Artificial Immune Systems: A Novel Paradigm to Pattern Recognition

L. N. de Castro and J. Timmis

Computing Laboratory

University of Kent at Canterbury

Kent, Canterbury, CT2 7NF, United Kingdom

e-mail: {L.N.deCastro, J.Timmis}@ukc.ac.uk

Abstract

This chapter introduces a new computational intelligence paradigm to perform pattern recognition, named *Artificial Immune Systems* (AIS). AIS take inspiration from the immune system in order to build novel computational tools to solve problems in a vast range of domain areas. The basic immune theories used to explain how the immune system perform pattern recognition are described and their corresponding computational models are presented. This is followed with a survey from the literature of AIS applied to pattern recognition. The chapter is concluded with a trade-off between AIS and artificial neural networks as pattern recognition paradigms.

Keywords: Artificial Immune Systems, Negative Selection, Clonal Selection, Immune Network.

1 Introduction

The vertebrate immune system (IS) is one of the most intricate bodily systems and its complexity is sometimes compared to that of the brain. With the advances in the biology and molecular genetics, the comprehension of how the immune system behaves is increasing very rapidly. The knowledge about the IS functioning has unravelled several of its main operative mechanisms. These mechanisms have demonstrated to be very interesting not only from a biological standpoint, but also under a computational perspective. Similarly to the way the nervous system inspired the development of artificial neural networks (ANN), the immune system has now led to the emergence of artificial immune systems (AIS) as a novel computational intelligence paradigm.

Artificial immune systems can be defined as abstract or metaphorical computational systems developed using ideas, theories, and components, extracted from the immune system. Most AIS aim at solving complex computational or engineering problems, such as pattern recognition, elimination, and optimisation. This is a crucial distinction between AIS and theoretical immune system models. While the former is devoted primarily to computing, the latter is focused on the modelling of the IS in order to understand its behaviour, so that contributions can be made to the biological sciences. It is not exclusive, however, the use of one approach into the other and, indeed, theoretical models of the IS have contributed to the development of AIS.

This chapter is organised as follows. Section 2 describes relevant immune theories for pattern recognition and introduces their computational counterparts. In Section 3, we briefly describe how to model pattern recognition in artificial immune systems, and present a simple illustrative example. Section 4 contains a survey of AIS for pattern recognition, and Section 5 contrast the use of AIS with the use of ANN when applied to pattern recognition tasks. The chapter is concluded in Section 6.

2 Biological and Artificial Immune Systems

All living organisms are capable of presenting some type of defence against foreign attack. The evolution of species that resulted in the emergence of the vertebrates also led to the evolution of the immune system of this species. The vertebrate immune system is particularly interesting due to its several computational capabilities, as will be discussed throughout this section.

The immune system of vertebrates is composed of a great variety of molecules, cells, and organs spread all over the body. There is no central organ controlling the functioning of the immune system, and there are several elements in transit and in different compartments performing complementary roles. The main task of the immune system is to survey the organism in the search for malfunctioning cells from their own body (e.g., cancer and tumour cells), and foreign disease causing elements (e.g., viruses and bacteria). Every element that can be recognised by the immune system is called an *antigen* (Ag). The cells that originally belong to our body and are harmless to its functioning are termed *self* (or *self antigens*), while the disease causing elements are named *nonself* (or *nonself antigens*). The immune system, thus, has to be capable of distinguishing between what is *self* from what is *nonself*; a process called *self*/*nonself discrimination*, and performed basically through pattern recognition events.

From a pattern recognition perspective, the most appealing characteristic of the IS is the presence of *receptor molecules*, on the surface of immune cells, capable of recognising an almost limitless range of antigenic patterns. One can identify two major groups of immune cells, known as B-cells and T-cells. These two types of cells are rather similar, but differ with relation to how they recognise antigens and by their functional roles. B-cells are capable of recognising antigens free in solution (e.g., in the blood stream), while T-cells require antigens to be presented by other accessory cells.

Fig. 1(a) illustrates that antigens are covered with molecules, named *epitopes*. These allow them to be recognised by the receptor molecules on the surface of B-cells, called *antibodies* (Ab). In contrast, Fig. 1(b) shows how for an antigen to be recognised by a T-cell receptor, it has to be processed and presented by an *accessory cell*.

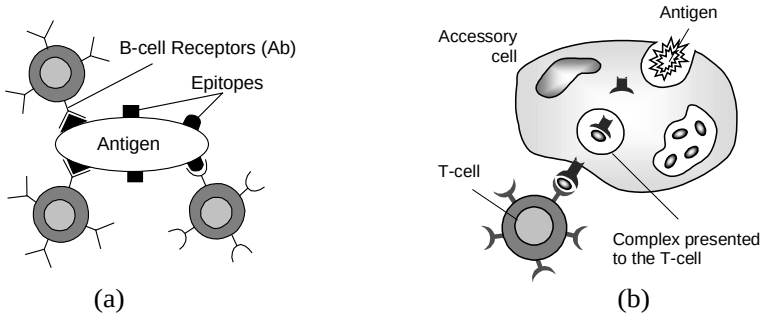

Figure 1: Pattern recognition in the immune system. (a) B-cell recognising an antigen (Ag) free in solution. (b) T-cell recognising an antigen presented by an accessory cell.

Antigenic recognition is the first pre-requisite for the immune system to be activated and to mount an *immune response*. The recognition has to satisfy some criteria. First, the cell receptor recognises an antigen with a certain *affinity*, and a *binding* between the receptor and the antigen occurs with strength proportional to this affinity. If the affinity is greater than a given threshold, named *affinity threshold*, then the immune system is activated. The nature of antigen, type of recognising cell, and the recognition site also influence the outcome of an encounter between an antigen and a cell receptor.

The human immune system contains an organ called *thymus* that is located behind the breastbone, which performs a crucial role in the *maturation* of T-cells. After T-cells are generated, they migrate into the thymus where they mature. During this maturation, all T-cells that recognise self-antigens are excluded from the population of T-cells; a process termed *negative selection*. If a B-cell encounters a nonself antigen with a sufficient affinity, it proliferates and differentiates into memory and effector cells; a process named *clonal selection*. In contrast, if a B-cell recognises a self-antigen, it might result in suppression, as proposed by the *immune network theory*. In the following subsections, each of these processes (negative selection, clonal selection, and network theory) will be described separately, along with their computational algorithms counterparts.

2.1 Negative Selection

The thymus is responsible for the maturation of T-cells; and is protected by a blood barrier capable of efficiently excluding nonself antigens from the thymic environment. Thus, most elements found within the thymus are representative of self instead of nonself. As an outcome, the T-cells containing receptors capable of recognising these self antigens presented in the thymus are eliminated from the repertoire of T-cells through a process named *negative selection* [34]. All T-cells that leave the thymus to circulate throughout the body are said to be *tolerant* to self, i.e., they do not respond to self.

From an information processing perspective, negative selection presents an alternative paradigm to perform pattern recognition by storing information about the complement set (nonself) of the patterns to be recognised (self). A negative selection algorithm [14] has been proposed in the literature with applications focused on the problem of anomaly detection, such as computer and network intrusion detection, time series prediction, image inspection and segmentation, and hardware fault tolerance.

Given an appropriate problem representation (Section 3), define the set of patterns to be protected and call it the *self- set* (**P**). Based upon the negative selection algorithm, generate a set of *detectors* (**M**) that will be responsible to identify all elements that do not belong to the self-set, i.e., the nonself elements. The negative selection algorithm runs as follows (Fig 2(a)):

1. Generate random candidate elements (**C**) using the same representation adopted;
2. Compare (match) the elements in **C** with the elements in **P**. If a match occurs, i.e., if an element of **P** is recognised by an element of **C**, then discard this element of **C**; else store this element of **C** in the detector set **M**.

After generating the set of detectors (**M**), the next stage of the algorithm consists in *monitoring* the system for the presence of nonself patterns (Fig 2(b)). In this case, assume a set **P*** of patterns to be protected. This set might be composed of the set **P** plus other new patterns, or it can be a completely novel set.

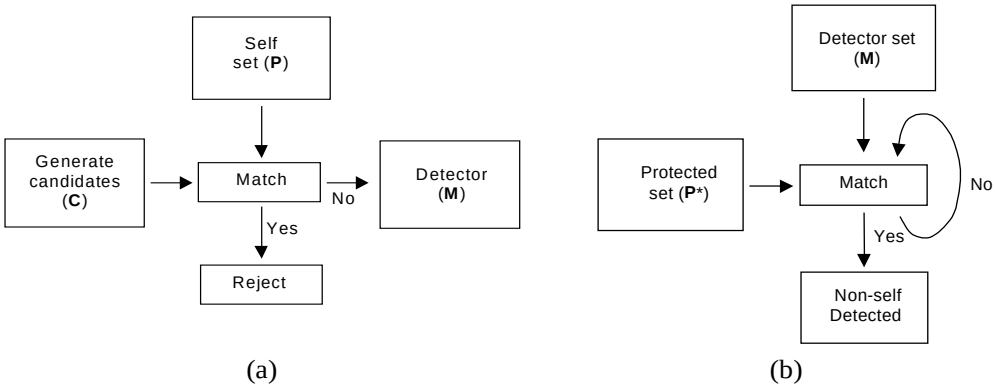

Figure 2: Pattern recognition via the negative selection algorithm. (a) Generating the set of detectors. (b) Monitoring for the presence of undesired (nonself) patterns.

For all elements of the detector set, that corresponds to the nonself patterns, check if it recognises (matches) an element of **P*** and, if yes, then a nonself pattern was recognised and an action has to be taken. The resulting action of detecting nonself varies according to the problem under evaluation and extrapolates the pattern recognition scope of this chapter.

2.2 Clonal Selection

Complementary to the role of negative selection, *clonal selection* is the theory used to explain how an immune response is mounted when a nonself antigenic pattern is recognised by a B-cell [1]. Fig. 3 illustrates the clonal selection, expansion (proliferation), and affinity maturation processes. In brief, when a B-cell receptor recognises a nonself antigen with a certain affinity, it is selected to proliferate and produce antibodies in high volumes. The antibodies are soluble forms of the B-cell receptors that are released from the B-cell surface to cope with the invading nonself antigen. Antibodies bind to antigens leading to their eventual elimination by other immune cells. Proliferation in the case of immune cells is asexual, a mitotic process; the cells divide themselves (there is no crossover). During reproduction, the B-cell progenies (clones) undergo a hyper *mutation* process that, together with a strong selective pressure, result in B-cells with antigenic receptors presenting higher affinities with the selective antigen. This whole process of mutation and selection is known as the *maturation of the immune response* [35] and is analogous to the natural selection of species [20]. In addition to differentiating into antibody producing cells, the activated B-cells with high antigenic affinities are selected to become memory cells with long life spans. These memory cells are pre-eminent in future responses to this same antigenic pattern, or a similar one.

Other important features of clonal selection relevant from the viewpoint of computation are:

1. An antigen selects several immune cells to proliferate. The proliferation rate of each immune cell is proportional to its affinity with the selective antigen: the higher the affinity, the higher the number of offspring generated, and vice-versa;
2. In complete opposition to the proliferation rate, the mutation suffered by each immune cell during reproduction is inversely proportional to the affinity of the cell receptor with the antigen: the higher the affinity, the smaller the mutation, and vice-versa.

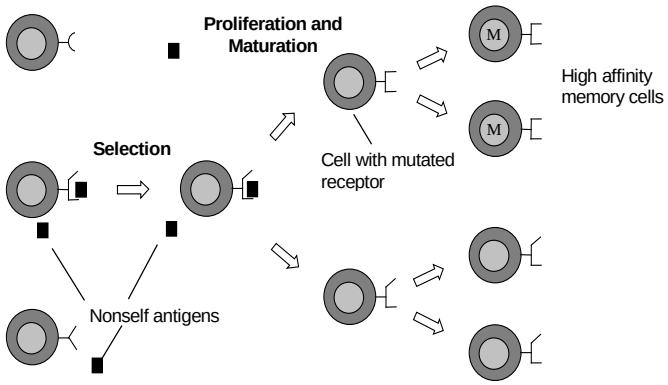

Figure 3: Clonal selection, expansion (proliferation), affinity maturation, and maintenance of memory cells. The highest affinity cells are selected to proliferate. Their progenies (clones) suffer mutation with high rates and those whose receptors present high affinity with the antigen are maintained as memory cells.

Some authors [15] have argued that a genetic algorithm without crossover is a reasonable model of clonal selection. However, the standard genetic algorithm does not account for important properties such as affinity proportional reproduction and mutation. Other authors [10] proposed a clonal selection algorithm, named CLONALG, to fulfil these basic processes involved in clonal selection. This algorithm was initially proposed to perform pattern recognition and then adapted to solve multi-modal optimisation tasks. Given a set of patterns to be recognised (**P**), the basic steps of the CLONALG algorithm are as follows:

1. Randomly initialise a population of individuals (**M**);
2. For each pattern of **P**, present it to the population **M** and determine its affinity (match) with each element of the population **M**;
3. Select n_1 of the best highest affinity elements of **M** and generate copies of these individuals proportionally to their affinity with the antigen. The higher the affinity, the higher the number of copies, and vice-versa;
4. Mutate all these copies with a rate proportional to their affinity with the input pattern: the higher the affinity, the smaller the mutation rate, and vice-versa.
5. Add these mutated individuals to the population **M** and re-select n_2 of these maturated (optimised) individuals to be kept as memories of the system;
6. Repeat Steps 2 to 5 until a certain criterion is met, such as a minimum pattern recognition or classification error.

Note that this algorithm allows the artificial immune system to become increasingly better at its task of recognising patterns (antigens). Thus, based upon an evolutionary-like behaviour, CLONALG learns to recognise patterns.

2.3 Immune Network

The *immune network theory* proposes that the immune system has a dynamic behaviour even in the absence of external stimuli [24]. It is suggested that the immune cells and molecules are capable of recognising each other, what endows the system with an eigen-behaviour that is not dependent on foreign stimulation. Several immunologists have refuted this theory, e.g. [32], however its computational aspects are relevant and it has proved itself to be a powerful model for computational systems.

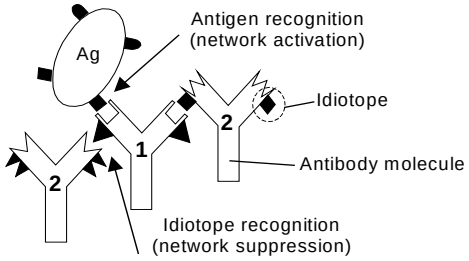

Figure 4: Immune network theory. The recognition of antigen by an antibody (cell receptor) leads to network activation, while the recognition of an idiotope by another antibody results in network suppression. Antibody Ab_2 is said to be the internal image of the antigen *Ag*, because Ab_1 is capable of recognising the antigen and also Ab_2.

According to the immune network theory, the receptor molecules contained in the surface of the immune cells present markers, named *idiotopes*, which can be recognised by receptors on other immune cells. These idiotopes are displayed in and/or around the same portions of the receptors that recognise nonself antigens. Fig. 4 provides a simple illustration of the immune network theory. To explain the network theory, assume that a receptor (antibody) Ab_1 on a B-cell recognises a nonself antigen *Ag*. Assume now, that this same receptor Ab_1 also recognises an idiotope i_2 on another B-cell receptor Ab_2. Keeping track of the fact that i_2 is part of Ab_2, Ab_1 is capable of recognising both *Ag* and Ab_2. Thus, Ab_2 is said to be the *internal image* of *Ag*, more precisely, i_2 is the internal image of *Ag*. The recognition of idiotopes on a cell receptor by other cell receptors, lead to ever increasing sets of *connected* cell receptors and molecules. Note that the network in this case, is a network of affinities, which different from the 'hardwired' network of the nervous system. As a result of the network recognition events, it was suggested that the recognition of a cell receptor by another cell receptor results in network suppression, whilst the recognition of an antigen by a cell receptor results in network activation and cell proliferation. The original theory did not account explicitly for the results of network activation and/or suppression, and the various artificial immune networks found in the literature model it in a particular form.

Recently, the most influential artificial immune network models found in the literature are [9] and [43]. Due to limited space, we will restrict ourselves to the description of only one of these two network models, for an overview of [43] refer to [44]. The work presented in [8] makes use of the clonal selection algorithm (CLONALG), described in Section 2.2 to explain how the immune network model responds to nonself antigens i.e. becomes activated. The recognition of cell receptors by other cell receptors results in network suppression. This is modelled by eliminating all but one of the self-recognising cells. Given a set of patterns (**P**) to be recognised, the basic algorithm runs as follows:

1. Randomly initialise the network population;
2. For each antigenic pattern in **P** apply the CLONALG algorithm that will return a set of memory cells (**M***) and their co-ordinates for the current antigen;
3. Determine the affinity (degree of matching) among all the individuals of **M***;
4. Eliminate all but one of the individuals in **M*** whose affinities are greater than a given threshold. The purpose of this process is to eliminate redundancy in the network by suppressing self-recognising elements;
5. Concatenate the remaining individuals of the previous step with the remaining individuals found for each antigenic pattern presented. This will result in a large population of memory individuals **M**;

6. Determine the affinity of the whole population **M** and suppress all but one of the self-recognising elements. This will result in a reduced final population of memory cells that recognise and follow the spatial distribution of the antigens.
7. Repeat Steps 2 to 6 until a pre-defined stopping criterion is met, such as a minimum pattern recognition or classification error.

Affinity in this case can be taken to mean the degree of recognition or match, between the elements of the artificial immune system itself (self), and among them and the environment (nonself).

3 Modelling Pattern Recognition in AIS

Up to this point, the most relevant immune principles and their corresponding computational counterparts to perform pattern recognition have been presented. In order to apply these algorithms to computational problems, there is a need to specify a limited number of other aspects of artificial immune systems, not as yet covered. The first aspect to introduce is the most relevant representations to be applied to model self and nonself patterns. Here the self-patterns correspond to the components of the AIS responsible for recognising the input patterns (nonself). Secondly, the mechanism by which the evaluation of the degree of match (affinity), or degree of recognition, of an input pattern by an element of the AIS has to be discussed.

To model immune cells, molecules, and the antigenic patterns, the *shape-space* approach proposed in [37] is usually adopted. As illustrated in Figs. 1 and 3, recognition of antigens by cell receptors occurs through a complementarity in the antigenic *shape* with relation to the shape of the cell receptor. Although AIS model recognition through pattern matching, given certain affinity functions to be described further, performing pattern recognition through complementarity or similarity is based more on practical aspects than on biological plausibility.

The shape-space approach proposes that an attribute string $s = \langle s_1, s_2, \ldots, s_L \rangle$ in an L-dimensional shape-space, S, ($s \in S^L$), can represent any immune cell or molecule. Each attribute of this string is supposed to represent a feature of the immune cell or molecule, such as its charge, van der Wall interactions, etc. In the development of AIS the mapping from the attributes to their biological counterparts is usually not relevant. The type of attributes used to represent the string will define partially the shape-space under study, and is highly dependent on the problem domain. Any shape-space constructed from a finite alphabet of length k constitutes a k-ary Hamming shape-space. As an example, an attribute string built upon the set of binary elements {0,1} corresponds to a binary Hamming shape-space [11]. It can be thought of, in this case, of a problem of recognising a set of characters represented by matrices composed of 0's and 1's. Each element of a matrix corresponds to a pixel in the character.

If the elements of s are represented by real-valued vectors, then we have an Euclidean shape-space. Most of the AIS found in the literature employ binary Hamming or Euclidean shape-spaces. Other types of shape-spaces are also possible, such as symbolic shape-spaces, which combine different (symbolic) attributes in the representation of a single string s. These are usually found in data mining applications, where the data might contain symbolic information like age, name, etc., of a set of patterns.

Another important characteristic of the artificial immune systems is that most of them are population based. It means that they are composed of a set of individuals, representing immune cells and molecules, which have to perform a given role; in our

context, pattern recognition. If we recapitulate the three immune processes reviewed, negative selection, clonal selection, and immune network, all of them rely on a population **M** of individuals to recognise a set **P** of patterns. The negative selection algorithm has to define a set of detectors for nonself patterns; clonal selection reproduces, maturates, and selects self-cells to recognise a set of nonself; and the immune network maintains a set of individuals, connected as a network, to recognise self and nonself.

Assume now the availability of a set of N patterns (antigens) p_i, $i = 1,\ldots N$ ($p_i \in \mathbf{P}$) to be recognised, and a set of M immune cells and/or molecules (antibodies) m_j, $j = 1,\ldots M$ ($m_j \in \mathbf{M}$) to be used as pattern recognisers (via negative, clonal or network algorithms). Assume also, that both have the same length L ($p_i, m_j \in S^L$).

Consider first the binary Hamming shape-space case, which is the most widely used. There are several expressions that can be employed in the determination of the degree of match or affinity between an element of **P** and an element of **M**. The simplest case is to simply calculate the Hamming distance (D_H) between these two elements, as given by Eq. (1). Another approach is to search for a sequence of r-contiguous bits [13], and if the number of r-contiguous matches between the strings is greater than a given threshold, then recognition is said to have occurred. As the last approach to be mentioned here, we can describe the affinity measure of Hunt [22], given by Eq. (2). This last method has the advantage that it favours sequences of complementary matches, thus searching for similar regions between the attribute strings (patterns).

$$D_H = \sum_{i=1}^{L} \delta, \text{where} \quad \delta = \begin{cases} 1 & \text{if } p_i \neq m_i \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

$$D = D_H + \sum_i 2^{l_i} , \tag{2}$$

where l_i is the length of the i-th sequence of matching bits longer than 2.

In the case of Euclidean shape-spaces, the Euclidean distance can be used to evaluate the affinity between any two components of the system. Other approaches such as the Manhattan distance may also be employed.

Note that all the methods described rely basically, on determining the match between strings. However, there are AIS in the literature that take into account other aspects, such as the number of patterns matched by each antibody (e.g. [8]).

3.1 A Simple Illustrative Example

Assume that an AIS capable of recognising the binary patterns illustrated in Fig. 5 needs to be designed. Each of these characters represent an attribute string of length $L = 400$ (resolution 20×20). The matrix **P** of patterns to be recognised has a dimension $\mathbf{P} \in S^{5\times400}$. It is now possible to use any of the three algorithms described to design a pattern recogniser for these characters.

Consider first the negative selection algorithm. Using any of the affinity measures described above for binary Hamming shape-spaces, the algorithm runs as follows. Generate (randomly) a set of candidate recognisers **C** using the same representation as that of **P**, and match them against the patterns in **P**. Select as detectors **M** all those elements from **C** that do not match any element of **P** given an *affinity threshold*. The affinity threshold controls the specificity of the elements of **M**: the higher the threshold, the more specific the elements, thus the more elements are required to recognise **P**. Note that, in this case, the detectors generated recognise every element that does not belong to **P**, instead of those who belong.

Figure 5: Illustrative input data set for an AIS pattern recogniser.

The clonal selection algorithm (CLONALG) would learn to recognise the patterns in **P** by reproducing, mutating, and selecting individuals already present in an initial population **M**.

Finally, the immune network model would learn to recognise the patterns in **P** by selecting randomly initialised elements from **M**, applying CLONALG to learn the patterns, and then performing the network activities to define inter-cell connectivity and the final number of individuals in the population.

4 A Survey of AIS for Pattern Recognition

The applications of artificial immune systems are vast, ranging from machine learning to robotic autonomous navigation. This section will review some of the works from the AIS literature applied to the pattern recognition domain. The rationale is to provide a guide to the literature and a brief description of the scope of applications of the algorithms. The section is divided into two parts for ease of comprehension: 1) computer security, and 2) other applications.

The problem of protecting computers (or networks of computers) from viruses, unauthorised users, etc., constitutes a rich field of research for pattern recognition systems. Due, mainly, to the appealing intuitive metaphor of building artificial immune systems to detect computer viruses, there has been a great interest from the computer science community to this particular application. The use of the negative and clonal selection algorithms have been widely tested on this application. The former because it is an inherent anomaly (change) detection system, constituting a particular case of a pattern recognition device. The latter, the clonal selection algorithm, has been used in conjunction to negative selection due to its learning capabilities.

Other more classical pattern recognition tasks, such as character recognition, and data analysis have also been studied within artificial immune systems. Section 4.2 reviews AIS applications to these problems.

4.1 Computer Security

Using the *r*-contiguous bit rule, the work presented in [14] compared the problem of protecting computer systems to that of learning to distinguish between self and nonself, and proposed the negative selection algorithm described in Section 2.1. Therefore, pattern recognition was performed by generating a set of patterns complementary to the ones to be recognised.

In the system developed in [25,26], a set of antibodies to previously not encountered computer viruses or worms (agents) was generated so as to promote a faster and stronger response to future infecting agents. The author was also concerned with minimising the risk of the computer immune system mistakenly identifying legitimate software as being undesirable (nonself). Thus, this particular AIS accounted for the recognition of self and nonself patterns.

In [42] the authors articulated a broad vision for the development of a computer immune system by discussing the immune system in terms of a set of organising principles and possible architectures for implementation. From a computational

standpoint, given the many interesting properties of the immune system, the authors described several possibilities to design pattern recognition systems based on direct mappings between immune components and current computer system architectures.

In [36], the authors proposed a distributed approach to computer virus detection and neutralisation by autonomous and heterogeneous immune agents. Their system detects viruses by matching a self-information, like the first few bytes of the head of a file, the file size and path, etc., against the current host files. Viruses were neutralised by overwriting the self-information on the infected files, and the recovering was attained by copying the same file from other uninfected hosts through the computer network. One of the interesting aspects of this work is that it accounts not only for the pattern recognition problem, but also for pattern elimination.

In [16], the authors proposed a new prevention AIS, called *antibody layer*, to actively recognise and put down various Internet hackers and viruses, i.e., Internet antigens. Additionally, they implemented a number of security classes for the antibody layer to efficiently counteract Internet antigens according to system and network resources. The aim of the antibody layer is to timely pre-empt the intruder and quickly recover the system on the basis of mechanisms extracted from the IS.

Several works have been published [13,18,19] pursuing the problem of developing an artificial immune system that is distributed, robust, dynamic, diverse and adaptive, with applications to computer network security. In these AIS, the several immune system cells and molecules were simplified by the definition of a basic type of detector that combined useful properties from these elements. The detectors were represented by bit strings in a binary Hamming shape-space. Detection was performed by a string match process that took into account the number of *r*-contiguous bits between two strings. The definition of self was performed by the negative selection algorithm described in Section 2.2. The maturation of naive detectors into memory detectors, together with the negative selection, was responsible for the learning part of the system.

Based on the works above, other authors [28] have also been trying to develop a network intrusion detection system inspired in the immune system. The authors reviewed and assessed the analogy between the vertebrate immune system and network intrusion detection systems. They aimed at unravelling the significant features of the IS that would be successfully applied to the task of detecting intrusions in computer networks. In a later work [27], the authors proposed that a hybridisation of negative selection with a clonal selection algorithm could result in more powerful AIS for network intrusion detection.

Framed on an agent-based paradigm, other authors [7] proposed a system for intrusion/anomaly detection and response in networked computers. In his approach, the immunity-based agents roamed around the nodes and routers monitoring the situation of the network. The most appealing properties of this system were mobility, adaptability and collaboration. The immune agents were able to interact freely and dynamically with the environment and each other.

4.2 Other Applications

In [2], the authors proposed an AIS aiming at integrating the distributed search of new agents and constraint relaxation among them. The authors applied a continuous immune network model, based upon a dynamic equation, to study the interactions among antibodies and among antibodies and antigens. The dynamic equation adopted took into account the stimulation and suppression among antibodies, their stimulation by antigens and a natural death rate.

In [33], the authors suggested that the pattern recognition task performed by the immune system has much in common with the aerial image segmentation problem. They used the negative selection algorithm to construct a set of detectors capable of recognising (detecting) everything but the desired class.

In [5], the authors proposed to apply the negative selection algorithm to detect novelties in time series data. They employed a binary Hamming shape-space to represent the elements of the system and the *r*-contiguous bit rule to determine the degree of recognition among the detectors and the encoded data. The authors reported results for two data sets: a simulated cutting dynamics of a milling operation and a synthetic signal. They observed that the number of *r*-contiguous bits chosen by the matching function was responsible for tuning the reliability of detection against the risk of false positives.

Hardware fault tolerance seeks to address the challenge of designing hardware systems that provide a high degree of reliability even in the presence of errors. The system must be protected from a variety of potential faults, manifesting in such forms as permanent stuck at faults or intermittent faults. In [3], the authors proposed what they called Immunotronics (immunological electronics) in order to implement a finite state machine based counter using immune principles. Their system relied upon the negative selection algorithm that was responsible for creating a set of tolerance conditions to monitor changes in hardware states. They employed a binary Hamming shape-space to represent the tolerance conditions.

A general form of a chemical reaction maps a set of reactants into a set of products. In [6], the authors used a binary Hamming shape-space, to describe each of the reactants and products for spectra recognition in chemical analysis.

Use was made of the immune network theory to produce a pattern recognition and classification system in [4]. This model consisted of T-cells, B-cells, antibodies, and an amino-acid library. The T-cells were used to control the production of B-cells. The B-cells would then compete for the recognition of the “unknowns”. The amino-acid library acts as a library of epitopes (or variables) currently in the system. When a new antigen (pattern) is introduced into the system, its variables are entered into this library. The T-cells then use the library to create their receptors that are used to recognise the new antigen. During the recognition stage of the algorithm, T-cells are matched against the antigen, and then a B-cell is created that match the antigen.

In [8], the authors proposed an artificial immune network model, summarised in Section 2.3, with the main goals of performing data clustering and filtering redundant data. An Euclidean shape-space model was used, in which the network units corresponded to antibodies and the input patterns were the antigens to be recognised and clustered. This network model was successfully applied to several clustering problems, including non-linearly separable tasks. Classification results comparable to supervised neural networks for the Iris data set of Fisher were presented in [9].

In [10], the authors applied the clonal selection algorithm of Section 2.2 to recognise a set of binary characters represented in a binary Hamming shape-space. This algorithm was then adapted to solve multi-modal optimisation tasks.

A version of clonal selection was used in [8] as an inspiration to develop a novel learning algorithm for a Boolean neural network. The resultant network, named ABNET (AntiBody NETwork), was applied to several binary and real-valued machine-learning and pattern recognition tasks. The results were compared to the self-organising feature map (SOM) introduced by [29], and to a pruning version of the SOM proposed in [12].

Work in [48] proposed an AIS that could be used for pattern discovery and classification in data. This AIS employed a number of high-level metaphors drawn from the immune system. These are: A B-cell is capable of recognising pathogens (antigenic

recognition); similar B-cells are linked together and these links form a network of B-cells (immune network theory); cloning and mutation operations are performed on B-cells (clonal selection and somatic hypermutation). A number of B-cells can be represented by an ARB (Artificial Recognition Ball) given the theory of shape space. The AIS evolves a network of ARBs that can be viewed via a specially developed tool aiVis [45]. This work was then investigated in [30] in an attempt to apply the AIS to a ore complex and large-scale domain. However, this work identified a different behavioural pattern not seen in the previous work, which in turn lead to a further investigation into the nature of the algorithm. The subsequent investigation discovered that the algorithm would naturally discover the strongest pattern within the data set that it was applied to. This new behaviour was deemed not to make a significant difference in the algorithms capability to discover patterns in data, but it was argued, enhances the usefulness of this algorithm.

5 AIS and ANN for Pattern Recognition

Similar to the use of artificial neural networks, performing pattern recognition with an AIS usually involves three stages: 1) defining a *representation* for the patterns; 2) *adapting* (*learning* or *evolving*) the system to identify a set of typical data; and 3) applying the system to *recognise* a set of new patterns (that might contain patterns used in the adaptive phase).

Refering to the three immune algorithms presented (negative selection, clonal selection, and immune network), coupled with the process of modelling pattern recognition in the immune system, as described in Section 3, this section will contrast AIS and ANN focusing the pattern recognition applications. Discussion will be based on computational aspects, such as basic components, adaptation mechanisms, etc. Common neural networks for pattern recognition will be considered, such as single and multi-layer perceptrons [40], associative memories [21], and self-organising networks [29]. All these networks are characterised by set(s) of units (artificial neurons); they adapt to the environment through a learning (or storage) algorithm, they can have their architectures dynamically adapted along with the weights, and they have the basic knowledge stored in the connection strengths [17].

Component: The basic unit of an AIS is an attribute string *s* (along with its connections in network models) represented in the appropriate shape-space. This string *s* might correspond to an immune cell or molecule. In an ANN, the basic unit is an artificial neuron composed of an activation function, a summing junction, connection strengths, and an activation threshold. While artificial neurons are usually processing elements, attribute strings representing immune cells and molecules are information storage and processing components.

Location of the components: In immune network models, the cells and molecules usually present a dynamic behaviour that tries to mimic or counteract the environment. This way, the network elements will be located according to the environmental stimuli. Unlike the immune network models, ANN have their neurons positioned in fixed pre-defined locations in the network. Some neural network models (e.g., [29]) also adopt fixed neighbourhood patterns for the neurons. If a network pattern of connectivity is not adopted for the AIS, each individual element will have a position in the population that might vary dynamically. Also, a *metadynamic* process might allow the introduction and/or elimination of particular units.

Structure: In negative and clonal AIS, the components are usually structured around matrices representing repertoires or populations of individuals. These matrices might have fixed or variable dimensions. In artificial immune networks and artificial neural networks, the components of the population are interconnected and structured around patterns of connectivity. Artificial immune networks usually have an architecture that follows the spatial distribution of the antigens represented in shape-space, while ANN usually have pre-defined architectures, and weights biased by the environment.

Memory: The attribute strings representing the repertoire(s) of immune cells and molecules, and their respective numbers, constitute most of the knowledge contained in an artificial immune system. Furthermore, parameters like the affinity threshold can also be considered part of the memory of an AIS. In artificial immune network models, the connection strengths among units also carry endogenous and exogenous information, i.e., they quantify the interactions of the elements of the AIS themselves and also with the environment. In most cases, memory is content-addressable and distributed. In the standard (earliest) neural network models, knowledge was stored only in the connection strengths of individual neurons. In more sophisticate strategies, such as constructive and pruning algorithms [31,39], and networks with self-adaptive parameters, the final number of network layers, neurons, connections, and the shapes of their respective activation functions are also part of the network knowledge. The memory is usually self-associative or content-addressable, and distributed.

Adaptation: Adaptation usually refers to the alteration or adjustment in the structure or behaviour of a system so that its pattern of response to other components of the system and to the environment changes. Although both evolutionary and learning processes involve adaptation, there is a conceptual difference between them. Evolution can be seen as a change in the genetic composition of a population of individuals during successive generations. It is a result of natural selection acting on the genetic variation among individuals. In contrast, learning can be seen as a long lasting change in behaviour as a result of previous experience. While AIS might present both types of adaptation, learning and evolution, ANNs adapt basically through learning procedures.

Plasticity and diversity: Metadynamics refers basically to two processes: 1) the recruitment of new components into the system, and 2) the elimination of useless elements from the system [46]. As consequences of metadynamics, the architecture of the system can be more appropriately adapted to the environment, and its search capability (diversity) increased. In addition, metadynamics reduces redundancy within the system by eliminating useless components. Metadynamics in the immune algorithms corresponds to a continuous insertion and elimination of the basic elements (cells/molecules) composing the system. In ANN, metadynamics is equivalent to the pruning and/or insertion of new connections, units, and layers in the network.

Interaction with other components: The interaction among cells and molecules in AIS occurs through the recognition (matching) of attribute strings by cell receptors (other attribute strings). In immune network models, the cells usually have weighted connections that allow them to interact with (recognise and be recognised by) other cells. These weights can be stimulatory or suppressive indicating the degree of interaction with other cells. Artificial neural networks are composed of a set (or sets) of interconnected neurons whose connection strengths assume any positive or negative values, indicating an excitatory or inhibitory activation. The interaction with other neurons in the network occurs explicitly through these connection strengths, where a single neuron receives and processes inputs from the environment (or network neurons) in the same or other layer(s). An individual neuron can also receive an input from itself.

Interaction with the environment: In pattern recognition applications, the environment is usually represented as a set of input patterns to be learnt, recognised, and/or classified. In AIS, an attribute string represents the genetic information of the immune cells and molecules. This string is compared with the patterns received from the environment. If there is an explicit antigenic population to be recognised (set of patterns), all or some antigens can be presented to the whole or parts of the AIS. At the end of the learning or recognition phase, each component of the AIS might recognise some of the input patterns. The artificial neurons have connections that receive input signals from the environment. These signals are processed by neurons and compared with the information contained in the artificial neural network, such as the connection strengths. After learning, the whole ANN might (approximately) recognise the input patterns.

Threshold: Under the shape-space formalism, each component of the AIS interacts with other cells or molecules whose complements lie within a small surrounding region, characterised by a parameter named *affinity threshold*. This threshold determines the degree of recognition between the immune cells and the presented input pattern. Most current models of neurons include a *bias* (or *threshold*). This threshold determines the neuron activation, i.e., it indicates how sensitive the neuron activation will be with relation to the input signal.

Robustness: Both paradigms are highly robust due mainly to the presence of populations or networks of components. These elements, cells, molecules, and neurons, can act collectively, co-operatively, and competitively to accomplish their particular tasks. As knowledge is distributed over the many components of the system, damage or failure to individual elements might not significantly deteriorate the overall performance. Both AIS and ANN are highly flexible and noise tolerant. An interesting property of immune network models and negative selection algorithms is that they are also self-tolerant, i.e., they learn to recognise themselves. In immune network models, the cells interact with each other and usually present connection strengths quantifying these interactions. In negative selection algorithms, the self-knowledge is performed by storing information about its complement.

State: At each iteration, time step or interval, the state of an AIS corresponds to the concentration of the immune cells and molecules, and/or their affinities. In the case of immune network models, the connection strengths among units are also part of the current state of the system. In artificial neural networks, the activation level of the output neurons determines the state of the system. Notice that this activation level of the output neurons takes into account the number of connection strengths and their respective values, the shape of activation functions and the network dimension.

Control: Any immune principle, theory or process can be used to control the types of interaction among the many components of an AIS. As examples, clonal selection can be employed to build an antibody repertoire capable of recognising a set of antigenic patterns, and negative selection can be used to define a set of antibodies (detectors) for the recognition of anomalous patterns. Differential or difference equations can be applied to the control of how an artificial immune network will interact with itself and the environment. Basically, three learning paradigms can be used to train an ANN: 1) supervised, 2) unsupervised, and 3) reinforcement learning.

Generalisation capability: In the AIS case, cells and molecules capable of recognising a certain pattern, can recognise not only this specific pattern, but also any structurally related pattern. This capability is attained by a process called *cross-reactivity* [41], and can be modelled using the affinity threshold. Any pattern lying in a ‘neighbourhood’ of a known pattern can be recognised by the same component of the AIS that recognise the

known pattern. Thus, a component of the AIS can generally recognise any other element whose affinity with is superior to ε. In addition to cross-reactivity, some immunologists (e.g. [23]), speculate that antibodies can also be multi-specific, in the sense that they can recognise antigens of relatively different structures, as far as enough interactions are established between them. Therefore, multispecificity contributes to the generalisation capability of AIS. ANNs are known to be efficient in generalising the training patterns, provided that an appropriate learning is performed. There are basically two ways in which an ANN can attain a satisfactory generalisation performance [38]: 1) by reducing the number of dimensions of the parameter space, or 2) by reducing the effective size of each dimension.

Non-linearities: Non-linearities in AIS appear basically in the use of activation functions that define the degree of recognition between two components of the system, proportionally to their affinity. As examples, a sigmoid or a simple threshold matching function might be used. Some immune network models [46] use Gaussian-like functions to make the maturation and proliferation probabilities dependent on the degree of connectivity of an immune cell with the current network configuration. Non-linearities in artificial neural networks reside basically in the activation functions of individual neurons. The ensemble operation of several non-linear neural units results in a network with great potentials to perform non-linear approximations and/or classifications.

6 Concluding Remarks

Artificial immune systems constitute an emergent biologically motivated computing paradigm. It is based upon the extraction of principles and metaphors from the immune system in order to design alternative computational tools to solve complex problems. Indeed, the main role of the immune system is to recognise what cells, molecules, and tissues belong to the organism and to distinguish them from the foreign elements. If the immune system were not so efficient in this self/nonself discrimination process, the body would have no problem with the rejection of graft tissues, for example. As a consequence, this great capability to recognise and eliminate specific patterns (nonself) serves as a good source of inspiration to develop novel computational paradigms for machine-learning and pattern recognition.

In this chapter three classes of artificial immune system algorithms to perform pattern recognition: 1) negative selection, 2) clonal selection, and 3) immune network models, have been reviewed. In negative selection, a pattern recognition system is designed by learning information about the complement set of the patterns to be recognised - a brand new paradigm. Clonal selection algorithms learn to recognise patterns through an evolutionary-like procedure. Finally, immune network models are peculiar because they carry information about the patterns to be recognised and, also, they have knowledge of themselves, i.e., a notion of self-identification. All algorithms are population based with the knowledge distributed among the components of the system.

The intuitive and appealing metaphor of engineering artificial immune systems to protect computers and networks of computers from viruses, unauthorised users, etc., led to the development of the so-called computational immunology. Most computational immunology algorithms, which compose particular cases of artificial immune systems, are based upon the negative selection algorithm. In the survey section of this chapter, the most influential works in computational immunology we reviewed. Additionally, the application of other models, including the immune network and clonal selection

algorithms, to other types of pattern recognition applications, such as character recognition, data analysis, clustering and classification were discussed.

The chapter followed with a theoretical comparison between artificial immune systems and neural network models for pattern recognition. Aspects such as the basic units composing each system, their respective types of adaptation mechanisms, the types of memory presented, and how they present generalisation capabilities were stressed.

There are also several works in the literature hybridising neural networks with artificial immune systems; a good review can be found in [8]. Although these were not included here due to a lack of space, these authors strongly believe that both approaches have much to profit from one another. Stretching speculations, it could be suggested that novel paradigms will soon emerge, such as *artificial neuroimmune systems*.

The aim of this chapter was to serve the purpose of introducing artificial immune systems to the neural network community, and also provided a basic guide to the literature. The algorithms presented could be directly employed and/or adapted as alternatives to solve the same types of pattern recognition problems as neural networks, or to complement their potentialities.

Acknowledgements

Leandro N. de Castro would like to thank the Computing Laboratory at UKC for the financial support.

References

[1] Ada, G. L. & Nossal, G. J. V. (1987), "The Clonal Selection Theory", *Scientific American*, 257(2), pp. 50-57.

[2] Aisu, H. & Mizutani H. (1996), "Immunity-Based Learning – Integration of Distributed Search and Constraint Relaxation", *Proc. of the IMBS'96*.

[3] Bradly, D. W. & Tyrrell, A. M. (2000), "Immunotronics: Hardware Fault Tolerance Inspired by the Immune System", *Lecture Notes in Computer Science*, 1801, pp. 11-20.

[4] Carter, J. H. (2000), "The Immune System as a Model for Pattern Recognition and Classification", *Journal of the American Medical Informatics Association*, 7(3), pp. 28-41.

[5] Dasgupta, D. & Forrest, S. (1996), "Novelty Detection in Time Series Data Using Ideas From Immunology", *Proc. of the ISCA'96*.

[6] Dasgupta, D., Cao, Y. & Yang, C. (1999), "An Immunogenetic Approach to Spectra Recognition", *Proc. of GECCO'99*, pp. 149-155.

[7] Dasgupta, D. (2000), "An Immune Agent Architecture for Intrusion Detection", *Proc. of GECCO'00 – Workshop Proceedings*, pp. 42-44.

[8] De Castro, L. N., & Von Zuben, F. J., (2001), "The Construction of a Boolean Competitive Neural Network Using Ideas From Immunology", *International Journal of Neurocomputing* (in print).

[9] De Castro, L. N., & Von Zuben, F. J., (2001), "aiNet: An Artificial Immune Network for Data Analysis", In *Data Mining: A Heuristic Approach*, H. A. Abbass, R. A. Sarker, and C. S. Newton (Eds.), Idea Group Publishing, USA, Chapter XII, pp. 231-259.

[10] De Castro, L. N. & Von Zuben, F. J. (2000), "The Clonal Selection Algorithm with Engineering Applications", *GECCO'00 – Workshop Proceedings*, pp. 36-37.

[11] De Castro, L. N. & Von Zuben, F. J. (1999), "Artificial Immune Systems: Part I – Basic Theory and Applications", *Technical Report – RT DCA 01/99*, p. 95.
[12] De Castro, L. N. & Von Zuben, F. J. (1999), "An Improving Pruning Technique with Restart for the Kohonen Self-Organizing Feature Map", *Proc. of the IJCNN*, 3, pp. 1916-1919.
[13] Forrest, S., Somayaji, A & Ackley, D. H. (1997), "Building Diversity Computer Systems", *Proc. of the 6th Workshop on Hot Topics in Operating Systems*, pp. 67-72.
[14] Forrest, S., A. Perelson, Allen, L. & Cherukuri, R. (1994), "Self-Nonself Discrimination in a Computer", *Proc. of the IEEE Symposium on Research in Security and Privacy*, pp. 202-212.
[15] Forrest, S., Javornik, B., Smith, R. E. & Perelson, A. S. (1993), "Using Genetic Algorithms to Explore Pattern Recognition in the Immune System", *Evolutionary Computation*, 1(3), pp. 191-211.
[16] Gu, J., Lee, D., Park, S. & Sim, K. (2000), "An Immunity-Based Security Layer Model", *Proc. of GECCO'00*, Workshop on Artificial Immune Systems and Their Applications, pp 47-48.
[17] Haykin S. (1999), *Neural Networks – A Comprehensive Foundation*, Prentice Hall, 2nd Ed.
[18] Hofmeyr S. A. & Forrest, S. (2000), "Architecture for an Artificial Immune System", *Evolutionary Computation*, 7(1), pp. 45-68.
[19] Hofmeyr S. A. & Forrest, S. (1999), "Immunity by Design: An Artificial Immune System", *Proc. of GECCO'99*, pp. 1289-1296.
[20] Holland, J. H. (1975), *Adaptation in Natural and Artificial Systems*, MIT Press.
[21] Hopfield, J. J. (1984), "Neurons with Graded Response Have Collective Computational Properties Like Those of Two-State Neurons", *Proc. Natl. Acad. Sci. USA*, 81, pp. 3088-3092.
[22] Hunt, J. E. & Cooke, D. E. (1996), "Learning Using an Artificial Immune System", *Journal of Network and Computer Applications*, 19, pp. 189-212.
[23] Inman, J. K. (1978), "The Antibody Combining Region: Speculations on the Hypothesis of General Multispecificity", *Theoretical Immunology*, (Eds.) G. I. Bell, A. S. Perelson & G. H. Pimbley Jr., Marcel Dekker Inc., pp. 243-278.
[24] Jerne, N. K. (1974), "Towards a Network Theory of the Immune System", *Ann. Immunol.* (Inst. Pasteur) 125C, pp. 373-389.
[25] Kephart, J. O., Sorkin, G. B. & Swimmer, M. (1997), "An Immune System for Cyberspace", *Proc. of the IEEE SMC'97*, pp. 879-884.
[26] Kephart, J. O. (1994), "A Biologically Inspired Immune System for Computers", R. A. Brooks & P. Maes (Eds.), *Artificial Life IV Proceedings of the Fourth International Workshop on the Synthesis and Simulation of Living Systems*, MIT Press, pp. 130-139.
[27] Kim, J. & Bentley, P. (1999), "Negative Selection and Niching by an Artificial Immune System for Network Intrusion Detection", *Proc. of GECCO'99*, pp. 149-158.
[28] Kim, J. and Bentley, P. J. (2001), "Towards an Artificial Immune System for Network Intrusion Detection: An Investigation of Clonal Selection with a Negative Selection Operator", *the Congress on Evolutionary Computation (CEC-2001), Seoul, Korea, May 27-30, 2001*
[29] Kohonen T. (1982), "Self-Organized Formation of Topologically Correct Feature Maps", *Biological Cybernetics*, 43, pp. 59-69.

[30] Knight, T and Timmis, J. (2001). "AINE: An immunological approach to data mining". To appear in Proceedings of IEEE International Conference on Data Mining. San Jose, CA. November.
[31] Kwok T. Y., & Yeung. D. Y. (1997), "Constructive Algorithms for Structure Learning in Feedforward Neural Networks for Regression Problems", *IEEE Trans. on Neural Networks*, 8(3), pp. 630-645.
[32] Langman, R. E. & Cohn, M. (1986), "The 'Complete' Idiotype Network is an Absurd Immune System", *Imm. Today*, 7(4), pp. 100-101.
[33] McCoy, D. F. & Devaralan, V. (1997), "Artificial Immune Systems and Aerial Image Segmentation", *Proc. of the SMC'97*, pp. 867-872.
[34] Nossal, G. J. V. (1994), "Negative Selection of Lymphocytes", *Cell*, 76, pp. 229-239.
[35] Nossal, G. J. V. (1993), "The Molecular and Cellular Basis of Affinity Maturation in the Antibody Response", *Cell*, 68, pp. 1-2.
[36] Okamoto, T. & Ishida, Y. (1999), "Multiagent Approach Against Computer Virus: An Immunity-Based System", *Proc. of the AROB'99*, pp. 69-72.
[37] Perelson, A. S. & Oster, G. F. (1979), "Theoretical Studies of Clonal Selection: Minimal Antibody Repertoire Size and Reliability of Self-Nonself Discrimination", *J. theor.Biol.*, 81, pp. 645-670.
[38] Prechelt L. (1998), "Automatic Early Stopping Using Cross Validation: Quantifying the Criteria", *Neural Networks*, 11(4), pp. 761-767.
[39] Reed R. (1993), "Pruning Algorithms – A Survey", *IEEE Trans. on Neural Networks*, 4(5), pp. 740-747.
[40] Rumelhart, D. E., McClelland, J. L. & The PDP Research Group, Eds. (1986), *Parallel Distributed Processing*, Cambridge MIT Press.
[41] Smith, D. J. (1997), "The Cross-Reactive Immune Response: Analysis, Modelling and Application to Vaccine Design", Ph.D. Dissertation, Computer Science, The University of New Mexico.
[42] Somayaji, A., Hofmeyr, S. A. & Forrest, S. (1997), "Principles of a Computer Immune System", *Proc. of the new Security Paradigms Workshop*, pp. 75-81.
[43] Timmis, J. (2000), *Artificial Immune Systems: A Novel Data Analysis Technique Inspired by the Immune Network Theory*, Ph.D. Dissertation, Department of Computer Science, University of Wales, September.
[44] Timmis, J and Neal, M. (2001), "A resource limited artificial immune system". *Knowledge Based Systems*, 14(3-4). Pp 121-130.
[45] Timmis, J. (2001). "aiVis: Artificial Immune Network Visualisation". In *EuroGraphics UK 2001 Conference Proceedings*, pages 61-69, University College London, April 2001. Eurographics.
[46] Varela, F. J., Coutinho, A. Dupire, E. & Vaz, N. N. (1988), "Cognitive Networks: Immune, Neural and Otherwise", *Theoretical Immunology*, Part II, A. S. Perelson (Ed.), pp. 359-375.