es poca y la DM es reacia a dar detalles o hacer suposiciones.
- La planificación en línea (ver [32]) no es posible, por lo que la secuencia de tareas se respetará estrictamente.
- Los camiones pueden llegar con adelanto o retardo respecto al momento de arribo esperado.
- No se consideran atrasos en el procesamiento de operaciones.

Con las consideraciones anteriores puede realizarse el análisis del problema para formular el programa de búsqueda de robustez. Los parámetros del problema son los tiempos de descarga y de mezcla. Las llegadas se supone que ocurrirán según un patrón definido sobre el cual la DM no tiene ingerencia, por lo que éstos son también parámetros del problema. Las decisiones a tomar incluyen fijar el tiempo de inicio de las descargas y de las transferencias y las máquinas que se usarán en la primera etapa. Sea $S(\mathbf{r}, \mathbf{u}, \mathbf{m})$ una secuencia definida a partir del vector de tiempos de arribo de los camiones $\mathbf{r}$, el vector de tiempos de descarga $\mathbf{u}$ y el vector tiempos de mezclado $\mathbf{m}$. Una secuencia es una colección de vectores $(J_j, M_i, t_{j,s})$, tal que la máquina M_i es designada para ejecutar la tareas J_j empezando en el tiempo $t_{j,s}$, donde los índices j, i, s denotan tareas, máquinas y etapas respectivamente. Las siguientes restricciones deben respetarse:

- Las descargas no pueden ejecutarse antes del tiempo de arribo, por tanto $t_{j,1} \geq \mathbf{r}_j$ (en la primera etapa) y $t_{j,2} \geq t_{j,1} + \mathbf{u}_j$ (en la segunda etapa).
- No están permitidas las interrupciones de las tareas y no es posible el almacenaje, de allí que el tiempo para completar una tarea sea $t_{j,1} + \mathbf{u}_j$ en la 1ª etapa y $t_{j,2} + \mathbf{m}_j$ en la 2ª etapa.
- Finalmente, para asegurar un servicio sostenido, el *makespan* no debe sobrepasar las 24 horas.

En presencia de incertidumbre parte de los parámetros se vuelven inciertos, en particular se acepta que el vector $\mathbf{r}$ está sujeto a incertidumbre epistémica. Obsérvese que es posible que los tiempos de arribo puedan ser representados como una variable aleatoria con una PDF asociada. No obstante, considerando la falta de información sobre el hecho, no existen evidencias que sostengan tal premisa. Por el contrario, los vectores $\mathbf{u}$ y $\mathbf{m}$ se consideran libres de incertidumbre. Del mismo modo, el vector de decisión $(J_j, M_i, t_{j,s})$ no posee incertidumbres en la asignación de tareas a máquinas.

Finalmente se considerará la existencia de funciones objetivo inciertas. Dado que se parte de la premisa de que la secuencia física ha de respetarse y solo pueden variar los tiempos asociados a los inicios de las tareas, interesa estudiar cómo variarán los objetivos frente a variaciones del vector $\mathbf{r}$. Para este caso se planteó una simulación Monte Carlo para generar N muestras del vector $\mathbf{r}+\delta_r$, donde δ_r representa la incertidumbre asociada a $\mathbf{r}$. Para todas las muestras del vector de tiempos de arribo (denotado $\delta_{r,k}$) y para cada secuencia física a evaluar, se estiman los objetivos TWT y *Makespan* como:

$$
\begin{aligned}
F_1 &= \frac{1}{N}\sum_{k}^{N} \mathrm{TWT}(S(\mathbf{r}+\delta_{r,k}, \mathbf{u}, \mathbf{m})) \\
F_2 &= \frac{1}{N}\sum_{k}^{N} \mathrm{Makespan}(S(\mathbf{r}+\delta_{r,k}, \mathbf{u}, \mathbf{m}))
\end{aligned}
$$

Según las categorías mostradas en la tabla B.10, el análisis muestra que el problema no puede ser catalogado como Clase 1, dado que no hay información suficiente para propagar la incertidumbre, puesto que ésta es epistémica. Del mismo modo, un programa Clase 2 tampoco es posible debido a que la DM no precisó la definición de las restricciones $I(\cdot)$. Por tanto, el problema debe ser atacado como uno de la Clase 3. En consecuencia, es necesario convertir el problema en uno de Clase 1 o Clase 2; para ello se planteó a la DM aceptar una desviación máxima de 10 minutos alrededor de las horas de arribo esperadas como punto de partida para permitir una propagación de tipo Clase 1. A partir de allí, es posible formular restricciones del tipo $I(\cdot)$ usando el frente de aproximación como hallado con el problema Clase 1 como referencia. A tal fin, se preguntó a la DM acerca del máximo período de tiempo que consideraba no significativo, fijándose este en 5 minutos.

Con la información recopilada y las restricciones originales, se planteó el siguiente programa de búsqueda de robustez de Clase 2: *Encuentre la secuencia S con máxima IB que minimice F_1 y F_2 simultáneamente, tal que $|F_1 - TWT(S)| \leq 2$ minutos y $|F_2 - Makespan(S)| \leq 2$ minutos (restricciones de desempeño) y $\overline{F_2} \leq 24$ horas (restricción de operabilidad).*

AUREO Etapa 2

La herramienta desarrollada en [59] para solucionar el problema original fue modificada para permitir simulaciones a partir de una secuencia específica para variaciones de $\mathbf{r}+\delta_r$. Dado que la heurística consume muchos recursos, la simulación Monte Carlo se limitó a 30 muestras (mínimo recomendado en [165]) por secuencia, con la verificación del tamaño del intervalo de confianza como criterio adicional. Los tiempos de arribo se muestrearon usando una distribución

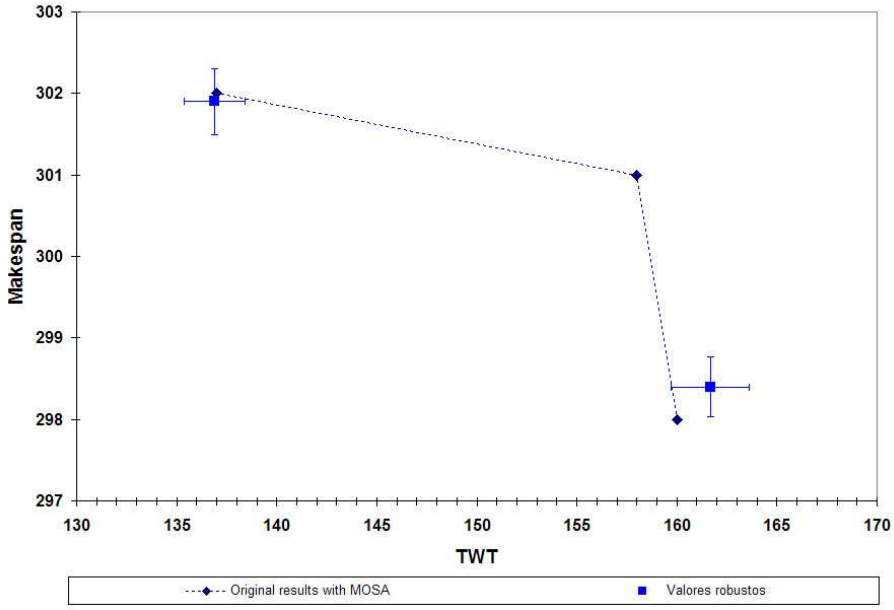


Figure B.27: Resultados filtrando los óptimos en dos etapas (en minutos).

uniforme, mientras que los objetivos TWT y *Makespan* se consideraban iguales si sus valores esperados yacían dentro del mismo hipercubo, luego de dividir el espacio usando el criterio de la ϵ-dominancia aditiva, con ϵ igual a un umbral de indiferencia. De este modo, con cajas de arista 1 minuto, todas las secuencias dentro de la misma caja se consideran iguales, y una desviación más allá de las cajas vecinas inmediatas resulta inaceptable para los criterios $I(\cdot)$.

Se estudiaron dos estrategias de almacenamiento de soluciones. Por una parte, se pueden clasificar las soluciones en términos de sus objetivos y su robustez, lo que equivale a resolver un problem de tres objetivos F_1, F_2 y MIB, mientras que por el otro lado, es posible encontrar el mayor número posible de soluciones que no violen las restricciones y luego aplicar un filtrado en términos del valor de MIB, lo que es semejante a aplicar un procedimiento lexicográfico, donde la optimización simultánea de F_1 y F_2 se realiza primero, y la optimización del MIB en la segunda etapa.

Para el cálculo del MIB se propuso un GA que trabaja con cromosomas $(t_1, t_2, \ldots, t_j, \ldots, t_{|J|})^t$ tal que la incertidumbre $\mathbf{r} + \delta_r$ es representada como $(r_1 \pm (10 \text{ min} + t_1), \ldots, r_j \pm (10 \text{ min} + t_j), \ldots, r_{|J|} \pm (10' + t_{|J|}))$ y $t_j \geq 0$ min$\forall j$, con MIB $= \prod_j t_j$ como función objetivo. De esta forma se pretende determinar el máximo rango de variacines en los tiempos de arribo que una secuencia puede soportar sin violar las restricciones. Se empleó una población de 10 individuos durante 10 generaciones, reportando como resultado el MIB más alto hallado.

La figura B.27 muestra los resultados obtenidos cuando los candidatos son filtrados en dos etapas. Los resultados originales se presentan como rombos, mientras que el valor esperado de los objetivos para las secuencias óptimas aparecen como cuadrados.

La figura B.28 presenta los resultados de resolver el problema como uno de

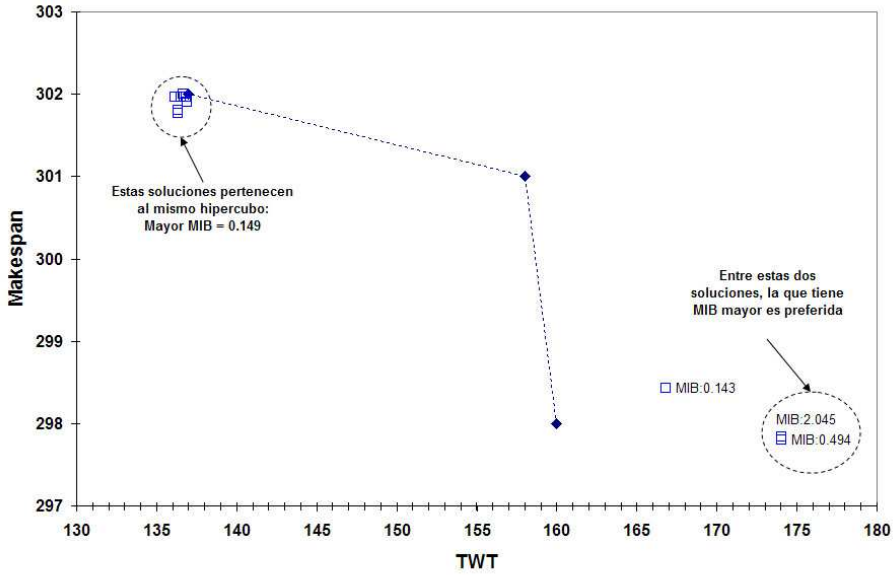


Figure B.28: Resultados resolviendo el problema como un programa de 3 objetivos (en minutos).

tres objetivos. Nótese que la solución abajo a la derecha viola las restricciones de desempeño en TWT, no obstante su representación es útil para ejemplificar cómo dos secuencias pueden tener casi los mismos valores de objetivos, y un valor robustez (MIB) muy diferente al mismo tiempo. Igualmente, el grupo superior izquierdo corresponde a soluciones óptimas que pertenecen al mismo hipercubo. Todas son consideradas no dominadas en un espacio de tres objetivos, no obstante aquella con MIB 0.149 sería la secuencia a elegir si se persigue robustez.

Los resultados muestran los muchos aspectos a considerar en problemas de Clase 3, y la utilidad de AUREO en problemas prácticos reales.

B.5.3 Aplicación de AUREO a problemas de análisis de vulnerabilidad

La última aplicación considerada en esta tesis corresponde al uso de AUREO para analizar y enfrentar la incertidumbre en un problema de análisis de vulnerabilidad. En particular, se pretendía desarrollar una metodología para analizar la vulnerabilidad en redes que pueden propagar un ataque realizado en algún nodo al resto de ellos.

Formalmente el problema se expresa en términos de una red genérica $G(N, E)$ compuesta por un conjunto $N = \{n\}$ de nodos conectados mediante arcos $E = \{e_{i,j}\}$, cada uno de los cuales relaciona dos nodos cualesquiera n_i y n_j en forma unidireccional o bidireccional [189]. A partir de aquí, se propone una identificación entre nodos con grupos de entes o personas susceptibles de ser atacados, y los arcos con canales de propagación de tal ataque [233, 232], como

sucede por ejemplo, en las redes de distribución de agua o de Internet.

Considérese como variables de decisión el conjunto $P^c = \{p_i^c\}$ de medidas defensivas que pueden ser implementadas, cada una con diferente coste $c(p_i^c)$ y capacidad de protección. Un defensor intentará minimizar el impacto de un supuesto ataque, sujeto a sus disponibilidad de recursos R_D. Paralelamente, el atacante dispone de P_N alternativas de ataque, sobre las cuales intentará maximizar el impacto, sujeto al total de recursos disponibles R_A.

Con los elementos anteriores, el problema del defensor se puede formular como *encontrar la asignación de protecciones* $\{(n, P^c(n))\}$ *de* P^c *a los nodos de* N *tal que se minimice el impacto en la red, s.a.* $\sum_{n \in N} \sum_{p_i^c \in P^c(n)} c(p_i^c) \leq R_D$. Naturalmente el patrón de ataques es desconocido, aunque sea posible conocer alguna información sobre las preferencias del atacante o su caudal de recursos R_A mediante labores de inteligencia.

En el modelo anterior falta especificar la manera de estimar el impacto. Algunos autores [109, 120, 121, 122] minimizar una función de utilidad, que típicamente corresponde al impacto o daño esperado. En cambio en [233, 232], los autores proponen dos índices de impacto:

Tiempo para alcanzar todos los nodos de destino (TTRAD) es el tiempo para afectar a todos los nodos de la red. Para este índice, valores bajos indican impactos altos.

Promedio de personas afectadas (ANPA) o promedio de entes afectados (ANEA), es la velocidad promedio de personas o entes afectados durante la propagación. Es proporcional al área debajo de la curva acumulada de personas afectadas en función del tiempo. Para este índice, valores altos implican impactos altos.

Los índices anteriores pueden ser empleados en expresiones de análisis de riesgo, donde la función de riesgo típicamente toma la forma $R = p \times q$, donde p denota la probabilidad de un evento y q sus consecuencias. Por tanto, para reducir el riesgo R de ataque, el defensor debe encontrar el sistema de protección que minimice el riesgo total. Dado que se deben considerar al menos $|P_N|$ escenarios en esa tarea, el riesgo esperado de implementar las protecciones $\{(n, P^c(n))\}$ puede ser estimado como $\sum_{a \in P_N} p(a)q(a)$, siendo a el escenario de ataque, $p(a)$ su probabilidad y $q(a)$ sus consecuencias (cf. [5]).

Problemas analizados

Se estudiaron dos redes, suponiendo en ambos casos que los recursos R_D y R_A estaban limitados a un punto de ataque y uno de defensa. Como primer problema se consideró la Red 1, compuesta de 52 nodos y 73 arcos bidireccionales, cuya topología corresponde a una red telefónica real [128] (ver fig. 6.3). El segundo caso corresponde a la Red 2, compuesta de 332 nodos y 2126 arcos bidireccionales, cuya estructura está tomada de la red aeroportuaria de Estados Unidos de América (http://vlado.fmf.uni-lj.si/pub/networks/pajek/data/gphs.htm) (ver fig. 6.4). No se consideró ningún otro elemento relativo a las redes reales de las cuales fue tomada la topología.

Los tiempos de transición o propagación para ambas redes se supuso uniforme U(0,10). El número de entes afectables en cada nodo de la Red 1 se supone

uniformemente distribuido según límites mostrados en la tabla 6.1. Para la Red 2 en todos los nodos se supuso una distribución U(10,40).

Para calcular las medidas de impacto ANPA y TTRAD, se empleó un simulador basado en Autómatas Celulares y simulación Monte Carlo, cuyos detalles se describen en el capítulo 6. Para efectos de AUREO, dicho simulador es un sistema de caja negra.

AUREO Etapa 1

En la etapa 1 de AUREO se pretende identificar las fuentes de incertidumbre, pero en el caso bajo estudio es necesario también analizar las posibles expresiones funcionales de impacto, dado que no existe un modelo único o universal.

Para los problemas descritos anteriormente, la variable de decisión se convierte en el nodo a defender, de allí que $\mathbf{x} = (x_1 = i)$ donde $i = \{1, 2, \ldots, |N|\}$ es el índice del nodo protegido n_i. El nodo de ataque, por otra parte, es un parámetro sometido a incertidumbre epistémica, mientras que los tiempos de propagación a través de los arcos y el número de personas en cada nodo, son parámetros que se aceptan como sujetos a incertidumbre aleatoria. De lo anterior, el vector de parámetro $\mathbf{p}$ se compone de $1 + |N| + |E|$ cantidades.

Considérese ahora el modelo de impacto, que debe ser coherente con el objetivo de minimizar riesgo. Entre las opciones posibles se encuentran:

$$F(\mathbf{x}, \mathbf{p}) = (f_1 = \text{ANPA}(\mathbf{x}, \mathbf{p}), f_2 = \text{TTRAD}(\mathbf{x}, \mathbf{p}))^t \tag{B.49}$$

$$F(\mathbf{x}, \mathbf{p}) = p(n_i) \times (w_1 I_{\text{ANPA}} + w_2 I_{\text{TTRAD}}), \sum_j w_j = 1 \tag{B.50}$$

$$F(\mathbf{x}, \mathbf{p}) = p(n_i) \times (I_{\text{ANPA}}, I_{\text{TTRAD}})^t \tag{B.51}$$

donde la ec. B.49 correspondería al caso de estimar el riesgo directamente con ANPA y TTRAD. En este caso ANPA debe ser minimizado mientras que TTRAD maximizado. Nótese también que como el punto de ataque es desconocido, la optimización debe hacerse sobre todos los escenarios posibles. Este modelo no depende de información probabilística, como en las ec. B.50 y B.51 donde se requiere conocer la probabilidad de ser atacado para cada nodo.

La ec. B.50 corresponde a una combinación lineal de índices de riesgo asociados a ANPA y a TTRAD. A tal fin, el TTRAD es expresado como un índice donde a menores tiempos de propagación el riesgo es mayor, adoptándose la forma $I_{\text{TTRAD}} = \max_N\{\text{TTRAD}\} - \text{TTRAD}$, donde el máximo es calculado respecto al TTRAD cuando se ataca un nodo n_i respecto al conjunto N. Por otro lado, si se quiere evitar la agregación de objetivos, se puede recurrir a un modelo multiobjetivo, como en la ec. B.51.

El nivel de incertidumbre respecto a las probabilidades de ataque $p(n_i)$ varía entre $[0 \leq \underline{p}(n_i), \overline{p}(n_i) \leq 1]$. Se desprende que, cuando la incertidumbre es máxima, el riesgo varía entre 0 y el impacto medido en términos de TTRAD o ANPA. En consecuencia, un modelo de Clase 1 correspondería a una comparación de intervalos, según algún criterio de entre los mencionados en la sección B.4.1. Entre ellos, el criterio del peor caso fue seleccionado como el más adecuado, dado que es empleado para propósitos defensivos [189] y porque una de las premisas del análisis considera la existencia de "una inteligencia malévola dirigida a producir la *máxima* desintegración social"[4, pg. 361] (las cursivas son nuestras). Luego, obsérvese que en referencia al modelo, adoptar el peor

caso en la ec. B.51 equivale a resolver la ec. B.49 usando comparaciones del peor caso si $[0 = \underline{p}(n_i), \overline{p}(n_i) = 1]$. Ese es el caso estudiado aquí ($p(n_i) \in [0, 1]$).

Mediante el análisis de robustez basado en las definiciones 17 a la 19 de B. Roy, se llegó a la misma conclusión de que el programa consiste en minimizar el máximo impacto.

Por último, dado que la evaluación de ANPA y TTRAD se realiza con un simulador Monte Carlo, la función objetivo es de naturaleza incierta, y los intervalos de confianza deben considerarse durante la jerarquización.

Con los elementos analizados, solo resta dividir el vector $\mathbf{p}$ en dos vectores, $\mathbf{p_a}$ y $\mathbf{p_G}$, que convenientemente representan el punto de ataque y los parámetros de las redes. Formalizando, el programa matemático consiste en hallar el nodo n_i a proteger ($\mathbf{x} = (i)$, $i = \{1, 2, \ldots, |N|\}$) tal que:

para ec. B.49:

$$\mathbf{x} = \arg\{\min_{\mathbf{x}} \max_{\mathbf{p_a}} \text{ANPA}(\mathbf{x}, \mathbf{p_a}, \mathbf{p_G}) \wedge \max_{\mathbf{x}} \min_{\mathbf{p_a}} \text{TTRAD}(\mathbf{x}, \mathbf{p_a}, \mathbf{p_G})\}$$

para ec. B.51:

$$\mathbf{x} = \arg\min_{\mathbf{x}} \max_{\mathbf{p_a}} \left(I_{\text{ANPA}}(\mathbf{x}, \mathbf{p_a}, \mathbf{p_G}), I_{\text{TTRAD}}(\mathbf{x}, \mathbf{p_a}, \mathbf{p_G})\right)^t$$

B.5.4 AUREO Etapa 2

El programa anterior implica dos tareas, por una parte se debe encontrar la frontera de Pareto para una protección particular, sobre el conjunto total de alternativas de ataque (máximo impacto), luego se debe intentar minimizar el impacto máximo hallando la frontera de menor impacto. A tal fin, se consideraron las siguientes opciones:

1. Implementar una estructura de lazos anidados, donde el lazo externo realiza una enumeración completa de nodos a proteger, mientras que el interno realiza una enumeración completa de los puntos a atacar.

2. El lazo externo realiza una enumeración completa mientras que el interno emplea MOEA.

3. El MOEA explora todo el dominio de soluciones.

Las tres alternativas fueron estudiadas, considerando además la no duplicidad de cromosomas. No fue posible resolver el problema directamente con MOEA, por lo que la estrategia de jerarquización usada queda abierta para seguir siendo estudiada en el futuro. Las dos estrategias iniciales por su parte, dieron buenos resultados, aunque claramente la primera solo es aceptable en redes de pequeñas dimensiones.

La Red 1 tiene solo 52 nodos, por lo que el uso de MOEA no está justificado, siendo más eficiente el realizar una enumeración completa de casos. La fig. B.29 muestra los resultados obtenidos, donde destacan los nodos 18, 22 y 24 como mejores soluciones. Obsérvese que no se trata de comparar puntos sino fronteras, de allí que surja una complicación adicional para decidir cuál es el nodo cuya protección maximiza la robustez. Se hizo un análisis mediante el ϵ-indicador, concluyéndose que los nodos 22 y 24 minimizan el máximo impacto, pero no fue posible decidir más allá, por lo que corresponde a la DM escoger entre esas dos alternativas de protección empleando algún criterio adicional.

Un resultado similar se obtiene al estudiar la Red 2, aunque la dimensionalidad de la misma sugiere el uso de MOEA. No obstante, se realizó la enumeración

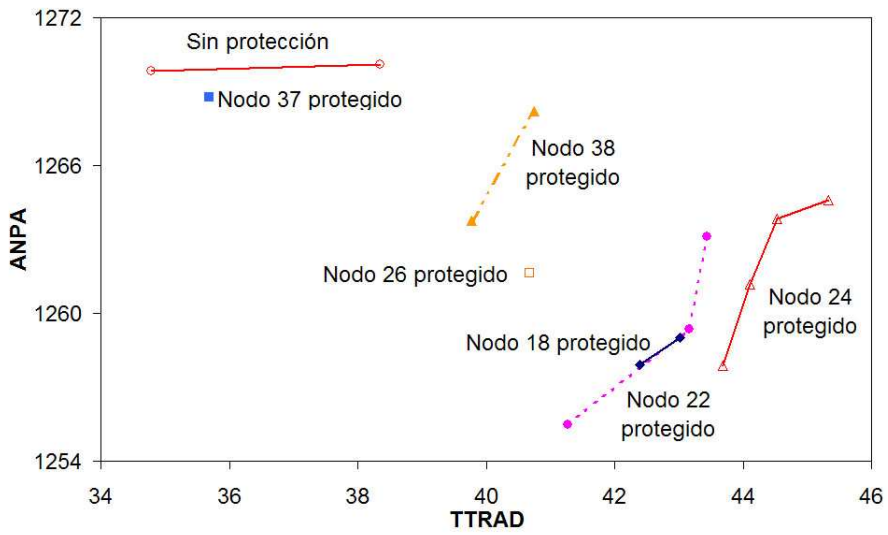


Figure B.29: Selección de fronteras efficientes para el atacante en función del nodo protegido para la Red 1 [161].

completa con el fin de obtener datos adicionales sobre el problema, dado que la metodología de evaluación de vulnerabilidad se encuentra en fase de desarrollo.

Por razones de seguridad, no se muestran las etiquetas de los nodos en el grafo de la Red 2 (fig. 6.4) La figura B.30 muestra el máximo impacto obtenible al atacar los nodos cuyas etiquetas se indican. Nótese que la totalidad de frentes está dominada por el nodo 8, de aquí que el uso del ϵ-indicador permita identificar dicho óptimo sin problema.

El análisis de algunos puntos no dominados revela la posibilidad de crear islas, es decir sub-redes aisladas, al proteger algunos nodos específicos, llamados nodos de corte [189]. Esto puede notarse también mediante inspección visual. La interpretación de la presencia de dichos nodos en términos de protección ante ataques está abierta para investigación futura. No obstante, a medida que el concepto de impacto evolucione, AUREO deberá aplicarse nuevamente para guiar la búsqueda de la protección más robusta.

B.6 Conclusiones y trabajo futuro

En la primera parte de esta tesis se ha presentado una revisión y análisis de los orígenes, fuentes y tipos de incertidumbre, junto con un resumen de los marcos teóricos empleados para representarla. Posteriormente se ha brindado una descripción de los elementos principales que componen a los MOEA y el estado del arte en manejo de incertidumbre en EMO. Con base en los elementos mencionados, se ha ofrecido un análisis de los efectos de la incertidumbre en toma de decisiones, desde el punto de vista del funcionamiento algorítmico de los MOEA. Específicamente se han estudiado los métodos de jerarquización y de almacenamiento de soluciones, cuando las últimas están caracterizadas por incertidumbre de tipo no probabilística, probabilística, posibilística y de

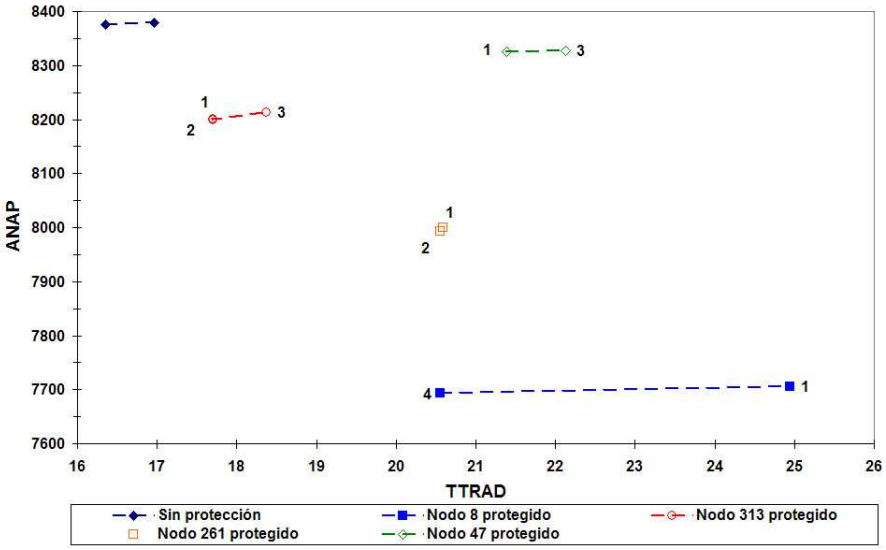


Figure B.30: Selección de fronteras efficientes para el atacante en función del nodo protegido para la Red 2 [159].

lógica difusa. A nuestro entender, dicho análisis constituye una contribución importante en la materia, pues los esfuerzos previos son escasos y se concentran en incertidumbre probabilística.

Entre las propuestas algorítmicas presentadas en esta tesis, se encuentra un algoritmo multiobjetivo para manejar incertidumbre no probabilística, llamado IP-MOEA. El mismo ha demostrado que es posible construir algoritmos eficientes para manejar ese tipo de incertidumbre. Por otro lado, también se propuso el tratamiento de incertidumbre no probabilística mediante la ϵ-dominancia, la cual puede ser empleada en conexión con el criterio de mínimo arrepentimiento. El análisis mostró que, bajo ciertas supuestos, el uso del mínimo arrepentimiento para comparar conjuntos es equivalente a aplicar el ϵ-indicador, lo que sugiere la relevancia de continuar estudiando estos conceptos en el futuro.

Paralelamente se ha propuesto un procedimiento para diseñar rutinas de almacenamiento, basado en el *hypergrid*. En el mismo, se sugiere una interpretación probabilística del concepto de densidad de hipercubo, que puede emplearse para trabajar con vectores objetivo inciertos, sujetos a incertidumbre aleatoria.

Un esfuerzo importante se dedicó a la exploración teórica y práctica del concepto de robustez, como herramienta para manejar incertidumbre. A partir de una revisión del estado del arte en optimización robusta y diseño robusto, se he propuesto una taxonomía de clases que permite sistematizar la idea de incertidumbre, desde la perspectiva de la información disponible para plantear y resolver problemas de decisión. Adicionalmente se encontró que el concepto de robustez no parece haber sido estudiado más allá de los límites de la incertidumbre probabilística y no probabilística, por lo que se han brindado algunas pistas para extender este concepto a casos posibilísticos, de lógica difusa o de

probabilidades imprecisas.

La primera parte de la tesis concluye con la propuesta de un marco metodológico para el *análisis de incertidumbre y robustez en Optimización Evolutiva* o AUREO. Esta metodología constituye la contribución más importante de esta tesis, respecto al uso de MOEA en toma de decisiones multicriterio. AUREO conjuga los aspectos teóricos y prácticos del manejo de incertidumbre con elementos algorítmicos de MOEA, permitiendo 1) formular los problemas de toma de decisiones como programas de búsqueda de robustez, y b) resolver dichos programas con ayuda de MOEA. Esta metodología es útil tanto para adaptar MOEA existentes para como para el diseño de nuevos algoritmos.

La aplicación de AUREO a tres problemas de seguridad de funcionamiento es abordada en la segunda parte de la tesis. Se estudió un problema de fiabilidad en el cual, tras usar AUREO, se aumentó significativamente la eficiencia del MOEA mediante el uso de la *representación porcentual*. Por otro lado, se estudió un problema de planificación en el cual, mediante el uso de AUREO, se redefinió tanto el programa matemático como el método heurístico empleado, a fin de producir soluciones robustas. Finalmente se aplicó AUREO al desarrollo de una metodología de análisis de vulnerabilidad, permitiendo formular el problema de optimizar la protección como un programa de robustez. Tras el análisis de dos ejemplos construidos a partir de redes reales, se obtuvieron resultados interesantes en materia de análisis de riesgos y se identificaron nuevos retos en diseño de MOEA. Los resultados obtenidos son de interés, no solamente para el área de optimización evolutiva multiobjetivo, sino también para la el área de seguridad de funcionamiento.

Los desarrollos futuros fueron claramente indicados a lo largo de la tesis. Por un parte, el concepto de densidad probabilística constituye una nueva perspectiva para el diseño de MOEA que operen bajo incertidumbre. Por otro lado, la utilidad del ϵ-indicador para manejar incertidumbre no probabilística debe ser estudiada más a fondo, pues podría permitir la construcción de un nuevo algoritmo para casos no probabilísticos, en los que la DM estuviese interesada en aplicar el criterio de mínimo arrepentimiento. De forma similar, la ϵ-dominancia también podría proveer de métodos para construir métricas que permitan comparar frente de aproximación con incertidumbre.

Entre los aspectos algorítmicos estudiados, quedan algunos no resueltos que deben ser abordados en el futuro. Por ejemplo, el asunto de la eficiencia en optimización multiobjetivo min-max, en el análisis de vulnerabilidad usando MOEA, no ha sido resuelto. Del mismo modo, un tema abierto en EMO y que afecta al problema mencionado es el de comparar frentes de aproximación. El análisis de incertidumbre plantea una dificultad particular pues el procedimiento min-max compara fronteras, y no siempre es posible identificar inequívocamente la mejor frontera, por lo que pueden obtenerse resultados no concluyentes.

En cuanto al diseño de MOEA, la taxonomía de problemas de búsqueda de robustez introducida en relación a AUREO, sugiere que los esfuerzos futuros pueden enfocarse en diseñar algoritmos eficientes y flexibles para cada una de las clases de problemas, en lugar de intentar diseñar algoritmos de propósito general como es habitual en EMO.

Finalmente, queda abierta la cuestión de cómo interpretar la idea de robustez en entornos posibilistas, difusos o probabilistas imprecisos.

Bibliography

[1] J. K. Allen, C. Seepersad, H. Choi, and F. Mistree. Robust Design for Multiscale and Multidisciplinary Applications. *Journal of Mechanical Design*, 128(4):832–843, July 2006.

[2] J. D. Andrews and T. R. Moss. *Reliability & Risk Assessment*. Longman Group United Kingdom, 1993.

[3] B. Aouni and O. Kettani. Goal programming model: A glorious history and a promising future. *European Journal of Operation Research*, 133(2):225–231, 2001.

[4] G. E. Apostolakis and D. M. Lemon. Screening methodology for the identification and ranking of infrastructures vulnerability due to terrorism. *Risk Anal*, 25(2):361–376, 2005.

[5] T. Aven. A unified framework for risk and vulnerability analysis covering both safety and security. *Reliab. Eng. Sys. Saf*, 92(6):745–754, 2007.

[6] B. M. Ayyub. From disecting ignorance to solving algebraic problems. *Reliab. Eng. Sys. Saf*, 85(1–3):223–238, July-September 2004.

[7] B. M. Ayyub and G. J. Klir. *Uncertainty Modelling and Analysis in Engineering and the Sciences*. Chapman & Hall/CRC, Boca Raton, FL, USA, 2006.

[8] T. Bäck. *Evolutionary Algorithms in THeory and Practice: Evolution Strategies, Evolutionary Programming, Genetic Algorithms*. Oxford Univ. Press, 1996.

[9] T. Bäck, D. Fogel, and Z. Michalewicz. *Handbook of Evolutionary Computation*. Oxford Univ. Press, 1997.

[10] C. Barrico and C. H. Antunes. Robustness Analysis in Multi-Objective Optimization Using a Degree of Robustness Concept. In *Proceedings of the 2006 IEEE Congress on Evolutionary Computation (CEC 2006)*, pages 1887–1892, Vancouver, BC, Canada, July 2006. IEEE Service Center.

[11] M. Basseur and E. Zitzler. Handling Uncertainty in Indicator-Based Multiobjective Optimization. *International Journal of Computational Intelligence Research*, 2(3):255–272, 2006.

[12] V. Batagelj and A. Mrvar. Pajek. http://vlado.fmf.uni-lj.si/pub/networks/pajek/, 2007.

[13] Y. Ben-Haim. Uncertainty, probability and information-gaps. *Reliab. Eng. Sys. Saf*, 85(1–3):249–266, July-September 2004.

[14] Y. Ben-Haim. Info-gap Decision Theory For Engineering Design. Or: Why 'Good' is Prefereble to 'Best'. In E. Nikolaides, D. Ghiocel, and S. Singhal, editors, *Engineering Design Reliability Handbook*. CRC Press, 2005.

[15] B. Besharati, L. Luo, S. Azarm, and P. K. Kannan. Multi-Objective Single Product Robust Optimization: An Integrated Design and Marketing Approach. *Journal of Mechanical Design*, 128(4):884–892, July 2006.

[16] H.-G. Beyer and B. Sendhoff. Robust opmitization - A comprehensive survey. *Comput. Methods Appl. Mech. Engrg*, 196:3190–3218, 2007.

[17] H.-G. Beyer and B. Sendhoff. Robust optimization - A comprehensive survey. *Comput. Methods Appl. Mech. Engrg*, 196:3190–3218, 2007.

[18] V. M. Bier, A. Nagaraj, and V. Abhichandani. Protection of simple series and parallel systems with components of different values. *Reliab. Eng. Sys. Saf*, 87(3):315–323, March 2005. doi:10.1016/j.ress.2004.06.003.

[19] J. Birchmeier. Systematic assessment of the degree of criticality of infrastructures. In T. Aven and J. E. Vinnem, editors, *Risk, Reliability and Societal Safety*, volume 1: Specialisation Topics, 2007.

[20] S. C. Brailsford, C. N. Potts, and C. N. Potts. Constraint satisfaction problem: Algorithms and applications. *European Journal of Operation Research*, 119(3):557–581, December 1999.

[21] A. Bronevich. The Maximal Variance of Fuzzy Interval. In *Proceedings of the Third International Symposium on Imprecise Probabilities and Their Applications ISIPTA'03*, Lugano, Switzerland, July 14-17 2003. Society for Imprecise Probability: Theories and Applications (SIPTA). Electronic proceedings: http://www.carleton-scientific.com/isipta/2003-toc.html.

[22] M. Bruns. Propagation of imprecise probabilities through black-box models. Master's thesis, Georgia Institute of Technology, May 2006.

[23] M. Bruns, C. Paredis, and S. Ferson. Computational methods for decision making based on imprecise information, 2006. NSF Workshop on Reliable Engineering Computing. Savannah, GA.

[24] O. C. C. García-Martínez and F. Herrera. A taxonomy and an empirical analysis of multiple objective ant colony optimization algorithms for the bi-criteria tsp. *European Journal of Operation Research*, 180(1):116–148, july 2007.

[25] C. Carlsson and R. Fullér. On possibilistic mean value and variance of fuzzy numbers. *Fuzzy Sets and Systems*, 122:315–326, 2001.

[26] C. Coello Coello and G. Toscano Pulido. Multiobjective optimization using a micro-genetic algorithm. In L. S. et al., editor, *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO'2001)*, 2001.

[27] C. A. Coello Coello. A comprehensive survey of evolutionary-based multiobjective optimization techniques. *Knowledge and Information Systems*, 1(3):269–308, August 1999.

[28] C. A. Coello Coello. Evolutionary Multi-Objective Optimization: A Historical View of the Field. *IEEE Computational Intelligence Magazine*, pages 28–36, February 2006.

[29] D. W. Coit, T. Jin, and N. Wattanapongsakorn. System Optimization With Component Reliability Estimation Uncertainty: A Multi-Criteria Application. *IEEE Transactions on Reliability*, 53(3):369–380, 2004.

[30] C. Colbourn. *The Combinatorics of Network Reliability.* Oxford University Press, 1987.

[31] D. Corne, K. Deb, P. Fleming, and J. Knowles. The good of the many outweights the good of the one: Evolutionary multiobjective optimization. *coNNectionS*, 1(1):9–13, 2003. ISSN 1543–4281.

[32] A. Davenport and J. Beck. A survey of techniques for scheduling with uncertainty. http://www.eil.utoronto.ca/profiles/chris/chris.papers.html.

[33] K. Deb. Multi-objective genetic algorithms: Problems difficulties and construction of test problems. *Evolutionary Computation*, 7(3):205–230, Fall 1999.

[34] K. Deb and H. Gupta. Searching for robust Pareto-optimal solutions in multi-objective optimization. In *Evolutionary Multi-Criterion Optimization*, number 3410 in LNCS, pages 150–164, Berlin, Germany, 2005. Springer-Verlag.

[35] K. Deb, D. Padmanabhan, S. Gupta, and A. K. Mall. Handling Uncertainties Through Reliability-Based Optimization Using Evolutionary Algorithms. KanGAL Report 2006009, Kanpur Genetic Algorithms Laboratory, 2006.

[36] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan. A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation*, 6(2):182–197, 2002.

[37] K. Deb, L. Thiele, M. Laumanns, and E. Zitzler. Scalable multi-objective optimization test problems. In *IEEE Congress on Evolutionary Computation (CEC 2002)*, volume 1, pages 825–830, Honolulu, HI, USA, May 2002. IEEE Service Center.

[38] K. Dehnad, editor. *Quality Control, Robust Design, and the Taguchi Method.* Wadsworth and Books, Pacific Grove, 1989.

[39] A. P. Dempster. Upper and lower probabilities induced by a multi-valued mapping. *Annals of Mathemical Statistics*, (38):325–339, 1967.

[40] W. Dong and F. Wong. Fuzzy weighted averages and implementation of the extension principle. *Fuzzy Sets and Systems*, 21(2):183–199, 1987.

[41] D. Dubois. Possibility theory and statistical reasoning. *Comp. Stat. Data Anal*, 51:47–69, 2006.

[42] D. Dubois and H. Prade. The mean value of a fuzzy interval. *Fuzzy Sets and Systems*, 24(3):279–300, December 1987.

[43] D. Dubois, H. Prade, and S. Sandri. On possibility/probability transformations. In R. Lowen and M. Roubens, editors, *Fuzzy Logic: State of the Art*, pages 103–112. Kluwer Academic, Dordrecht, 1993.

[44] J. S. Dyer, P. C. Fishburn, R. E. Steruer, J. Wallenius, and S. Zionts. Multiple Criteria Decision Making, Multiattibute Utility Theory: The Next Ten Years. *Management Science*, 38(5):645–654, May 1992.

[45] M. Emmerich, N. Beume, and B. Naujoks. An EMO Algorithm Using the Hypervolume Measure as Selection Criterion. In C. A. Coello Coello, A. Hernández Aguirre, and E. Zitzler, editors, *Evolutionary Multi-Criterion Optimization. Third International Conference, EMO 2005, Guanajuato, Mexico, March 9-11, 2005. Proceedings*, volume 3410 of *Lecture Notes in Computer Science*, pages 62–76, Germany, 2005. Springer-Verlag.

[46] S. Ferson, L. Ginzburg, V. Kreinovich, and J. Lopez. Absolute Bounds on the Mean of Sum, Product, etc.: A Probabilistic Extension of Interval Arithmetic. In *Extended Abstracts of the 2002 SIAM Workshop on Validated Computing*, pages 70–72, Toronto, Canada, May 23-25 2002.

[47] S. Ferson and L. R. Ginzburg. Different methods are needed to propagate ignorance and variability. *Reliab. Eng. Sys. Saf*, 54, 1996.

[48] S. Ferson, J. Hajagos, D. Berleant, J. Zhang, W. T. Tucker, L. Ginzburg, and W. Oberkampf. Dependence in dempster-shafer theory and probability bounds analysis. Technical report, Sandia National Laboratories, 2004.

[49] S. Ferson and J. G. Hajagos. Arithmetic with uncertain numbers: rigorous and (often) best possible answers. *Reliab. Eng. Sys. Saf*, 85(1–3):135–152, July-September 2004.

[50] S. Ferson, V. Kreinovich, L. Ginzburg, D. S. Myers, and K. Sentz. Constructing probability boxes and dempster-shafer structures. Technical report, Sandia National Laboratories, 2002.

[51] J. E. Fieldsend and R. M. Everson. Multi-objective Optimisation in the Presence of Uncertainty. In *Proceedings of the 2005 IEEE Congress on Evolutionary Computation (CEC 2005)*, volume 1, pages 243–250, Edinburgh, Scotland, September 2005. IEEE Service Center.

[52] P. Fleming, R. C. Purshouse, and R. J. Lygoe. Many-Objective Optimization: An Engineering Design Perspective. In C. A. Coello Coello, A. Hernández Aguirre, and E. Zitzler, editors, *Evolutionary Multi-Criterion Optimization. Third International Conference, EMO 2005, Guanajuato, Mexico, March 9-11, 2005. Proceedings*, volume 3410 of *Lecture Notes in Computer Science*, pages 14–32, Germany, 2005. Springer-Verlag.

[53] C. M. Fonseca and P. J. Fleming. Genetic Algorithms for Multiobjective Optimization: Formulation, Discussion and Generalization. In S. Forrest, editor, *Proceedings of the Fifth International Conference on Genetic Algorithms*, pages 416–423, San Mateo, California, USA, 1993. University of Illinois at Urbana-Champaign, Morgan Kauffman Publishers.

[54] C. M. Fonseca, V. Grunert da Fonseca, and L. Paquete. Exploring the performance of stochastic multiobjective optimisers with the second-order attainment function. In C. A. Coello Coello, A. Hernández Aguirre, and E. Zitzler, editors, *Evolutionary Multi-Criterion Optimization. Third International Conference, EMO 2005, Guanajuato, Mexico, March 9-11, 2005. Proceedings*, volume 3410 of *Lecture Notes in Computer Science*, pages 250–264, Germany, 2005. Springer-Verlag.

[55] C. M. Fonseca, J. Knowles, L. Thiele, and E. Zitzler. A tutorial on the performance assessment of stochastic multiobjective optimizers. http://dbkgroup.org/knowles/emo-tutorial-2up.pdf, 2005. Invited talk at the Evolutionary Multi-Criterion optimization Conference (EMO 2005).

[56] C. M. Fonseca, L. Paquete, and M. L.-I. nez. An Improved Dimension-Sweep Algorithm for the Hypervolume Indicator. In *IEEE Congress on Evolutionary Computation*, pages 1157–1163, Sheraton Vancouver Wall Center Hotel, Vancouver, BC, Canada, July 16-21 2006.

[57] P. Fortemps and M. Roubens. Ranking and defuzzification methods based on area compensation. *Fuzzy Sets and Systems*, 82:319–330, 1996.

[58] B. Galván, G. Winter, D. Greiner, and D. Salazar. New evolutionary methodologies for integrated safety system design and maintenance optimization. In G. Levitin, editor, *Computational Intelligence in Reliability Engineering: Evolutionary Techniques in Reliability Analysis and Optimization*, volume 39. Springer-Verlag, 2007.

[59] X. Gandibleux, J. Jorge, S. Martínez, and N. Sauer. Premier Retour d'Experience sur le Flow-Shop Biobjectif et Hybride a Deux Etages avec une Contrainte de Blocage Particuliere. In 6e *Conférence Francophone de MOdélisation et SIMulation (MOSIM'06)*, Rabat, Maroc, April 2006.

[60] X. Gandibleux, J. Jorge, D. Salazar Aponte, and S. Martinez. Solving methods and robustness methodology for a two stage bi-objective hybrid flowshop problem with a particular blocking constraint, November 2006. http://www.lina.sciences.univ-nantes.fr/Publications/2006/GJAM06.

[61] B. J. Garrick. Perspective on the Use of Risk Assessment to Address Terrorism. *Risk Anal*, 22(3):421–423, 2002.

[62] B. J. Garrick, J. E. Hall, M. Kilger, J. C. McDonald, T. O'Toole, P. S. Probst, E. Rinskopf Parker, R. Rosenthal, A. W. Trivelpiece, L. A. Van Arsdale, and E. Zebroski. Confronting the risks of terrorism: making the right decisions. *Reliab. Eng. Sys. Saf*, 86:129–176, 2004.

[63] G. Gerla and L. Scarpati. Extension principles for fuzzy set theory. *Journal of Information Science*, 106:49–69, 1998.

[64] D. D. Goldberg. *Genetics algorithms in search, optimization & machine learning*. Addison-Wesley Publishing Company, Inc, USA, 1989.

[65] S. Graves. Probabilistic dominance criteria for comparing uncertain alternatives: A tutorial. *Omega*, 2007. doi:10.1016/j.omega.2007.03.001.

[66] D. Greiner, J. Emperador, B. Galván, and G. Winter. Robust design of frames under uncertainty loads by multiobjective genetic algorithms. In B. Topping, G. Montero, and R. Montenegro, editors, *Proceedings of the Eighth International Conference on Computational Structures Technology*. Civil-Comp Press, 2006.

[67] J. N. Gupta and E. F. Stafford. Flowshop scheduling research after five decades. *European Journal of Operation Research*, 169:699–711, 2006.

[68] Y. Y. Haimes. On the Definition of Vulnerabilities in Measuring Risks to Infrastructures. *Risk Anal*, 26(2):293–296, 2006.

[69] Y. Y. Haimes and T. Longstaff. The Role of Risk Analysis in the Protection of Critical Infrastructures Against Terrorism. *Risk Anal*, 22(3):439–444, 2002.

[70] E. Hansen and G. W. Walster. *Global Optimization Using Interval Analysis*, volume 264 of *Pure and Applied Mathematics*. CRC Press, second edition, 2003.

[71] M. P. Hansen and A. Jaszkiewicz. Evaluating the quality of approximations to the non-dominated set. Technical Report IMM-REP-1998-7, Technical University of Denmark, March 1998.

[72] A. J. Hatfield and K. W. Hipel. Understanding and managing uncertainty and information. In *Proceedings of the IEEE International Conference on systems, Man, and Cybernetics, 1999*, volume 5, pages 1007–1012, Tokyo, Japan, 1999. ISBN:0-7803-5731-0.

[73] K. Hausken. Protecting infrastructures from strategic attackers. In T. Aven and J. E. Vinnem, editors, *Risk, Reliability and Societal Safety*, volume 1: Specialisation Topics, pages 881–887. Taylor & Francis, 2007. ISBN: 978-0-415-44783-6.

[74] B. Hayes. A Lucid Interval. *American Scientist*, 91(6):484–488, November-December 2003.

[75] T. Hayward and J. Preston. Chaos theory, economics and information: the implications for strategic decision-making. *Journal of Information Science*, 25(3):173–182, 1999.

[76] E. M. Hendrix, C. J. Mecking, and T. H. Hendriks. Finding robust solutions for product design problems. *European Journal of Operation Research*, 92:28–36, 1996.

[77] F. Hernandes, M. T. Lamata, J. L. Verdagay, and A. Yamakami. The shortest path problem on networks with fuzzy parameters. *Fuzzy Sets and Systems*, 158(14):1561–1570, 2007.

[78] A. G. Hernández-Díaz, L. V. Santana-Quintero, C. Coello Coello, R. Caballero, and J. Molina. A new proposal for multi-objective optimization using differential evolution and rough sets theory. In M. K. et al., editor, *Genetic and Evolutionary Computation Conference (GECCO'2006)*, volume 1, 2006.

[79] A. Herreros López. *Diseño de controladores robustos multiobjetivo por medio de algoritmos genéticos.* PhD thesis, Universidad de Valladolid, 2000.

[80] R. Hites. *Robustness and Preferences in Combinatorial Optimization.* PhD thesis, Université Libre de Bruxelles, 2005.

[81] R. Hites, Y. De Smet, N. Risse, M. Salazar-Neumann, and P. Vincke. About the aplicability of MCDA to some robustness problems. *European Journal of Operation Research*, 174(1):322–332, October 2005.

[82] Å. Holgrem, E. Jenelius, and J. Westin. Evaluating Strategies for Defending Electric Power Networks Against Antagonistic Attacks. *IEEE Transactions on Power Systems*, 22(1):76–84, February 2007.

[83] Å. J. Holgrem. Using Graph Models to Analyze the Vulnerability of Electric Power Networks. *Risk Anal*, 26(4):955–968, 2006.

[84] J. Holland. *Adaptation in Natural and Artificial Systems.* The University of Michigan Press, Ann Arbor, 1975.

[85] H. Hoogeveen. Multicriteria scheduling. *European Journal of Operation Research*, 167:592–623, 2005.

[86] O. Hryniewicz. Fuzzy sets in the evaluation of reliability. volume 40. Springer-Verlag, 2007.

[87] H.-Z. Huang, X. Tong, and M. J. Zuo. Posbist reliability theory for coherent systems. volume 40. Springer-Verlag, 2007.

[88] E. J. Hughes. Multi-Objective Probabilistic Selection Evolutionary Algorithm. Technical Report DAPS/EJH/56/2000, Department of Aerospace, Power, & Sensors, Cranfield University, September 2000.

[89] E. J. Hughes. Constraint Handling With Uncertain and Noisy Multi-Objective Evolution. In *Proceedings of the Congress on Evolutionary Computation 2001 (CEC 2001)*, volume 2, pages 963–970, Piscataway, New Jersey, May 2001. IEEE Service Center.

[90] N. N. Jeffrey Horn and D. E. Goldberg. A Niched Pareto Genetic Algorithm for Multiobjective Optimization. In *Proceedings of the First IEEE Conference on Evolutionary Computation IEEE World Congress on Computational Intelligence*, 1994.

[91] M. T. Jensen. *Robust and Flexible Scheduling with Evolutionary Computation.* PhD thesis, University of Aarhus, Denmark, October 2001.

[92] Y. Jin and J. Branke. Evolutionary optimization in uncertainty environments - a survey. *IEEE Transactions on Evolutionary Computation*, 9(3):303–317, 2005.

[93] Y. Jin and B. Sendhoff. Trade-off between optimality and robustness: An evolutionary multiobjective approach. In *Evolutionary Multi-Criterion Optimization*, number 2632 in LNCS, pages 237–251, Berlin, Germany, 2003. Springer-Verlag.

[94] C. W. Johnson. Understanding the interaction between public policy, managerial decision-making and the engineering of critical infrastructures. *Reliab. Eng. Sys. Saf*, 92(9):1141–1154, September 2007.

[95] A.-L. Jousselme, P. Maupin, and Éloi Bossé. Uncertainty in a situation analysis perspective. In *Proceedings of the Sixth International Conference of Information Fusion, 2003*, 2003.

[96] S. Kaplan. The Words of Risk Analysis. *Risk Anal*, 17(4):407–417, 1997.

[97] S. Kaplan and B. J. Garrick. On the Quantitative Definition of Risk. *Risk Anal*, 1(1):11–27, 1981.

[98] A. Kaufmann and M. Gupta. *Introduction to Fuzzy Arithmetic*. Van Nostrand Reinhold, New York, 1985.

[99] R. B. Kearfott. Interval Computations: Introduction, Uses, and Resources. *Euromath Bulletin*, 2(1):95–112, 1996.

[100] R. Keeney and H. Raiffa. *Decisions with Multiple Objectives: Preferences and Value Trade-offs*. Wiley, 1976.

[101] G. Klir and Y. Pan. Constrained fuzzy arithmetic: Basic questions and some answers. 2, 1998.

[102] G. Klir and B. Yuan. *Fuzzy Sets and Fuzzy Logic: Theory and Applications*. Prentice Hall, Upper Saddle River, NJ, USA, 1995.

[103] G. J. Klir. Principles of uncertainty: What are they? Why do we need them? *Fuzzy Sets and Systems*, 74(1):15–31, 1995.

[104] G. J. Klir. Generalized information theory: aims, results, and open problems. *Reliab. Eng. Sys. Saf*, 85(1):12–38, 2004.

[105] G. J. Klir and T. A. Folger. *Fuzzy Sets, Uncertainty, and Information*. Prentice-Hall International, USA, 1988.

[106] J. Knowles and D. Corne. Properties of an adaptive archiving algorithm for storing nondominated vectors. *IEEE Transactions on Evolutionary Computation*, 7(2):100–116, April 2003.

[107] J. D. Knowles and D. W. Corne. Approximating the Nondominated Front Using the Pareto Archived Evolution Strategy. *Evolutionary Computation*, 8(2):149–172, 2000.

[108] J. D. Knowles, D. W. Corne, and M. J. Oates. On the Assessment of Multi-objective Approaches to the Adaptive Distributed Database Management Problem. In M. Schoenauer, K. Deb, G. Rudolph, X. Yao, E. Lutton, J. J. Merelo, and H.-P. Schwefel, editors, *Proceedings of the Sixth International Conference on Parallel Problem Solving from Nature (PPSN VI)*, pages 869–878, Berlin, Germany, 2000. Springer.

[109] E. Korczak, G. Levitin, and H. B. Haim. Survivability of seriesŰparallel systems with multilevel protection. *Reliab. Eng. Sys. Saf*, 90(1):45–54, October 2005.

[110] J. R. Koza. *Genetic Programming: On the Programming of Computer by Means of Natural Selection*. MIT Press, USA, 1992.

[111] V. Kreinovich, J. Beck, C. Ferregut, A. Sanchez, G. R. Keller, M. Averill, and S. A. Starks. Monte-carlo-type techniques for processing interval uncertainty, and their engineering applications. In *Proceedings of the Workshop on Reliable Engineering Computing*, pages 139–160, Savannah, Georgia, September 15-17 2004.

[112] V. Kreinovich, G. Xiang, and S. Ferson. Computing mean and variance under Dempster-Shafer uncertainty: Towards faster algorithms. *International Journal of Approximate Reasoning*, 42(3):212–227, August 2006.

[113] R. Kröger. On the variance of fuzzy random variables. *Fuzzy Sets and Systems*, 92(1):83–93, November 1997.

[114] H. Kunreuther. Risk Analysis and Risk Management in an Uncertain World. *Risk Anal*, 22(4):655–664, 2002.

[115] M. Laumanns. *Analysis and Applications of Evolutionary Multiobjective Optimization Algorithms*. PhD thesis, Swiss Federal Institute of Technology, 2003.

[116] M. Laumanns, L. Thiele, K. Deb, and E. Zitzler. On the Convergence and Diversity-Preservation Properties of Multi-Objective Evolutionary Algorithms. TIK-Report 108, Computer Engineering and Networks Laboratory (TIK), Swiss Federal Institute of Technology (ETH), Gloriastrasse 35, CH-8092 Zurich, Switzerland, May 2001.

[117] M. Laumanns, L. Thiele, K. Deb, and E. Zitzler. Combining Convergence and Diversity in Evolutionary Multi-objective Optimization. *Evolutionary Computation*, 10(3):263–282, Fall 2002.

[118] M. Laumanns, L. Thiele, E. Zitzler, and K. Deb. Archiving with Guaranteed Convergence and Diversity in Multi-Objective Optimization. In E. C.-P. et al, editor, *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO'2002)*, pages 439–447, San Francisco, California, USA, July 2002. Morgan Kaufmann Publishers.

[119] M. Laumanns, E. Zitzler, and L. Thiele. On the Effects of Archiving, Elitism, and Density Based Selection in Evolutionary Multi-objective Optimization. In E. Zitzler, K. Deb, L. Thiele, C. A. Coello Coello, and

D. Corne, editors, *First International Conference on Evolutionary Multi-Criterion Optimization*, volume 1993 of *Lecture Notes in Computer Science*, pages 181–196. Springer-Verlag, 2001.

[120] G. Levitin. Optimal Defense Strategy Against Intentional Attacks. *IEEE Transactions on Reliability*, 56(1):148–157, March 2007.

[121] G. Levitin and H. Ben-Haim. Importance of protections against intentional attacks. *Reliab. Eng. Sys. Saf*, 2007. doi:10.1016/j.ress.2007.03.016.

[122] G. Levitin and K. Hausken. Protection vs. redundancy in homogeneous parallel systems. *Reliab. Eng. Sys. Saf*, 2007. doi:10.1016/j.ress.2007.10.007.

[123] M. Li, S. Azarm, and A. Boyars. A New Deterministic Approach Using Sensitivity Region Measures for Multi-Objective Robust and Feasibility Robust Design Optimization. *Journal of Mechanical Design*, 128(4):874–883, July 2006.

[124] D. Lim, Y.-S. Ong, Y. Jin, B. Sendhoff, and B. S. Lee. Inverse Multi-objective Robust Evolutionary Design. *Genetic Programming and Evolvable Machines Journal*.

[125] P. Limbourg. Multi-objective Optimization of Problems with Epistemic Uncertainty. In C. A. Coello Coello, A. Hernández Aguirre, and E. Zitzler, editors, *Evolutionary Multi-Criterion Optimization. Third International Conference, EMO 2005, Guanajuato, Mexico, March 9-11, 2005. Proceedings*, volume 3410 of *Lecture Notes in Computer Science*, pages 413–427, Germany, 2005. Springer-Verlag.

[126] P. Limbourg and D. E. Salazar Aponte. An Optimization Algorithm for Imprecise Multi-Objective Problem Functions. In *Proceedings of the 2005 IEEE Congress on Evolutionary Computation (CEC 2005)*, volume 1, pages 459–466, Edinburgh, Scotland, September 2005. IEEE Service Center.

[127] S.-H. Lossin. Decision Making with Imprecise and Fuzzy Probabilities - a Comparison. In *ISIPTA'05 Electronic Proceedings. Proceedings of the 4th International Symposium on Imprecise Probabilities and Their Applications*, Pittsburg, Pennsylvania, July 20-23 2005.

[128] E. Manzi, M. Labbé, G. Latouche, and F. Maffioli. Fishman's sampling plan for computing network reliability. *IEEE Transactions on Reliability*, 50(1):41–46, March 2001.

[129] R. Marler and J. Arora. Survey of multi-objective optimization methods for engineering. *Struct Multidisc Optim*, 26:369–395, 2004. DOI 10.1007/s00158-003-0368-6.

[130] M. Marseguerra and E. Zio. *Basics of the Monte Carlo Method with Application to System Reliability*. LiLoLe- Verlag GmbH (Publ. Co. Ltd.), 2002.

[131] M. Marseguerra, E. Zio, and F. Bosi. Direct Monte Carlo availability assessment of a nuclear safety system with time-dependent failure characteristics. In *Proceedings of the Third International Conference on Mathematical Methods in Reliability MMR*, Trondheim, Norway, 2002.

[132] M. Marseguerra, E. Zio, L. Podofillini, and D. W. Coit. Optimal Design of Reliable Network Systems in Presence of Uncertainty. *IEEE Transactions on Reliability*, 54(2):243–253, 2005.

[133] S. Martínez. *Ordonnancement de Systèmes de Production avec contraintes de blocage*. PhD thesis, University of Nantes, France, 2005.

[134] S. Martorell, S. Carlos, J. Villanueva, A. Sanchez, B. Galvan, D. Salazar, and M. Cepin. Use of multiple objective evolutionary algorithms in optimizing surveillance requirements. *Reliab. Eng. Sys. Saf*, 91(9):1027–1038, 2006.

[135] S. Martorell, B. Galván, and D. Salazar. Utilización de algoritmos genéticos para resolver problemas de optimización en ingeniería de seguridad. Tutorial del VI Congreso de Confiabilidad, Madrid, 2004.

[136] S. Martorell, A. Sanchez, and S. Carlos. A tolerance interval based approach to address uncertainty for rams+c optimization. *Reliab. Eng. Sys. Saf*, 92(4):408–422, April 2007.

[137] R. Menon, L. H. Tong, L. Zhijie, and Y. Ibrahim. Robust desgin of a spindle motor: a case study. *Reliab. Eng. Sys. Saf*, 75:313–319, 2002.

[138] Z. Michalewicz. *Genetic Algorithms + Data Structures = Evolution Programs*. Springer-Verlag, 1994.

[139] M. Milanese and M. Taragna. h_∞ set membership identification: A survey. *Automatica*, 41:2019–2032, 2005.

[140] M. Milanese and A. Vicino. Optimal estimation theory for dynamic systems with set membership uncertainty: An overview. *Automatica*, 27(6):997–1009, 1991.

[141] R. E. Moore. *Methods and Applications of Interval Analysis*. SIAM Studies in Applied Mathematics. Society for Industrial Mathematics, 1987.

[142] M. G. Morgan and M. Henrion. *Uncertainty: A Guide to Dealing with Uncertainty in Qualitative Risk and Policiy Analysis*. Cambridge University Press, USA, 1990.

[143] Z. P. Mourelatos and J. Liang. A Methodology for Trading-Off Performance and Robustness Under Uncertainty. *Journal of Mechanical Design*, 128(4):856–863, July 2006.

[144] B. Naujoks, N. Beume, and M. Emmerich. Multi-objective optimization using s-metric selection: Application to three-dimensional solution spaces. In *Proceedings of the 2005 IEEE Congress on Evolutionary Computation (CEC 2005)*, volume 2, pages 1282–1289, Edinburgh, Scotland, September 2005. IEEE Press.

[145] W. L. Oberkampf, J. C. Helton, C. A. Joslyn, S. F. Wojtkiewicz, and S. Ferson. Challenge problems: uncertainty in system response given uncertain parameters. *Reliab. Eng. Sys. Saf*, 85(1–3):11–19, July-September 2004.

[146] A. O'Hagan and J. E. Oakley. Probability is perfect, but we can't elicit it perfectly. *Reliab. Eng. Sys. Saf*, 85(1–3):239–248, July-September 2004.

[147] Y.-S. Ong, P. B. Nair, and K. Y. Lum. Max-Min Surrogate-Assisted Evolutionary Algorithm for Robust Design. *IEEE Transactions on Evolutionary Computation*, 10(4):392–404, August 2006.

[148] I. Paenke, J. Branke, and Y. Jin. Efficient Search for Robust Solutions by Means of Evolutionary Algorithms and Fitness Approximations. *IEEE Transactions on Evolutionary Computation*, 10(4):405–420, August 2006.

[149] B. Rai, N. Singh, and M. Ahmed. Robust design of an interior hard trim to improve occupant safety in a vehicle crash. *Reliab. Eng. Sys. Saf*, 89:296–304, 2005.

[150] U. K. Rakowsky. Fundamentals of the Dempster-Shafer theory and its applications to reliability modeling. *International Journal of Reliability, Quality and Safety Engineering*, 14(6):579–601, 2007.

[151] H. Ratschek and J. Rokne. *Computer Methods for the Range of Functions.* Ellis Horwood Limited, England, 1984.

[152] T. Ray. Constrained robust optimal design using a multiobjectiveevolutionary algorithm. In *IEEE Congress on Evolutionary Computation (CEC 2002)*, volume 1, pages 419–424, Honolulu, HI, USA, May 2002. IEEE Service Center.

[153] M. Reyes Sierra and C. A. Coello Coello. Multi-objective particle swarm optimizers: A survey of the state-of-the-art. *International Journal of Computational Intelligence Research*, 2(3):287–308, 2006.

[154] T. Robic and B. Filipic. Demo: Differential evolution for multiobjective optimization. In C. A. Coello Coello, A. Hernández Aguirre, and E. Zitzler, editors, *Evolutionary Multi-Criterion Optimization. Third International Conference, EMO 2005*, volume 3410 of *Lecture Notes in Computer Science*, pages 520–533. Springer, 2005.

[155] C. M. Rocco, J. A. Moreno, and N. Carrasquero. Robust design using a hybrid-cellular-evolutionary and interval-arithmetic approach: a reliability application. *Reliab. Eng. Sys. Saf*, 79(2):149–159, February 2003.

[156] C. M. Rocco S. A hybrid approach based on evolutionary strategies and interval arithmetic to perform robust designs. In F. Rothlauf, J. Branke, S. Cagnoni, D. W. Corne, R. Drechsler, Y. Jin, P. Machado, E. Marchiori, J. Romero, G. D. Smith, and G. Squillero, editors, *Applications of Evolutionary Computing, EvoWorkshops2005: EvoBIO, EvoCOMNET, EvoHOT, EvoIASP, EvoMUSART, EvoSTOC*, volume 3449 of *LNCS*, pages 623–628, Lausanne, Switzerland, 30 March-1 April 2005. Springer Verlag.

[157] C. M. Rocco S. and J. A. Moreno. Network reliability assessment using a cellular automata approach. *Reliab. Eng. Sys. Saf*, 78(3):289–295, 2002.

[158] C. M. Rocco S. and D. E. Salazar A. A hybrid approach based on evolutionary strategies and interval arithmetic to perform robust designs. In S. Yang, Y.-S. Ong, and Y. Jin, editors, *Evolutionary Computation in Dynamic and Uncertain Environments*, volume 51, chapter 24, pages 543–564. Springer-Verlag, 2007.

[159] C. M. Rocco S., D. E. Salazar A., and E. Zio. Safety assessment of network systems exposed to harmful events based on Soft Computing Techniques: A case study. *International Journal of Performability Engineering*, In Review.

[160] C. M. Rocco S. and E. Zio. Solving advanced network reliability problems by means of cellular automata and Monte Carlo sampling. *Reliab. Eng. Sys. Saf*, 89(2):219–226, 2005.

[161] C. M. Rocco S., E. Zio, and D. E. Salazar A. Multi-objective evolutionary optimisation of the protection of complex networks exposed to terrorist hazard. In T. Aven and J. E. Vinnem, editors, *Risk, Reliability and Societal Safety*, volume 1: Specialisation Topics, pages 899–905. Taylor & Francis, 2007. ISBN: 978-0-415-44783-6.

[162] N. Rolander, J. Rambo, Y. Joshi, J. K. Allen, and F. Mistree. An Approach to Robust Design of Turbulent Convective Systems. *Journal of Mechanical Design*, 128(4):844–855, July 2006.

[163] C. Romero. *Teoría de la decisión multicriterio: Conceptos, técnicas y aplicaciones.* Alianza Editorial, Madrid, Spain, 1993.

[164] J. Rosenhead. Robustness analysis. Newsletter of the European Working Group «Multiple Criteria Decision Aiding», Fall 2002. Series 3, No 6.

[165] S. M. Ross. *Simulation.* Academic Press Inc., 1997.

[166] W. D. Rowe. Understanding uncertainty. *Risk Anal*, 14(5):743–750, 1994.

[167] B. Roy. Des critères multiples en recherche opérationnelle: Pourquoi? In G. Rand, editor, *Operational Research '87*, pages 829–842. North-Holland, Amsterdam, 1988.

[168] B. Roy. Decision-aid and decision-making. In C. A. Bana e Costa, editor, *Readings in Muliple Criteria Decision Aid*, pages 17–35. Springer-Verlag, Berlin, 1990.

[169] B. Roy. Robustesse de quoi et vis-à-vis de quoi mais aussi robustesse pourquoi en aide à la décision? Newsletter of the European Working Group «Multiple Criteria Decision Aiding», Fall 2002. Series 3, No 6.

[170] B. Roy. La robustesse en recherche opérationnelle et aide à la décision : une préoccupation multi-facettes. In B. Roy, M. A. Aloulou, and R. Kalaï, editors, *Robustness in OR-DA. Annales du LAMSADE 7*, pages 209–235. Université Paris-Dauphine, Paris, 2007.

[171] B. Roy and D. Bouyssou. *Aide Multicritère à la Décision: Méthodes et Cas*. Economica, Paris, France, 1993.

[172] D. Salazar, N. Carrasquero, and B. Galván. Exploiting Comparative Studies Using Criteria: Generating Knowledge from an Analyst's Perspective. In C. A. Coello Coello, A. Hernández Aguirre, and E. Zitzler, editors, *Evolutionary Multi-Criterion Optimization. Third International Conference, EMO 2005, Guanajuato, Mexico, March 9-11, 2005. Proceedings*, volume 3410 of *Lecture Notes in Computer Science*, pages 221–234, Germany, 2005. Springer-Verlag.

[173] D. Salazar and P. Limbourg. Optimización multiobjetivo en presencia de incertidumbre. In M. G. Arenas, F. Herrera, M. Lozano, J. J. Merelo, G. Romero, and A. M. S'anchez, editors, *IV Congreso Español sobre Metaheurísticas, Algoritmos Evolutivos y Bioinspirados [MAEB'05]*, volume 2, pages 937–944, Granada, Spain, September 2005.

[174] D. Salazar, C. M. Rocco, and B. Galván. Optimization of constrained multiple-objective reliability problems using evolutionary algorithms. *Reliab. Eng. Sys. Saf*, 91(9):1057–1070, 2006.

[175] D. E. Salazar, S. S. Martorell, and B. J. Galván. Analysis of representation alternatives for a multiple-objective floating bound scheduling problem of a nuclear power plant safety system. In R. Schilling, W. Haase, J. Periaux, H. Baier, and G. Bugeda, editors, *Evolutionary and Deterministic Methods for Design,Optimisation and Control with Applications to Industrial and Societal Problems EUROGEN 2005*, 2005.

[176] D. E. Salazar A. and C. M. Rocco S. Solving advanced multi-objective robust designs by means of multiple objective evolutionary algorithms (MOEA): A reliability application. *Reliab. Eng. Sys. Saf*, 92(6):697–706, 2007.

[177] D. E. Salazar A., C. M. Rocco S., and B. J. Galván G. Robustness analysis: An information-based perspective. Newsletter of the European Working Group «Multiple Criteria Decision Aiding», Fall 2006. Series 3, No 14.

[178] D. E. Salazar A., C. M. Rocco S., and E. Zio. Multiple objective evolutionary optimisation for robust design. In D. Ruan, P. D'hondt, P. Fantoni, M. De Cock, M. Nachtegael, and E. Kerre, editors, *Applied Artificial Intelligence. Proceedings of The 7th Interntional FLINS Conference*, pages 930–937, Genova, Italy, August 29-31 2006. FLINS, World Scientific.

[179] D. E. Salazar A., C. M. Rocco S., and E. Zio. Robust reliability design of a nuclear system by multiple objective evolutionary optimization. *International Journal of Nuclear Knowledge Management*, 2(3):333–345, 2007.

[180] D. Salazar Aponte, X. Gandibleux, J. Jorge, and M. Sevaux. A robust-solution-based methodology to solve multiple-objective problems with uncertainty, August 2006. To appear (MOPGP'06 post-conference volume - LNEMS, Springer).

[181] D. E. Salazar Aponte. Evaluación de métodos evolutivos de optimización multiobjetivo de segunda generación. Master's thesis, Universidad Central de Venezuela, November 2003.

[182] S. Sayin. Robustness analysis: Robustness algorithms for multiple criteria optimization. Newsletter of the European Working Group «Multiple Criteria Decision Aiding», Spring 2005. Series 3, No 11.

[183] K. Sentz and S. Ferson. Combination of evidence in dempster-shafer theory. Technical report, Sandia National Laboratories, 2002.

[184] M. Sevaux and K. Sörensen. Genetic algorithm for robust schedules. In *Proceedings of 8th International Workshop on Project Management and Scheduling, PMS'2002*, pages 330–333, Valencia, Spain, 3-5 April 2002. ISBN 84-921190-5-5.

[185] M. Sevaux and K. Sörensen. Robustness analysis: Optimisation. Newsletter of the European Working Group «Multiple Criteria Decision Aiding», Fall 2004. Series 3, No 10.

[186] G. Shafer. *A Mathematical Theory of Evidence*. Princeton University Press, 1976.

[187] C. Shannon. The mathematical theory of communication. *The Bell System Technical Journal*, 27, 1948.

[188] W. P. Sheridan. Coping with ignorance in the information age. http://www3.sympatico.ca/cypher2/ChapterFive.htm.

[189] D. R. Shier. *Network Reliability and Algebraic Structures*. Oxford University Press, New York, 1991.

[190] M. Sim. *Robust Optimization*. PhD thesis, Massachusetts Institute of Technology, June 2004.

[191] Z. Skolicki, M. Wadda, M. Houck, and T. Arciszewski. Reduction of Physical Threats to Water Distribution Systems. *Journal of Water Resources Planning and Management*, 132(4):221–217, July-August 2006. Special Issue: Drinking Water Systems Security.

[192] P. Smets. What is dempster-shafer's model? In *Advances in the Dempster-Shafer theory of evidence*, pages 5–34. John Wiley & Sons, Inc., New York, NY, USA, 1994.

[193] K. Sörensen. A framework for robust and flexible optimisation using metaheuristics. *4OR: A Quarterly Journal of Operations Research*, 1(4):341–345, 2003.

[194] K. Sörensen. *A framework for robust and flexible optimisation using metaheuristics with applications in supply chain design*. PhD thesis, Universiteit Antwerpen, April 2003.

[195] N. Srinivas and K. Deb. Multiobjective Optimization Using Nondominated Sorting in Genetic Algorithms. *Evolutionary Computation*, 2(3):221–248, 1994.

[196] R. Storn and K. Price. Differential Evolution - A Simple and Efficient Heuristic for Global Optimization over Continuous Spaces. *Journal of Global optimization*, 11:341–359, 1997.

[197] J. Teich. Pareto-Front Exploration with Uncertain Objectives. In E. Zitzler, K. Deb, L. Thiele, C. A. Coello Coello, and D. Corne, editors, *First International Conference on Evolutionary Multi-Criterion Optimization*, volume 1993 of *Lecture Notes in Computer Science*, pages 314–328. Springer-Verlag, 2001.

[198] P. M. Todd. How much information do we need? *European Journal of Operation Research*, 177(3):1317–1332, 2007.

[199] M. C. Troffaes. Decision making under uncertainty using imprecise probabilities. *International Journal of Approximate Reasoning*, 45:17–29, 2007.

[200] S. Tsuitsui and A. Ghosh. Genetic algorithms with a robust solution searching scheme. *IEEE Transactions on Evolutionary Computation*, 1(3):201–208, 1997.

[201] E. Ulungu, J. Teghem, P. Fortemps, and D. Tuyttens. MOSA Method: A Tool for Solving Multiobjective Combinatorial Optimization Problems. *Journal of Multi-Criteria Decision Analysis*, 8:221–236, 1999.

[202] L. V. Utkin and F. P. A. Coolen. Imprecise reliability: An introductory overview. In G. Levitin, editor, *Intelligence in Reliability Engineering: New Metaheuristics, Neural and Fuzzy Techniques in Reliability*, volume 40. Springer-Verlag, 2007.

[203] D. A. Van Veldhuizen. *Multiobjective Evolutionary Algorithms: Classification, analyses, and New Innovations*. PhD thesis, Air Force Institute of Technology, USA, 1999.

[204] J. Villanueva, A. Sanchez, S. Carlos, and S. Martorell. Genetic algorithm-based optimization of testing and maintenance under uncertain unavailability and cost estimation: A survey of strategies for harmonizing evolution and accuracy. In Press.

[205] P. Vincke. About robustness analysis. Newsletter of the European Working Group «Multiple Criteria Decision Aiding», Fall 2003. Series 3, No 8.

[206] M. Wadda, Z. Skolicki, and T. Arciszewski. Generation of Terrorist Scenarios for Water Distribution Systems: An Evolutionary Computation Approach. In A. Woodcock and C. Pommerening, editors, *Proceedings of the workshop "Critical Infrastructure Protection Project"*, pages 162–170. George Mason University, 2004.

[207] P. Walley. *Statistical Reasoning with Imprecise Probabilities*. Chapman and Hall, London, UK, 1991.

[208] P. Walley. Coherent upper and lower previsions. http://ensmain.rug.ac.be/~ipp, 1997.

[209] P. Walley. Towards a unified theory of imprecise probability. *International Journal of Approximate Reasoning*, 24:125–148, 1999.

[210] P. Walley and T. L. Fine. Towards a frequentist theory of upper and lower probability. *The Annals of Statistics*, 10(3):741–761, 1982.

[211] P. Walley, L. Gurrin, and P. Burton. Analysis of clinical data using imprecise prior probabilities. *The Statistician*, 45(4):457–485, 1996.

[212] L. While, L. Bradstreet, L. Barone, and P. Hingston. Heuristics for Optimising the Calculation of Hypervolume for Multi-objective Optimisation Problems. In *Proceedings of the 2005 IEEE Congress on Evolutionary Computation (CEC 2005)*, volume 3, pages 2225–2232, Edinburgh, Scotland, September 2005. IEEE Service Center.

[213] L. While, P. Hingston, L. Barone, and S. Huband. A faster algorithm for calculating hypervolume. *IEEE Transactions on Evolutionary Computation*, 10(1):29–38, February 2006.

[214] Wikipedia. Portal: Artificial intelligence.

[215] R. C. Williamson. *Probabilistic Arithmetic*. PhD thesis, University of Queensland, 1989.

[216] R. C. Williamson and T. Downs. Probabilistic Arithmetic: Numerical Methods for Calculating Convolutions and Dependency Bounds. *International Journal of Approximate Reasoning*, 3:89–158, 1990.

[217] S. Wolfram. Origins of randomness in physical systems. *Physical Review Letters*, 55:449–452, 1985.

[218] G. Xiang. *Fast algorithms for computing statistics under interval uncertainty, with applications to computer science and to electrical and computer engineering*. PhD thesis, Department of Computer Science of The University of Texas at El Paso, Texas, El Paso, USA, December 2007.

[219] F. Xue. *Multi-Objective Differential Evolution: Theory and Applications*. PhD thesis, Rensselaer Polytechnic Institute, Troy, New York, USA, September 2004.

[220] R. Yager. On choosing between fuzzy subsets. *Kybernetics*, 9:151Ű–154, 1980.

[221] R. Yager. A procedure for ordering fuzzy subsets of the unit interval. *Information Science*, 24:143Ű–161, 1981.

[222] R. Yager. On ordered weighted averaging aggregation operators in multicriteria decision making. *IEEE Transactions on Systems, Man and Cybernetics*, 18:183Ű–190, 1988.

[223] R. R. Yager. On general class of fuzzy connectivities. *Fuzzy Sets and Systems*, 4:235Ű–242, 1980.

[224] R. R. Yager and V. Kreinovich. Decision Making Under Interval Probabilities. *International Journal of Approximate Reasoning*, 22(3):195Ű–215, 1999.

[225] Y. Yoo and N. Deo. A comparison algorithm for terminal-pair reliability. *IEEE Transactions on Reliability*, 37(2):210–215, June 1988.

[226] L. A. Zadeh. Fuzzy sets. *Information and Control*, 8:338–353, 1965.

[227] L. A. Zadeh. Generalized theory of uncertainty (GTU)–principal concepts and ideas. *Comp. Stat. Data Anal*, 51(1):15–46, 2006.

[228] M. Zeleny. Compromise Programming. In M. Zeleny and J. Cochrane, editors, *Multiple Criteria Decision Making*, pages 262–301. University of South Carolina Press, Columbia, SC, USA, 1973.

[229] M. Zeleny. A Concept of Compromise Solutions and the Method of the Displaced Ideal. *Computers and Operations Research*, 1(4):479–496, 1974.

[230] M. Zeleny. *Multiple Criteria Decision Making*. McGraw-Hill, New York, USA, 1982.

[231] M. Zeleny. Cognitive equilibrium: A knowledge-based theory of fuzziness and fuzzy sets. *International Journal of General Systems*, 19:359–381, 1991.

[232] E. Zio and C. Rocco. Security assessment in complex networks exposed to terrorist hazard: A methodological approach based on cellular automata and monte carlo simulation. In *International Workshop on Complex Network and Infrastructure Protection CNIP 2006*, Rome, March 28-29 2006.

[233] E. Zio and C. M. Rocco S. Security assessment in complex networks exposed to terrorist hazard: a simulation approach. *International Journal of Critical Infrastructures*, 4(1/2):80–95, 2008.

[234] E. Zio, C. M. Rocco S., D. E. Salazar A., and G. Müller. Complex networks vulnerability: A multiple-objective optimization approach. In *2007 Proc. Ann. Reliability & Maintainability Symp.*, Orlando, Florida, USA, January, 22-25 2007. RAMS.

[235] E. Zitzler. *Evolutionary Algorithms for Multiobjective optimization: Methods and Applications*. PhD thesis, Swiss Federal Institute of Technology, 1999.

[236] E. Zitzler, K. Deb, and L. Thiele. Comparison of multiobjective evolutionary algorithms: Empirical results. *Evolutionary Computation*, 8:173–195, 2000.

[237] E. Zitzler and S. Künzli. Indicator-based Selection in Multiobjective Search. In X. Y. et al, editor, *Parallel Problem Solving from Nature - PPSN VIII*, volume 3242 of *Lecture Notes in Computer Science*, pages 832–842, Birmingham, UK, September 2004.

[238] E. Zitzler, M. Laumanns, and S. Bleuler. A tutorial on evolutionary multiobjective optimization. In Springer-Verlag, editor, *Workshop on Multiple Objective Metaheuristics (MOMH 2002)*, 2004.

[239] E. Zitzler, M. Laumanns, and L. Thiele. SPEA2: Improving the Strength Pareto Evolutionary Algorithm. In J. P. P. P. K. Giannakoglou, D. Tsahalis and T. Fogarty, editors, *EUROGEN 2001, Evolutionary Methods for Design, Optimization and Control with Applications to Industrial Problems*, pages 95–100, Athens, Greece, 2002.

[240] E. Zitzler and L. Thiele. Multiobjective Evolutionary Algorithms: A Comparative Case Study and the Strength Pareto Approach. *IEEE Transactions on Evolutionary Computation*, 3(4):257–271, November 1999.

[241] E. Zitzler, L. Thiele, M. Laumanns, C. M. Fonseca, and V. Grunert da Fonseca. Performance Assessment of Multiobjective Optimizers: An Analysis and Review. *IEEE Transactions on Evolutionary Computation*, 7(2):117–132, April 2003.

VITA

Daniel E. Salazar Aponte was born in Caracas, Venezuela on August 9th, 1977. He joined the Institute for Intelligent Systems and Numerical Applications in Engineering of the Universidad de Las Palmas de Gran Canaria (ULPGC) as a PhD candidate in November 2003. He was appointed as Coordinator and is part of the academic staff of the online *Master in Reliability and Risk Engineering* of this institution (taught in Spanish).

His main research areas include Evolutionary Multiple-Criteria Optimization, Decision-Making, Robustness and Uncertainty Handling and Reliability and Risk Engineering. He has co-edited one book and published more than 20 contributions in scientific international journals, conferences and book chapters.

In the past, he has received a BSc. in Chemical Engineering and a MSc. in Operation Research degrees from Universidad Central de Venezuela (UCV) in 2001 and 2003 respectively. He has also lectured in pre-graduate and post-graduate courses at the Department of Operations Research and Computing of this institution where he retains a close scientific collaboration.