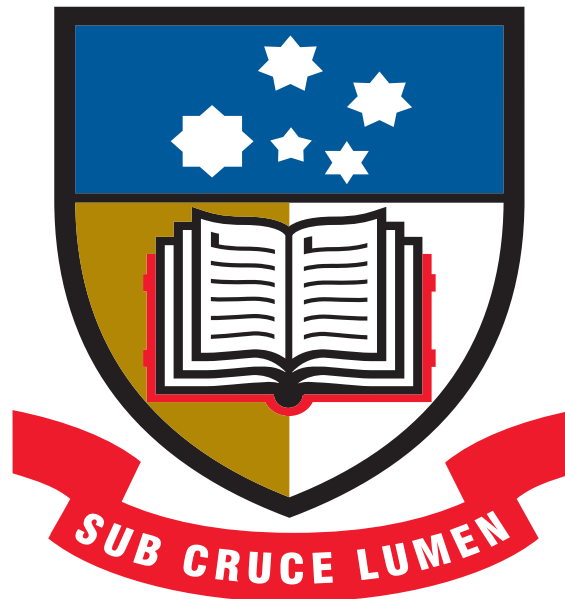


Application of Memory-Based Approach to Multi-Objective Optimisation on Dynamic Resource-Constrained Project Scheduling with Time-varying Number of Tasks

by

Manuel Blanco Abello, MEng. Sc. Elec. Eng.



Submitted to the Graduate Center of the

University of Adelaide,

this thesis is a partial fulfillment for

the prerequisites to the degree of

Doctor of Philosophy

Supervisor: Prof. Zbigniew Michalewicz

Associate Supervisor: Dr. Lam Thu Bui

13th of June 2014

To the loving memory of my mother,

Conсорcia Blanco Abello

and of my father,

Florencio Raz Abello

whose last words to me were:

“Finish your Ph.D.”

... Done it, Dad.

Contents

List of Reserved Symbols	x
List of Tables	xii
List of Figures	xiv
Abstract	xx
Copyright	xxii
Acknowledgements	xxiii
1 Introduction	1
1.1 Motivation	3
1.2 Notations and Definitions	4
1.2.1 Tasks and Resources	4
1.2.2 General Static Problem	5
1.2.3 Dynamic Environment	6
1.2.4 Dynamic Problems	7
1.2.5 Subsets of \mathcal{M}	8

1.2.6	Set of Important Techniques	9
1.3	Goals	9
1.4	Methodology	11
1.4.1	The Effectiveness of McBAR	12
1.4.2	Tri-Objective Problems	13
1.4.3	Resiliency of Techniques	14
1.4.4	Component Legitimation	14
1.4.5	Intelligence of Algorithms	15
1.5	Contributions	16
1.6	Thesis Outline	19
2	Background Knowledge	24
2.1	Evolutionary Algorithms	25
2.1.1	Individual Representation	26
2.1.2	Initial Population	27
2.1.3	Fitness Measure	29
2.1.4	Modifying Operations	29
2.1.4.1	Crossover Operator	30
2.1.4.2	Mutation Operators	30
2.1.5	Selection Criteria	31
2.1.6	Termination Condition	31
2.2	Genetic Programming (GP)	31
2.3	Resource Constrained Project Scheduling	36

2.4	Dynamic Optimisation Problem (DOP)	38
2.4.1	Finding/Tracking Optima	40
2.4.2	Adaptation	41
2.5	Multi-Objective Optimisation (MOO)	45
2.5.1	Hypervolume	47
2.5.2	Set Coverage	48
2.5.3	The Multi-Objective Evolutionary Algorithm (MOEA)	50
2.5.3.1	Non-Dominated Sorting Genetic Algorithm II (NSGA-II)	50
2.5.3.2	Fast Non-Dominated Sorting	51
2.5.3.3	Crowding Distance	53
2.5.3.4	Selection	57
2.6	Estimation Distribution Algorithm	58
2.6.1	Characteristics of EDA	59
2.6.2	EDA in Scheduling	61
2.7	Measure of Risk	62
2.8	Response Surface Methodology (RSM)	63
2.8.1	Design of Experiment (DOE)	65
2.8.2	Model Significance and Adequacy	67
2.8.3	Screening of Factors	68
2.8.4	Model Building	69
2.8.5	Model Prediction	70
2.8.6	Optimal Treatment	71

2.9	Lagrange Optimisation	73
2.10	Measures of Intelligence	76
2.10.1	Reinforcement Learning (RL)	76
2.10.2	Universal Intelligence (UI)	78
3	Literature Review	80
3.1	Genetic Programming (GP)	80
3.2	Resource Constrained Project Scheduling	83
3.3	Dynamic Optimisation Problem (DOP)	84
3.3.1	Explicit Memory-Based Approach	84
3.3.2	Implicit Memory-Based Approach	86
3.3.3	Time-Varying Task Number	87
3.4	Multi-Objective Optimisation (MOO)	88
3.5	Estimation Distribution Algorithm (EDA)	89
3.5.1	EDA in Scheduling and Planning	90
3.5.2	Multi-Objective EDA	90
3.5.3	Dynamic EDA	91
3.6	Robust Proactive Approaches	91
3.6.1	Support for Schedule Robustness	92
3.6.2	Distribution-Based	93
3.6.3	Scenario-Based	94
3.6.4	EA on Robustness	95
3.6.5	Multi-Objective Optimisation on Robustness	95

3.6.6	Robust Reactive-Predictive Approaches	96
3.6.7	Anticipating Risk	97
3.7	DOE and RSM	98
3.7.1	Performance Comparison	99
3.7.2	Search for Optimising Parameters	100
3.7.3	Other Usages of RSM Models	101
3.8	Reinforcement Learning in Scheduling	102
3.9	Multi-Agent System (MAS)	102
3.9.1	Hierarchical Systems	103
3.9.2	Heterarchical Systems	103
3.9.3	Holonic Systems	104
3.9.4	MAS Applications	106
4	Test Environments	107
4.1	Military Mission Planning (MMP)	107
4.1.1	Inputs	111
4.1.2	Objective Functions	113
4.1.3	Constraints	114
4.1.4	Output	114
4.1.5	Dynamic Factors of \mathcal{M}	116
4.2	Problems from \mathcal{L}^2 and \mathcal{L}^3	119
4.2.1	Labels	120
4.2.2	Tables as Functions	125

4.3	Problems from \mathcal{O}^2	127
4.3.1	Tasks	127
4.3.2	Resources	130
4.3.3	Slight and Extreme Changes	130
4.4	Simulations	131
5	Techniques from \mathcal{T}	133
5.1	CBAR	133
5.1.1	Centroid Repair	134
5.1.2	Initial Population	136
5.1.3	Genetic Operators	138
5.1.4	Schedule Formation	139
5.1.5	The Algorithm of CBAR	140
5.1.6	Chosen Schedule	141
5.2	Inapplicability of CBAR	142
5.3	Partial Remedy	143
5.4	Gene-Insertion Side-Effect	145
5.4.1	Sample case	145
5.4.2	Large ID Discontinuity	147
5.4.3	Analytical Investigation	147
5.4.4	Cluster Representation	148
5.4.5	Empirical Investigation	150
5.4.6	Intuitive Explanation	154

5.5	Resolution of Effects	157
5.6	Algorithm of McBAR	160
5.6.1	Computing System	160
5.6.2	First Appearance of New Tasks	161
5.6.3	No New Task in Succeeding Number of Snapshots	162
5.6.4	Other Batches of New Tasks	164
5.6.5	Centroid Repair	165
5.6.6	Initial Population for a Current Static Sub-Problem	166
5.6.7	Summary of Algorithm	167
5.7	MedianBAR	170
5.8	EDA on \mathcal{P}^2 Problem	171
5.9	Techniques	173
6	Effectiveness of McBAR	176
6.1	Averages	176
6.2	Relative Performance Under Task Duration Change	178
6.3	Legitimation	180
6.4	Relative Performance Under Type of Changes	183
6.5	Dynamic Performance	186
6.5.1	Figure Arrangement	187
6.5.2	S_1 -NDLPOP block	191
6.5.3	S_2 -NDLPOP block	195
6.5.4	S_3 -NDLPOP block	196

6.5.5	Other Blocks	198
6.5.6	Other Figures	199
6.6	Conclusions and Future Work	200
7	Performance of McBAR on \mathcal{L}^3	202
7.1	McBAR-Proactive	203
7.2	Models and Averages	204
7.2.1	Convergence Model	204
7.2.2	Difference in Average Makespan and Cost	205
7.3	Optimal Parameters	207
7.4	Relevance of Proactiveness	210
7.5	Difference in Makespan	212
7.6	Difference in Cost	214
7.7	Conclusions and Future Work	215
8	Response Surface Methodology (RSM)	217
8.1	Technique Performance Optimisation	218
8.2	Optimising Parameters	222
8.3	Voxel Models	231
8.3.1	Factors	232
8.3.2	Resolutions, Levels and Ranges	234
8.3.3	Multiple Models	235
8.3.4	Voxels	236
8.3.5	Problems from \mathcal{X}^2	238

8.3.6	Form of the Dynamical Sub-Model	240
8.3.7	Composite Model	241
8.4	Dynamical Model Development	242
8.4.1	Model Building	243
8.4.2	Data Diagnostics	244
8.4.3	Model Predictions	245
8.5	Legitimising the Components of McBAR	252
8.6	Dynamics of Average Performance	254
8.7	Conclusions and Future Work	264
9	Other Facets of the Techniques	266
9.1	Lagrangian Optimisation	266
9.2	Resiliency	270
9.3	Conclusions on Resiliency	274
9.4	Intelligence of Techniques	274
9.5	Measured Intelligence	276
9.6	Conclusions on Technique's Intelligence	278
10	Conclusions	279
10.1	Recalled Hypothesis and Goals	279
10.2	Summary of Methodology and Results	280
10.3	Future Work	283
A	Appendix	285

A.1 Mathematical Formulation of the Static Sub-Problems from Γ_i 285

A.2 Mathematical Formulation of the Dynamic Problems from \mathcal{M} 288

Bibliography **289**

List of Reserved Symbols

N	Total number of tasks
\mathcal{P}^2	Class of bi-objective dynamic RCPS problems, each with increases and without decrease in total number of tasks and has four types of resources
\mathcal{P}^3	Class of tri-objective dynamic RCPS problems, each with increases and without decrease in total number of tasks and has four types of resources
\mathcal{C}^2	Class of bi-objective dynamic RCPS problems, each with fixed total number of tasks
\mathcal{L}^2	Set of problems from \mathcal{P}^2 that have labels of one to 30 listed in Table 4.5
\mathcal{L}^3	Set of tri-objective counterparts of \mathcal{L}^2
\mathcal{M}	Class of multi-objective dynamic RCPS problems, each with increasing total number of tasks
\mathcal{O}^2	Subset of \mathcal{Q}^2
\mathcal{Q}^2	Class of bi-objective dynamic RCPS problems, each with increases and without decrease in the total number of tasks and has two types of resources
\mathcal{X}^2	a subset of \mathcal{O}^2 that is comprised of DOE-designed problems
BBD	Box-Behnken Design
CBAR	Centroid-Based Adaptation with Random Immigrants

CCD	Central Composite Design
EA	Evolutionary Algorithm
EDA	Estimation of Distribution Algorithm
McBAR	Mapping of Task IDs for CBAR
MCS	Monte Carlo Simulation
MOE	Military Operation Environment
MOO	Multi-Objective Optimisation
NSGA-II	Non-Dominated Sorting Genetic Algorithm-II
PNT	Precedence Network of Tasks
RCPS	Resource-Constrained Project Scheduling
RSM	Response Surface Methodology
SOSA	Sequential Order of State Alteration
TNIS	Task Number Increase Sequence
TSC	Type of sequence of changes

List of Tables

4.1	Properties of tasks in MOEs to where the sub-problems of problems from \mathcal{L}^2 and \mathcal{L}^3 are set	112
4.2	Types of environmental changes	121
4.3	Types of sequence of changes (TSC)	121
4.4	Types of TNIS	122
4.5	Problem labels	126
4.6	Properties of tasks in the MOE where the problems from \mathcal{O}^2 are set	129
5.1	Sets of techniques	175
6.1	Average differential set coverage $\mathbf{E}[dSC\delta]$ with (a) $\delta = 3.0$ (b) $\delta = 6.0$	179
6.2	Average execution time	182
6.3	Average $\mathbf{E}_t^r[\bullet]$ performance under change types 0 to 2	184
6.4	Average $\mathbf{E}_t^r[\bullet]$ performance under Change Types 3 to 5	185
6.5	Average $\mathbf{E}_t^r[\bullet]$ performance under Change Type 6	187
8.1	Coefficients of models determined through TPO	223
8.2	Optimal parameters	231
8.3	Range types for various voxels and SOSAs	238

8.4	Pairs of techniques from \mathcal{T} whose $\text{ARP}^{\circ}\text{S}^{\text{b}}\text{S}^{\text{i}}$ s are being modelled	242
8.5	Sample ANOVA	244
8.6	Legitimisation of the sub-algorithms of McBAR	253
9.1	Resiliency	271
9.2	Average Universal Intelligence with respect to δ	277
9.3	Average Universal Intelligence with respect to change type	277

List of Figures

1.1	Contributions, goals and thesis structure relationships	23
2.1	Evolutionary Algorithm	26
2.2	Genotype as an ordered set of integers	27
2.3	Sample Schedule	28
2.4	One-point crossover	30
2.5	Two-point crossover	30
2.6	Mutation of a gene	31
2.7	Sample GP tree	32
2.8	Sample program generation	35
2.9	Original precedence network of tasks	37
2.10	Serial Schedule Generation Scheme (SSGS) Algorithm	38
2.11	Hypersurface of Pareto Front	46
2.12	Hypercube corresponding to objective vector (a) \mathbf{x}_1 (b) \mathbf{x}_2 (d) \mathbf{x}_4 (e) \mathbf{x}_5 . .	48
2.13	Pareto Fronts	52
2.14	Fast Non-Dominated Sorting (FNDS)	54
2.15	First Pareto Front with Cuboid	55

2.16	NSGA-II selection	58
2.17	Sample task ID distribution	60
2.18	Estimation of Distribution Algorithm	61
2.19	Sample confidence interval	71
2.20	Contour graphs	74
2.21	Sample Reinforcement Learning system	77
4.1	Precedence network of tasks for TNIS T_4	115
4.2	Retained schedule for the fifth snapshot	117
4.3	Retained schedule for the sixth snapshot	118
4.4	Precedence network of tasks for TNIS T_3	122
4.5	Precedence network of tasks for TNIS T_5	123
4.6	Precedence network of tasks for TNIS T_6	124
4.7	Precedence network of tasks for TNIS T_7	125
4.8	Precedence network of tasks of problems from \mathcal{O}^2	128
5.1	Algorithm of Centroid-Based Adaptation with Random Immigrants (CBAR)	141
5.2	Sample gene insertion	144
5.3	The SOSAs of the MOE at which genes are inserted	145
5.4	GIBAR degradation	146
5.5	Clusters	148
5.6	Cluster representation in distribution format of cluster/s (a) C_1 (b) C_1 and C_3 (c) C_0 to C_3	149

5.7	Distribution of IDs in genotypes of phenotypes/solutions to sub-problem 1_1^2 obtained by GIBAR	152
5.8	Distribution of IDs in genotypes of phenotypes/solutions to sub-problem 1_{11}^2 obtained by GIBAR	153
5.9	Clusters of solutions to various sub-problems without new tasks	155
5.10	Clusters of solutions to various sub-problems with one sub-problem set in the sixth snapshot with new tasks	155
5.11	Clusters of solutions to various sub-problems with gene-inserted corresponding genotypes	156
5.12	Distribution of mapped IDs in genotypes of phenotypes/solutions to sub-problem 1_{11}^2 obtained by McBAR	159
5.13	Algorithm of McBAR	169
6.1	Dynamics of $\mathbf{E}_{j,k}^\sigma [\bullet]$ with $\delta = 3.0$ with (a) McBAR (b) McBA (c) McBAS as basis techniques	189
6.2	Dynamics of $\mathbf{E}_{j,k}^\sigma [\bullet]$ with $\delta = 6.0$ with (a) McBAR (b) McBA (c) McBAS as basis techniques	192
7.1	Sample evolution dynamics of hypervolume	205
7.2	Crossover and mutation rate as a function of (a) hypervolume (b) convergence rate	209
7.3	Dynamics of average difference in (a) makespan with $\delta = 6.0$ (b) makespan with $\delta = 3.0$ (c) cost with $\delta = 6.0$ (d) cost with $\delta = 3.0$	213
7.4	Starting times	214
8.1	Observed hypervolume of solutions determined by RI under TPO	225

8.2	Observed hypervolume of solutions determined by CBAM under TPO . . .	226
8.3	Observed hypervolume of solutions determined by McBA under TPO . . .	226
8.4	Observed hypervolume of solutions determined by McBAR under TPO . . .	227
8.5	Observed hypervolume of solutions determined by MedianBar under TPO . . .	228
8.6	Observed hypervolume of solutions determined by McBAS under TPO . . .	228
8.7	Observed hypervolume of solutions determined by NDLPOP under TPO . . .	229
8.8	Observed hypervolume of solutions determined by GIBAR under TPO . . .	230
8.9	Performance of EDA/ \mathcal{P}^2 at various values of its parameters	232
8.10	Leverage for the final \mathbf{M}_7 (McBAR, MedianBAR) sub-model	246
8.11	Cooks Distance for the final \mathbf{M}_7 (McBAR, MedianBAR) sub-model	246
8.12	DFFITs for the final \mathbf{M}_7 (McBAR, MedianBAR) sub-model	247
8.13	Normal plot of Studentized Residue for the final \mathbf{M}_7 (McBAR, MedianBAR) sub-model	248
8.14	Predicted versus actual for the final \mathbf{M}_7 (McBAR, MedianBAR) sub-model	249
8.15	Residue versus predicted for the final \mathbf{M}_7 (McBAR, MedianBAR) sub-model	250
8.16	Predictions	251
8.17	SOC average	252
8.18	Dynamics of $\mathbf{E}_{s,t}^{3:6} [dSC (\text{McBAR}, \text{CBAM}, \tau, n, c)]$ under the changes in the number of tasks	257
8.19	Dynamics of $\mathbf{E}_{s,t}^{3:6} [dSC (\text{McBAR}, \text{RI}, \tau, n, c)]$ under the changes in the num- ber of tasks	258
8.20	Dynamics of $\mathbf{E}_{s,t}^{3:6} [dSC (\text{McBAR}, \text{McBA}, \tau, n, c)]$ under the changes in the number of tasks	259

8.21	Dynamics of $\mathbf{E}_{s,t}^{3:6} [dSC (\text{McBAR}, \text{McBAS}, \tau, n, c)]$ under the changes in the number of tasks	259
8.22	Dynamics of $\mathbf{E}_{s,t}^{3:6} [dSC (\text{McBAR}, \text{MedianBAR}, \tau, n, c)]$ under the changes in the number of tasks	260
8.23	Dynamics of $\mathbf{E}_{s,t}^{3:6} [dSC (\text{McBAR}, \text{CBAM}, \rho_1, n, c)]$ under the changes in resource type R_1	261
8.24	Dynamics of $\mathbf{E}_{s,t}^{3:6} [dSC (\text{McBAR}, \text{RI}, \rho_1, n, c)]$ under the changes in resource type R_1	262
8.25	Dynamics of $\mathbf{E}_{s,t}^{3:6} [dSC (\text{McBAR}, \text{McBA}, \rho_1, n, c)]$ under the changes in resource type R_1	262
8.26	Dynamics of $\mathbf{E}_{s,t}^{3:6} [dSC (\text{McBAR}, \text{McBAS}, \rho_1, n, c)]$ under the changes in resource type R_1	263
8.27	Dynamics of $\mathbf{E}_{s,t}^{3:6} [dSC (\text{McBAR}, \text{MedianBAR}, \rho_1, n, c)]$ under the changes in resource type R_1	264

Abstract

Many, if not all, manufacturing processes in industry require scheduling activities; such activities are very important as often they determine the success or failure of some companies. For example, in wine production, grapes are planted, mature fruits are harvested, transported and crushed then the juice obtained is placed in tanks, which are managed during fermentation, and finally the wine is bottle. A schedule can be a solution to a problem which has several, possibly conflicting objectives, e.g. the minimisation of production costs and delays while meeting customer-imposed wine delivery times; the problem also has constraints, e.g. a bottling line cannot be used without being cleaned to process white wine when it last processed red wine. As can be expected, the problem has variables such as the number of wine bottles ordered.

The environment (e.g. wine factory) in which the schedule is implemented may change (e.g. one bottling line breaks down) whereby this schedule becomes infeasible. Consequently, there could be a need to solve a new scheduling problem to obtain a new schedule best suited to the new state of the environment. The number of variables in this new problem may be the same as that of the previous problem. A large proportion of research effort has been directed towards scheduling problems with a constant number of variables despite changes in the environments where the problems are set. However, there are important scheduling problems where the number of variables could vary. For example, in some models of job-shop scheduling problems there are occurrences of additional rush jobs and job cancellations.

This thesis deals with one particular class of scheduling problems, each being multi-objective, resource constrained, and having numbers and values of variables which vary over time. Various traditional operation research methods as well as a few Artificial Intelligence-based techniques, such as Multi-Agent Systems and Evolutionary Algorithms (EA), have been applied to solve this type of problem. In this thesis, a memory-based EA technique was applied to solve problems from the class. Being memory-based, this technique utilises the solutions to problems set in previous states of an environment in

order to solve a problem set in the current state of this environment.

The memory-based EA technique, referred to as *Centroid-Based Adaptation with Random Immigrants* (CBAR), is applicable only to solve multi-objective, resource-constrained problems with a constant number of variables. In this thesis, CBAR is extended to become applicable to solve all problems from the above-mentioned class. The result of this extension is a technique referred to as *Mapping of Task IDs for CBAR* (McBAR).

This thesis investigates the performance, the performance stability over environmental dynamics, and the efficiency of McBAR for solving various problems from the above class, legitimises the sub-algorithms that constitute McBAR and extends McBAR to become proactive (anticipative of future environmental changes).

Compared to the other techniques investigated in this thesis, results showed McBAR to have the best and most stable performance, and to be most efficient for determining solutions to problems from the above class. All of the sub-algorithms of McBAR are shown to be legitimate, while McBAR having been made proactive is shown to be beneficial in some applications.

Copyright

Application of Memory-Based Approach to Multi-Objective Optimisation on Dynamic Resource-Constrained Project Scheduling with Time-varying Number of Tasks

I certify that this work contains no material which has been accepted for the award of any other degree or diploma in my name, in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. In addition, I certify that no part of this work will, in the future, be used in a submission in my name, for any other degree or diploma in any university or other tertiary institution without the prior approval of the University of Adelaide and where applicable, any partner institution responsible for the joint-award of this degree.

I give consent to this copy of my thesis when deposited in the University Library, being made available for loan and photocopying, subject to the provisions of the Copyright Act 1968.

The author acknowledges that copyright of published works contained within this thesis resides with the copyright holder(s) of those works.

I also give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library Search and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

Manuel Blanco Abello
13th of June 2014

Acknowledgements

Although unworthy to state the following, I have to give credence to whom it is due: First and foremost, I would like to thank the triune God, maker of heaven and earth, for all the graces He gave me before and during my study; and to the Blessed Virgin Mary and my dear friend Saint Anthony of Padua¹ for all their intercessions. Sincere gratitude goes to my late parents for their great love and labour for my education.

I never expected to be accepted as a Ph.D. student in Evolutionary Algorithm, considering that my background is in Electrical Engineering. However, Professor Zbigniew Michalewicz supported me to be accepted for Ph.D. student in the Computer Science department. My sincere thanks to him, my supervisor. I also would like to thank Dr. Lam Thu Bui for his great assistance to the publications leading to this theses as my second supervisor. In addition, it is a great honour working with him. Earnest gratitude also goes to Dr. William Winser for his tireless and patient proof-reading of this thesis. Further, thanks to Tracey Young for her kind assistance during my first year of this study. The computing facilities afforded by e-Research South Australia (ERSA) is extremely helpful for the computation of the results presented in this thesis, thanks to this institution as well.

Thanks go to my former lecturers from University of Queensland, Dr. Nicholas Shuley and Dr. David Lovell. Also, thanks go to my former lecturers from the University of San Carlos, Engr. Roger Bajarias, Engr. Tomas Javellana; further, to my lecturers on advanced Mathematics, Dr. Jose T. de Luna, the late Dr. Jose Bernaldez and to my grade school teacher, Sir Labitigan, who awaken my interest on Mathematics. Thanks also to my former classmate, Geneth Remedio, for all her financial assistance.

Several authors who were kind enough to share with me their papers/research results and provided additional explanation when I asked for: André R. Goncalves from School of Electrical and Computer Engineering, University of Campinas, Brazil; Luís Paquete from Faculdade de Economia & CSI Centro de Sistemas Inteligentes, Universidade do

¹The final edition of this thesis was made exactly on the feast day of Saint Anthony.

Algarve, Portugal; Feili Yu from the University of Connecticut; and Gary W. Oehlert from the University of Minnesota.