

The Artificial Immune System with Evolved Lymphocytes

by

Alexander J. Graaff

Supervisor: Prof. A.P. Engelbrecht

Department of Computer Science

Degree: Magister Scientiae

KEYWORDS:

Self, non-self, artificial lymphocytes, artificial immune system, affinity distance, threshold, life counter, transition function, genetic algorithms, classification system, negative selection, positive selection, mature, memory

Submitted in partial fulfillment of the requirements for the degree

Magister Scientia

in the Faculty of Engineering, Built Environment and Information Technology

University of Pretoria

November 2003

15.000 919A.5
77 01/2003

Abstract

The main purpose of the natural immune system is to protect the body against any unwanted foreign cells that could infect the body and lead to devastating results. The natural immune system has different lymphocytes to detect and destroy these unwanted foreign patterns. The natural immune system can be modeled into an artificial immune system that can be used to detect any unwanted patterns in a non-biological environment. One of the main tasks of an immune system is to learn the structure of these unwanted patterns for a faster response to future foreign patterns with the same or similar structure. The artificial immune system (AIS) can therefore be seen as a pattern recognition system. The AIS contains artificial lymphocytes (ALC) that classify any pattern either as part of a predetermined set of patterns or not. In the immune system, lymphocytes have different states: Immature, Mature, Memory or Annihilated. Lymphocytes in the annihilated state needs to be removed from the active set of ALCs. The process of moving from one state to the next needs to be controlled in an efficient manner. This dissertation presents an AIS for detection of unwanted patterns with a dynamical active set of ALCs and proposes a threshold function to determine the state of an ALC. The AIS in the dissertation uses evolutionary computation techniques to evolve an optimal set of lymphocytes for better detection of unwanted patterns and removes ALCs in the annihilated state from the active set of ALCs.

Supervisor: Prof. A.P. Engelbrecht

Department of Computer Science

Degree: Magister Scientiae

KEYWORDS:

Self, non-self, artificial lymphocytes, artificial immune system, affinity distance threshold, life counter transition function, genetic algorithms, classification system, negative selection, positive selection, mature, memory

Opsomming

Die hoofdoel van die biologiese immuun stelsel is om die liggaam te beskerm teen enige ongewenste vreemde selle wat die liggaam kan binnedring en sodoende skade aan die liggaam veroorsaak. Die immuun stelsel in die menslike liggaam het verskillende limfosiete wat die ongewenste selle raaksien en vernietig. Dit is moontlik om die biologiese immuun stelsel te modeleer as 'n kunsmatige immuun stelsel wat gebruik kan word om enige ongewenste patrone in 'n nie-biologiese omgewing raak te sien. Een van die hoof funksies van die biologiese immuun stelsel is om die struktuur van die ongewenste selle aan te leer om sodoende 'n vinniger immuun reaksie teenoor moontlike toekomstige ongewenste selle met naastenby of presies dieselfde struktuur te hê. Die kunsmatige immuun stelsel (KIS) kan dus gesien word as 'n patroonherkenningstelsel. Die KIS gebruik kunsmatige limfosiete (KLS) wat enige patroon kan klassifiseer as deel van 'n vooraf-bepaalde stel patrone al dan nie. In die immuun stelsel het die limfosiete verskillende toestande: Onvolwasse, Volwasse, Geheue en Vernietig. Die KLS'e wat in die vernietig-toestand is, moet van die aktiewe stel van KLS'e verwyder word. Die proses om van een toestand na 'n ander toestand oor te gaan moet op 'n doeltreffende wyse bepaal en beheer word. Die verhandeling lê 'n KIS voor om ongewenste patrone met 'n dinamiese aktiewe stel van KLS'e te herken en stel 'n toestands-veranderingsfunksie voor om die toestand van 'n KLS te bepaal. Die KIS in die verhandeling maak gebruik van evolusionêre komputasie tegnieke om 'n optimale stel van KLS'e te evolueer wat ongewenste patrone beter kan herken, en verwyder KLS'e in die vernietig-toestand vanuit die aktiewe stel van KLS'e.

Studieleier: Prof. A.P. Engelbrecht

Departement Rekenaarwetenskap

Graad: Magister Scientiae

Acknowledgments

- My Father in heaven for all my privileges and His mercy.
- My dad, mother and Dollie for their support.
- Prof. Andries Engelbrecht for his guidance and patience.
- CIRG Team, especially Werner van Rensburg and Nelis Franken.
- Telkom SA Ltd. - Centre of Excellence for their financial support.

Contents

Introduction	1
The Natural Immune System	5
1.1 Antibodies and Antigens	5
1.2 The White Blood Cells	6
1.2.1 The Lymphocytes	6
1.2.1.1 The B-Cell	6
1.2.1.2 The Helper-T Cell (HTL)	9
1.2.1.3 The Natural-Killer-TCell (NKTC)	9
1.3 The Cloning Process of the Lymphocyte	10
1.4 Learning the Antigen Structure	10
1.5 Immunity Types	11
1.6 Conclusion	11
Evolutionary Computation	13
2.1 A General Evolutionary Algorithm	14
2.2 EC Paradigms	14
2.2.1 Genetic Algorithms (GA)	15
2.2.2 Genetic Programming (GP)	15
2.2.3 Evolutionary Programming (EP)	15
2.2.4 Evolutionary Strategies (ES)	16
2.3 The Chromosome	16
2.4 Calculating the Fitness	17
2.5 Reproduction	18
2.5.1 Selection	18
2.5.2 Crossover	20

Contents

1	Introduction	1
2	The Natural Immune System	5
2.1	Antibodies and Antigens	6
2.2	The White Cells	7
2.2.1	The Lymphocytes	7
2.2.1.1	The B-Cell	8
2.2.1.2	The Helper-T-Cell (HTC)	9
2.2.1.3	The Natural-Killer-T-Cell (NKTC)	9
2.3	The Cloning Process of the Lymphocyte	10
2.4	Learning the Antigen Structure	10
2.5	Immunity Types	11
2.6	Conclusion	11
3	Evolutionary Computation	13
3.1	A General Evolutionary Algorithm	14
3.2	EC Paradigms	14
3.2.1	Genetic Algorithms (GA)	15
3.2.2	Genetic Programming (GP)	15
3.2.3	Evolutionary Programming (EP)	15
3.2.4	Evolutionary Strategies (ES)	15
3.3	The Chromosome	16
3.4	Calculating the Fitness	17
3.5	Reproduction	18
3.5.1	Selection	18
3.5.2	Crossover	20

CONTENTS

vii

3.5.2.1	Continuous-valued Chromosomes	20
3.5.2.2	Nominal-valued Chromosomes	21
3.6	Mutation	22
3.6.1	Continuous-valued Chromosomes	23
3.6.2	Nominal-valued Chromosomes	23
3.7	The Initial Population	24
3.8	Convergence	25
3.9	Conclusion	25
4	The Artificial Immune System	27
4.1	Natural to Artificial	28
4.2	The Network Theory	31
4.3	The Genetic Artificial Immune System	32
4.4	Applications and Other Models	32
4.5	Conclusion	33
5	The Artificial Lymphocyte Life Counter	34
5.1	Introduction	34
5.2	The Artificial Lymphocyte	36
5.2.1	Life Cycle of the Artificial Lymphocyte	36
5.2.2	Training the ALC to Cover Non-Self Space	37
5.2.2.1	Adapted Negative Selection	37
5.2.2.2	Positive Selection	38
5.2.3	Matching a Non-Self Pattern	38
5.2.3.1	Adapted Negative Selection Hit Counter	39
5.2.3.2	Positive Selection Hit Counter	40
5.2.3.3	The Hit Ratio	40
5.3	The Life Counter	41
5.3.1	Life Counter Threshold Function	42
5.4	The Expected Matching Ratio	44
5.5	Conclusion	45
6	Evolving Artificial Lymphocytes	46
6.1	Genetic Algorithm to Evolve an ALC	46
6.1.1	ALC as Chromosome	48

CONTENTS

viii

6.1.2	The Initial Population	49
6.1.3	Evaluating the Chromosome's Fitness	49
6.1.3.1	Adapted Negative Selection Fitness	49
6.1.3.2	Positive Selection Fitness	50
6.1.4	Parent Selection	50
6.1.5	Mutation	51
6.1.6	Selection of the New Population	52
6.1.7	Convergence of the GA	52
6.2	The GAIS Algorithm	52
6.2.1	Initialisation of the ALC Set	55
6.2.2	Convergence of the ALC Set	55
6.3	Conclusion	55
7	Experimental Results	57
7.1	Iris	59
7.1.1	Setosa	59
7.1.2	Versicolor	63
7.1.3	Virginica	63
7.1.4	Conclusion: Iris	71
7.2	Wisconsin Breast Cancer	71
7.2.1	Benign	71
7.2.2	Malignant	73
7.2.3	Conclusion: Wisconsin Breast Cancer	76
7.3	Mushroom	79
7.3.1	Edible	79
7.3.2	Poisonous	82
7.3.3	Conclusion: Mushroom	82
7.4	Glass	86
7.4.1	Building-window-float	86
7.4.2	Building-window-nonfloat	91
7.4.3	Containers	91
7.4.4	Headlamps	96
7.4.5	Tableware	96
7.4.6	Vehicle-window-float	102
7.4.7	Conclusion: Glass	109

CONTENTS

ix

7.5	Car Evaluation	109
7.5.1	Acceptable	109
7.5.2	Good	112
7.5.3	Unacceptable	112
7.5.4	Very Good	118
7.5.5	Conclusion: Car Evaluation	118
7.6	Comparing the Results	123
8	Conclusion and Future Work	125
8.1	Conclusion	125
8.2	Future work	126
	Bibliography	126
A	Publications	135
B	Symbols	136
C	Abbreviations	138
D	Glossary	140
3.2	Venn-diagram of Adapted Negative Selection ALC	18
3.3	Venn-diagram of Positive Selection ALC	19
3.3	Hypobolic Tangent Function and Sigmoid Function	6
6.1	Flow layout of the Artificial Immune System	23
7.1	Sezen - Negative selection with IS=37 and W1=50	61
7.2	Seton - Positive selection with IS=37 and W1=50	64
7.3	Wasseler - Negative selection with IS=37 and W1=75	65
7.4	Wasseler - Positive selection with IS=37 and W1=75	68
7.5	Reign - Negative selection with IS=524 and W1=25	75
7.6	Malignant - Positive selection with IS=175 and W1=75	78
7.7	Potomac - Positive selection with IS=2031 and W1=100	85
7.8	Building window-floor - Negative selection with IS=106 and W1=25	88
7.9	Building window-floor - Positive selection with IS=53 and W1=25	90

List of Figures

2.1	Antigen-Antibody-Complex	6
2.2	White Cell types	7
2.3	B-Cell develops into plasma cell, producing antibodies	8
2.4	Macrophage and NKTC	10
2.5	Life cycle of a lymphocyte	12
3.1	One-point crossover	21
3.2	Two-point crossover	21
3.3	Uniform crossover	22
3.4	Random Mutation	24
3.5	In-order Mutation	24
5.1	Venn-diagram of Adapted Negative Selection ALC	38
5.2	Venn-diagram of Positive Selection ALC	39
5.3	Hyperbolic Tangent Function and Sigmoid Function	42
6.1	Flow layout of the Artificial Immune System	53
7.1	<i>Setosa</i> - Negative selection with IS=37 and W1=50	61
7.2	<i>Setosa</i> - Positive selection with IS=37 and W1=50	64
7.3	<i>Versicolor</i> - Negative selection with IS=37 and W1=75	66
7.4	<i>Versicolor</i> - Positive selection with IS=37 and W1=75	68
7.5	<i>Benign</i> - Negative selection with IS=524 and W1=25	73
7.6	<i>Malignant</i> - Positive selection with IS=175 and W1=75	78
7.7	<i>Poisonous</i> - Positive selection with IS=2031 and W1=100	85
7.8	<i>Building-window-float</i> - Negative selection with IS=106 and W1=25	88
7.9	<i>Building-window-float</i> - Positive selection with IS=53 and W1=25	90

LIST OF FIGURES

xi

7.10	<i>Containers</i> - Positive selection with IS=53 and W1=25	97
7.11	<i>Headlamps</i> - Negative selection with IS=53 and W1=25	99
7.12	<i>Tableware</i> - Negative selection with IS=53 and W1=50	103
7.13	<i>Tableware</i> - Positive selection with IS=106 and W1=50	105
7.14	<i>Vehicle-window-float</i> - Negative selection with IS=106 and W1=25	107
7.15	<i>Vehicle-window-float</i> - Positive selection with IS=106 and W1=25	110
7.16	<i>Good</i> - Negative selection with IS=432 and W1=50	115
7.17	<i>Very Good</i> - Negative selection with IS=432 and W1=50	121
7.18	<i>Amica</i> - Negative selection	60
7.19	<i>Amica</i> - Positive selection	62
7.20	<i>Amicolor</i> - Negative selection	63
7.21	<i>Amicolor</i> - Positive selection	67
7.22	<i>Wrylinka</i> - Negative selection	69
7.23	<i>Wrylinka</i> - Positive selection	73
7.24	<i>Radian</i> - Negative selection	75
7.25	<i>Radian</i> - Positive selection	78
7.26	<i>Malignant</i> - Negative selection	80
7.27	<i>Malignant</i> - Positive selection	82
7.28	<i>Edible</i> - Negative selection	83
7.29	<i>Edible</i> - Positive selection	87
7.30	<i>Polygonal</i> - Negative selection	89
7.31	<i>Polygonal</i> - Positive selection	94
7.32	<i>Building-window-float</i> - Negative selection	97
7.33	<i>Building-window-float</i> - Positive selection	99
7.34	<i>Building-window-nonfloat</i> - Negative selection	102
7.35	<i>Building-window-nonfloat</i> - Positive selection	104
7.36	<i>Containers</i> - Negative selection	104
7.37	<i>Containers</i> - Positive selection	105
7.38	<i>Headlamps</i> - Negative selection	108
7.39	<i>Headlamps</i> - Positive selection	109
7.40	<i>Tableware</i> - Negative selection	109
7.41	<i>Tableware</i> - Positive selection	110
7.42	<i>Vehicle-window-float</i> - Negative selection	110
7.43	<i>Vehicle-window-float</i> - Positive selection	110

List of Tables

7.1	<i>Setosa</i> - Negative selection	60
7.2	<i>Setosa</i> - Positive selection	62
7.3	<i>Versicolor</i> - Negative selection	65
7.4	<i>Versicolor</i> - Positive selection	67
7.5	<i>Virginica</i> - Negative selection	69
7.6	<i>Virginica</i> - Positive selection	70
7.7	<i>Benign</i> - Negative selection	72
7.8	<i>Benign</i> - Positive selection	74
7.9	<i>Malignant</i> - Negative selection	75
7.10	<i>Malignant</i> - Positive selection	77
7.11	<i>Edible</i> - Negative selection	80
7.12	<i>Edible</i> - Positive selection	81
7.13	<i>Poisonous</i> - Negative selection	83
7.14	<i>Poisonous</i> - Positive selection	84
7.15	<i>Building-window-float</i> - Negative selection	87
7.16	<i>Building-window-float</i> - Positive selection	89
7.17	<i>Building-window-nonfloat</i> - Negative selection	92
7.18	<i>Building-window-nonfloat</i> - Positive selection	93
7.19	<i>Containers</i> - Negative selection	94
7.20	<i>Containers</i> - Positive selection	95
7.21	<i>Headlamps</i> - Negative selection	98
7.22	<i>Headlamps</i> - Positive selection	100
7.23	<i>Tableware</i> - Negative selection	101
7.24	<i>Tableware</i> - Positive selection	104
7.25	<i>Vehicle-window-float</i> - Negative selection	106
7.26	<i>Vehicle-window-float</i> - Positive selection	108

LIST OF TABLES

xiii

7.27	<i>Acceptable</i> - Negative selection	111
7.28	<i>Acceptable</i> - Positive selection	113
7.29	<i>Good</i> - Negative selection	114
7.30	<i>Good</i> - Positive selection	116
7.31	<i>Unacceptable</i> - Negative selection	117
7.32	<i>Unacceptable</i> - Positive selection	119
7.33	<i>Very Good</i> - Negative selection	120
7.34	<i>Very Good</i> - Positive selection	122
7.35	Summarised results	123
B.1	Table of symbols	136
B.2	Table of symbols (continued)	137

Classification is the process to orderly separate a group of similar patterns into classes according to patterns or characteristics common to each class. Various classification models have been developed, which can be broadly divided into two classes of algorithms. The first class of algorithms include decision trees [70], rough sets [58, 61] and rule induction [83], while the second class includes algorithms that model a natural process. Examples of these are the artificial neural network that models the biological neural network in the brain [6], evolutionary computation techniques that model the natural evolution of organisms [1] and swarm intelligence that models the behavior of a structured collection of interacting organisms to solve a global objective [87]. The modeling of natural processes has also proven to be successful in classification problems, optimisation problems, control, pattern matching and data mining. These computational techniques are usually trained with negative and positive examples that have been pre-classified according to a specific concept or rule. This training method is known as supervised learning and the trained-model must be able to correctly predict or classify any pattern not seen before. Classification models have also been developed using unsupervised learning algorithms where the learning process consists of automatically discovering similar patterns in data without relying on an external teacher [51].

Recently, artificial immune systems (AIS) have been developed as an alternative classification algorithm. An AIS is modeled after the natural immune system (NIS) to detect foreign patterns in a non-biological environment. The NIS has the ability to not only learn valid patterns and recognize foreign patterns (or anomalies), but also has the ability to memorize general foreign patterns (antibodies) [56, 63]. Contrary to standard classification algorithms, the AIS can be trained on positive patterns alone. After training on positive patterns, the artificial immune system can