Coevolving a Computer Player for Resource Allocation Games - Using the game of TEMPO as a test space

by

© **Phillipa Avery, BCompSci(Hons)**

A thesis submitted to the
Graduate Centre
in partial fulfilment of the
requirements for the degree of
Doctor of Philosophy

Supervisor: Prof. Dr. Zbigniew Michalewicz
Associate Supervisor: Dr. Charles Lakos

School of Computer Science
University of Adelaide

September, 2008

Adelaide Australia

For Ben and Rachael.
Your support has gotten me to where I am today.
Thank you.

Contents

Abstract **iv**

Acknowledgements **vii**

List of Tables **viii**

List of Figures **ix**

List of publications **xii**

1 Introduction **1**
1.1 Motivation 2
1.2 Objectives 4
1.3 Methodology 5
1.4 Contributions 6
1.5 Thesis outline 8

2 Background **11**
2.1 Strategic Decision Making 11
2.2 Resource Allocation 13
2.3 Game Theory 14
2.3.1 Game Theory and Early AI 15
2.3.2 The Prisoner's Dilemma 16
2.3.3 Nash Equilibrium 18
2.3.4 Sequential vs. Simultaneous Games 18
2.4 Computers and Games 22
2.4.1 The Evolving Field of AI 23
2.4.2 Commonly Used Methods 24
2.4.3 Expert Systems 25

2.5 Evolutionary Algorithms . . . 26
2.5.1 The General Evolutionary Algorithm . . . 27
2.5.2 Variations of Evolutionary Algorithms . . . 29
2.6 Coevolutionary Algorithms . . . 34
2.7 Fuzzy Logic Systems . . . 37
2.7.1 Fuzzy Logic . . . 37
2.7.2 Fuzzy Control Systems . . . 38

3 The game of TEMPO **40**
3.1 Background of TEMPO . . . 40
3.2 The Game Objectives . . . 41
3.3 How to Play . . . 42

4 Literature Review **47**
4.1 Memory in Coevolutionary Systems . . . 48
4.2 The Creation of Artificial Intelligence Strategies . . . 51
4.3 Evolving Computer Players . . . 53
4.4 Related Games . . . 55
4.5 Adapting to Humans . . . 57

5 The TEMPO Coevolutionary System **60**
5.1 The Early TEMPO Computer Player . . . 60
5.2 Experimental Settings . . . 62
5.2.1 The Representation . . . 62
5.2.2 The Coevolutionary Algorithm Implemented . . . 65

6 Short and Long Term Memory in Coevolution **69**
6.1 Introduction . . . 69
6.2 Experimental Settings . . . 70
6.3 Seeding the Populations . . . 71
6.3.1 Experimenting with the Alien Expert . . . 71
6.4 The Inclusion of a Simple Memory Structure . . . 72
6.5 A Discussion on the Psychology of Human Memory . . . 78
6.6 Implementation of a Short and Long Term Memory Structure . . . 80
6.7 Unique vs. Not Unique Memory Structure . . . 82
6.8 Ranked “Gladiator” Selection . . . 83
6.9 Investigating the Inclusion of Migration . . . 83

6.10 Creating a Dynamic Enemy 86
6.11 Reassessment of the Alien Expert 92
6.12 Experiment Parameter Comparisons 97
6.13 Conclusions and Further Work 98

7 Intelligence and Counter Intelligence 100
7.1 Introduction 100
7.2 A Discussion on the Role of Intelligence 102
7.2.1 Intelligence (INTEL) 102
7.2.2 Counter Intelligence (CI) 103
7.3 Intelligence and Counter Intelligence in the TEMPO Military Planning Game 103
7.3.1 Deficiencies of Intelligence in TEMPO 104
7.4 Applying a more Realistic Mechanism for INTEL and CI 105
7.4.1 Redesigning Offensive and Defensive Intelligence 105
7.4.2 Effects on Counter Intelligence 107
7.4.3 Implementation Methods Used 108
7.5 Experiments and Results 111
7.5.1 Experiments with the new INTEL Mechanism 111
7.5.2 Addition of Counter Intelligence 117
7.6 Conclusions and Future Work 125

8 Adapting to Human game-play 126
8.1 Introduction 126
8.2 Coevolving with Humans 128
8.3 Representing Human Strategies 130
8.4 The Human Adaptive Coevolutionary Process 132
8.5 User Study 135
8.5.1 User study results 136
8.6 Conclusions and future work 138

9 Conclusions and Future Work 141

Bibliography 145

Abstract

decision-making in resource allocation can be a complex and daunting task. Often there exist circumstances where there is no clear optimal path to choose, and instead the decision maker must predict future need and allocate accordingly. The application of resource allocation can be seen in many organizations, from military, to high end commercial and political, and even individuals living their daily life. We define resource allocation as follows: the allocation of owner's assets to further the particular cause of the owner.

We propose two ways that computers can assist with the task of resource allocation. Firstly they can provide decision support mechanisms, with alternate strategies for the allocations that might not have been previously considered. Secondly, they can provide training mechanisms to challenge human decision makers in learning better resource allocation strategies. In this research we focus on the latter, and provide the following general hypothesis: *Coevolutionary algorithms are an effective mechanism for the creation of a computer player for strategic decision-making games*.

To address this hypothesis, we present a system that uses coevolution to learn new strategies for the resource allocation game of TEMPO. The game of TEMPO provides a perfect test bed for this research, as it abstracts real-world military resource allocation, and was developed for training Department of Defence personnel. The environment created allows players to practice their strategic decision-making skills, providing an opportunity to analyse and improve their technique. To be truly effective in this task, the computer player the human plays against must be continuously challenging, so the human can steadily improve. In our research the computer player is represented as a fuzzy logic rule base, which allows us investigation into the strategies being created. This provides insight into the ways the coevolution addresses strategic decision-making.

Importantly, TEMPO also gives us an abstraction of another component of strategic decision-making that is not directly available in other games – that of intelligence (INTEL) and counter intelligence (CI). When resource allocation is occurring in a competitive circumstance, it is often beneficial to gain insight into what your opponent is doing through intelligence. In turn, an opponent may seek to halt or skew the information being gained.

The use of INTEL and CI in TEMPO allows research into the effects this has on the resource allocation process and the coevolved computer player.

The development of a computer player for the game of TEMPO gives us endless possibilities of research. In this research, we have focused on the creation a computer player that can provide a fun and challenging environment for humans learning resource allocation strategies. We investigate the addition of memory to a coevolutionary algorithm for strategy creation. This includes mechanisms to select memory individuals for evaluation of coevolutionary individuals. We describe a successful strategy of selection, based on the way a human's short and long term memory works. We then investigate the use of INTEL and CI in the game of TEMPO, and the way it is used by the coevolved computer players. Through this work, we present a new version of the TEMPO game that more realistically represents INTEL and CI. Finally, we describe a process that uses coevolution to adapt to a human player real-time, to create a tailored game-play experience. This process was tested in a user study, and showed a distinct advantage through the adaptive mechanism. Overall, we have made some important discoveries, and described some limitations that leave future research open. Ultimately, we have shown that our hypothesis is an achievable goal, with an exciting future.

Coevolving a Computer Player for Resource Allocation Games - Using the game of TEMPO as a test space

This work contains no material which has been accepted for the award of any other degree or diploma 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.
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.

Phillipa Melanie Avery
4th September 2008

Acknowledgements

First and foremost, I would like to sincerely thank my mentor and supervisor Zbigniew Michalewicz. His guidance and endless patience has made this work possible. When I first knocked on his door and asked him to be my supervisor, I could not have imagined the challenging and exciting path that lay ahead of me. It has been a true honour to work with him, and I could not have asked for a better supervisor.

I would also like to thank all the people who have made this research possible. Firstly, I would like to thank Martin Schmidt for helping to form the early stages of the research, and Charles Lakos for his input in the latter stages. Thanks also go to Garrison Greenwood, who came along for the ride and made it better, and I am privileged to call a colleague and friend. Additionally, I would like to thank SAPAC for their continued support and provision of resources.

Thanks go to all the friends and reviewers who have helped shape the work over the years. I would especially like to thank all my colleagues, who have always been there to provide me with a pat on the back, slap to the forehead, and occasional knock on the head. Thanks also go to my family, for the encouragement and support.

Special thanks go to my husband Ben, who has given me the support and strength to keep going. I would also like to thank Rachael, who has kept me sane for all these years. This thesis is dedicated to them.

Finally, I would like to thank Mike Brooks and Dave Munro, who gave a stranger at the door a chance. None of this would have been possible without your trust in me, and I am truly grateful.

List of Tables

3.1 Example TEMPO net offensive util scoring 44
3.2 Util adjustments to reflect diminishing returns 45

6.1 Summary of the experiments 97

7.1 Example values and function results 111

8.1 User Study Stage 1 Results 137
8.2 User Study Stage 2 Results 137
8.3 User Study Stage 3 Results 138

List of Figures

2.1 Penalty matrix for the Prisoner's Dilemma 17
2.2 Payoff matrix for the Prisoner's Dilemma 17
2.3 Example game tree 19
2.4 A generic evolutionary algorithm 27
2.5 Example crossover implementation 29
2.6 Example of solutions in multiobjective space 32
2.7 A generic coevolutionary algorithm 35
2.8 Example fuzzy set membership 38

3.1 Example screen of a year's game-play 43
3.2 Diminishing return distribution for util adjustments 45

5.1 The early TEMPO coevolutionary representation 61
5.2 New structure of a chromosome 63
5.3 Membership function example 64
5.4 The coevolutionary algorithm implemented 66

6.1 Baseline success ratio against the static expert 72
6.2 The coevolutionary algorithm with memory implemented 73
6.3 Success ratio against the expert using random selection from history 75
6.4 Probability distribution by function of time 76
6.5 Success ratio against the expert using single memory with linear distribution 77
6.6 Success ratio against the expert for a single run using linear distribution . . 77
6.7 Success ratio against the expert using Laplace probability distribution . . . 78
6.8 Success ratio against the expert using short and long term memory 81
6.9 Success ratio against the expert for a single run using short and long term memory 82
6.10 Success ratio against the expert using gladiator system 84

6.11 Success ratio against the expert with migration incorporated into the gladiator system 85
6.12 Success ratio against the expert for two single runs with migration incorporated into the gladiator system 86
6.13 Using the clustered memory mechanism against the *pwar* expert 90
6.14 Success ratio against the *pwar* expert using linear time probability long term memory 91
6.15 Success ratio against the expert with no alien expert 93
6.16 Success ratio against the expert with individuals evaluated against the expert 95
6.17 Success ratio against the expert with alien expert seeding and evaluation against the expert 96

7.1 Example screen of a year's game-play with the new inputs and data 107
7.2 Example probability density function used for INTEL. 109
7.3 Success ratio against the 'intelligent' expert with same expert inserted in populations 114
7.4 Success ratio against the 'intelligent' expert with no expert inserted into the populations 116
7.5 Success ratio against the old expert with no expert inserted into the populations 117
7.6 Success ratio against the 'intelligent' expert with new fitness function with the expert inserted into the population 118
7.7 Success ratio against the 'intelligent' expert with new fitness function without the expert inserted into the population 119
7.8 Success ratio with population A only allowed weapon and INTEL rules, and population B also allowed CI rules, measured against the old baseline expert 120
7.9 Success ratio with population A only allowed weapon and INTEL rules, and population B also allowed CI rules 120
7.10 Success ratio with population A only allowed weapon rules, and population B also allowed INTEL rules, measured against the old baseline expert . . . 121
7.11 Success ratio for both populations with population A only allowed weapon rules, and population B also allowed INTEL rules 121
7.12 Success ratio for both populations with no penalty to the first five INTEL rules 122

7.13 Mean and standard deviation for each population with no penalty to the first five INTEL rules 123
7.14 Success ratio for the experiments with lowered INTEL costs, against the 'intelligent' expert 124
7.15 Success ratio for the experiments with lowered INTEL costs, against the 'intelligent' expert 124

8.1 The human adaptive coevolutionary process 132
8.2 Screenshot of the TEMPO GUI 134

List of publications

The work carried out for this thesis has been partially published in a number of conference proceedings and a journal article.
The research for Chapter 6 has been published in the following conference proceedings:

1. P. Avery, Z. Michalewicz, and M. Schmidt. *“A historical population in a coevolutionary system”*. In *IEEE Symposium on Computational Intelligence and Games*, Honolulu, Hawaii, USA, 2007.

2. P. Avery and Z. Michalewicz. *“Static experts and dynamic enemies in coevolutionary games”*. In *IEEE Proceedings for Congress on Evolutionary Computation*, Singapore, 2007.

As well as the following journal article:

1. P. Avery, Z. Michalewicz, and M. Schmidt. *“Short and long term memory in coevolution”*. International Journal of Information Technology and Intelligent Computing, 3(1), 2008.

The research for Chapter 7 was published in the following conference proceedings, with the accomplishment of winning best student paper for the conference:

1. P. Avery, G.W. Greenwood and Z. Michalewicz. *“Coevolving Strategic Intelligence”*. In *IEEE Proceedings for Congress on Evolutionary Computation*, Hong Kong, China, 2008.

The research for Chapter 8 has been used for the following paper, which was under review at the time of submission:

1. P. Avery and Z. Michalewicz. *”Adapting to human game play with Coevolution”*. Submitted to *IEEE Symposium on Computational Intelligence and Games*, in August 2008.

It will also be used to contribute to the following book chapter, which is currently in production:

1. P. Avery and Z. Michalewicz. *"Using coevolution to adapt to human gamers "*. In preparation for *Studies in Computation Intelligence*, with the volume *Recent Advances in Machine Learning*. Springer, 2009.