Computing with Idiotypic Networks

Francisco Martins[1] and Neva Slani[2]

[1] Departamento de Matemática, Universidade dos Açores,
Rua da Mãe de Deus, Ponta Delgada, Portugal
[2] Katedra za matematiku i nacrtnu geometriju,
Fakultet strojarstva i brodogradnje, Ivana Lučića, Zagreb, Croatia

Abstract. The paper presents a computer experiment inspired on the immune metaphor and based on the work of Farmer, Packard, and Perelson [FPP86]. We develop a model influenced by the way the immune system works that is well-suited to address a class of NP-hard problems. We discuss the results obtained when applying the model to an artificial vision problem denoted *the museum problem*, where artificial agents successfully accomplish a surveillance assignment to protect the pieces of an art exhibition from bad behaved visitors.

1 Introduction

The immune system is a highly complex system responsible for the defense of our body (and of vertebrates in general) against foreign material (or *antigens*). This task is accomplished identifying each molecular shape encountered and mounting the adequate immune response. Therefore, it needs to distinguish the molecules of the body from antigens, *i.e.*, it must classify each molecule as *self* or *non-self*. The self–non-self classification problem is one of the major assignments of the immune system.

The model we present is inspired by the clonal selection theory introduced by Sir John MacFarlane Burnet [Jer55,Bur59,AN87] and in the idiotypic network hypothesis, introduced by Niels Jerne [Jer74]. These two hypothesis provide the mathematical framework for describing the immune system as a dynamic system of interacting species.

From a computational point of view, the immune system is comparable to the neural system, namely, it possesses the capacity to learn and to memorise and recognise patterns, which are emergent properties of the idiotypic network and of the clonal selection theory.

In this paper, we exploit a model based on the immune system to solve a surveillance problem (artificial vision) of a museum, referred as *the museum problem*. The model follows Farmer *et al* [FPP86] and is used to train a population of artificial agents to detect bad behaved museum visitors.

The experiments we did with the developed model produced very promising results. We were able to evolve and maintain a population of artificial agents, well fitted to the environment, that classified successfully self and non-self material. For a complete discussion on the model and on the results we obtained please refer to [Mar00].

The paper is organised as follows. The next section presents a brief introduction to the immune system focusing on its dynamics and meta-dynamics. Section 3 is devoted to the experiment. We characterise the problem and explain the conducted computer simulation. Last section discusses the results we have achieved with the simulation of the model.

2 Dynamics and meta-dynamics of the Idiotypic Network

The model we present belongs to a class of models introduced by Farmer *et al* [FPP86] and emphasises the two major aspects of the immune system: its *dynamics* and its *meta-dynamics*. The *dynamics* describes the changes over time of the concentration of lymphocyte species (and secreted molecules). The *meta-dynamics* models the continuous process of clonal insertion and elimination of lymphocyte species in the population. Notice that the model does not intend to describe the immune system, but simply to exploit the computational aspects of it.

We model the characteristics of the idiotypic network using a system free of antigens, since these do not interfere directly in the idiotypic phenomena. Thus, the environment has no longer the role of training the system (by presenting antigens) and exerts the role of moderator. The repertoire adaptation is done by *reinforcement*, *i.e.* without receiving any information from the outside world (via antigens).

2.1 The artificial immune system

In order to define a computational model based on the immune system we must define its actors, namely, the antigenic environment, lymphocytes, antigens, and the interactions between them.

Since all molecular shapes are settled by polypeptidic chains and all interactions results from chemical bindings of such chains, we encode these two features in our model. We adopt the representation in [FPP86] and code the specificity of the immune actors as binary strings. Unlike Farmer we consider only an antigenic determinant for lymphocytes. Therefore, all the immune actors are coded similarly.

In what concerns molecular interactions, there must exists a big enough affinity between two molecules in order for them to react. We measure the molecular affinity between two artificial molecules as the complementarity of the binary strings that represent them. A natural choice is to consider that 0 and 1 are complementary units. So, complementarity between two binary stings is the result of sum of the logical *exclusive or* between them.

In the model we present, each lymphocyte species has an activation threshold, D_i, that sets its lower bound of reaction. If the distance between lymphocyte species i and another molecule is superior to D_i, then the lymphocyte is activated. This means that it is stimulated to clone itself and start an immune response.

2.2 Dynamics

Following [FPP86] we model the variation on the concentration of lymphocyte species as a dynamic system. The concentration of lymphocyte species i, denoted by x_i, is described by the following non-linear differential equation:

$$\dot{x}_i = x_i\left(k_1 S_i + k_2 F_i - k_3 x_i\right), \qquad x_i(0) = c_{min}.$$

The first summand $(k_1 S_i)$ is the *suppression coefficient* resulting from the idiotypic interactions (S_i is always a negative value). The second—the *fitness effect*—contributes to increase the species concentrations. At last, $k_3 x_i$ corresponds to a *death factor* and models the ageing of cells. The constants k_1, k_2, and k_3 allow us to fine tune the relative influence of each part in the equation. Value c_{min} is the initial concentration of each lymphocyte species.
Idiotypic interactions. Niels Jerne hypothesis is a central notion in our work and provides important intuitions on how the immune system regulates itself via idiotypic interactions. The idiotypic interaction measures the suppression coefficient of one species relatively to another. Its intensity reflects the pattern recognition intersection between species. Therefore, we model it as a function of the activation threshold and of the hamming distance $(d(i,j))$ between the bit strings representing species.

The suppression that lymphocyte species j exerts on i is denoted by m_{ij} and computed as

$$m_{ij} = \begin{cases} 0, & \text{if } D_i + D_j - d(i,j) \leq 0 \\ \frac{d(i,j) - D_i - D_j}{2D_i}, & \text{if } 0 < D_i + D_j - d(i,j) \leq 2D_i \\ -1, & \text{if } 2D_i < D_i + D_j - d(i,j) \end{cases}.$$

The suppression is minimal (0) when species i and j do not interfere (recognise) each other, and is maximal (-1) when species i is totally overcome by species j. When species intersect with each other, the suppression coefficient is proportional to the intersecting area $d(i,j) - D_i - D_j$.

Notice that the suppression value is not symmetric. The species with a bigger activation threshold exerts more suppression over the other one.

The idiotypic suppression that the network exercises on lymphocyte species i is given by

$$S_i = \sum_{j \neq i} m_{ij} x_j,$$

where x_j is the concentration of species j.
Fitness. The fitness measures the adaptability of a lymphocyte species to the antigenic environment. The balance between the activation threshold and

the distance to self material seems a reasonable fitness value, since the model rejects auto-immune species.

Let p_i denotes the bit string that codes the receptor for species i and e_j the coding of a self element. The fitness $f(i,j)$ between p_i and e_j is given by

$$f(i,j) = \frac{D_i}{d(p_i, e_j)}.$$

The fitness of a lymphocyte species i is computed as

$$F_i = \frac{\sum_{j \in \mathcal{P}} f(i,j)}{\#\mathcal{P}}.$$

where $\mathcal{P}$ denotes the set of self material.

2.3 Meta-dynamics

An essential aspect of the immune system is the constant adjustment of its immune repertoire, since the diversity of the antigenic environment is by far superior than the recognition capacity of the immune system. Therefore, besides the variation in the concentration of the species, the immunologic repertoire is itself dynamic. This phenomenon is referred as *meta-dynamics*.

The meta-dynamics of an idiotypic network consists in the study of the topological variations induced by inclusion—*recruitment*—and exclusion of lymphocyte species. The impact in the dynamics includes the change of the number of equations that describe the concentration flow and the idiotypic interactions.

To exploit this double plasticity of the immune system the meta-dynamic phase should occur sparsely when compared to dynamics variation. Hence, permitting the dynamic system to acquire some stability before suffering new perturbations caused by the variation of its population. The stabilisation of the system is important, since it quantitatively reflects each lymphocyte species' fitness and the suppression from the idiotypic network in its concentration.

Recruitment. The genesis of new lymphocyte species copies the biological processes of the immune system: (a) bone marrow production and (b) clonal selection. Bone marrow production of lymphocyte species is simulated by the generation of random sequences that represent their receptors. The goal is to maintain a diverse repertoire and to uncover new evolutionary streams.

Clonal selection, associated with high mutation rates, is an important mechanism in the plasticity of the immune system. We use processes inspired in biological procedures (genetic algorithms) to mate and produce new species from the immune repertoire, namely, selection techniques, crossover, and mutation.

The generation of new species, either by random means or using genetic operators, aims at the refinement of the immune repertoire, searching for new species that are more suitable to cope with the antigenic environment. The selection between both reproduction processes is random, but we give priority to the reproduction using genetic operators.

New lymphocyte species undergo a maturation phase that filters those that are auto-immune. We test new species against self material and eliminate the ones that react with self fragments (mimicking apoptose).

Elimination. The elimination of lymphocyte species allows the system to dispose species that were overcome by new ones. There are two ways in which a lymphocyte species may be eliminated: (a) it reveals as auto-immune or (b) it possesses a null concentration. The former models the removing (by apoptose) of the agents that interfere with self material. There may be some auto-immune agents in the population because the recruitment phase only shows a fragment of self material. So, when we detect an auto-immune agent it is removed immediately from the population. The latter contributes to an optimisation in the computational algorithm, since a species with no (or very low) concentration has no influence in the system.

3 An experiment with the immune system

This section defines the problem we want to address and explains the decisions we made when planning the computer simulation.

3.1 The museum problem

The problem we propose is an instance of a class of NP-hard problems that can be suitable addressed using techniques inspired in the immune system.

The *museum problem* is defined as follows. Imagine that we want to install a surveillance system on a museum visited regularly by different groups of people. A classical approach implies the recruitment of well-trained security officers, helped with surveillance devices such as cameras and motion detectors. But suppose that we want to perform this task using artificial agents that, after a training phase, become capable to accomplish the surveillance assignment with satisfactory results.

The proposed problem consists in learning the behaviour of groups of visitors during a *guided tours to the museum*, where artificial agents must be able to detect possible "non-normal" behaviours among the visitors. We want to develop an artificial vision system capable to analyse and classify behaviours as self or non-self. From an algorithmic point of view, the problem consists in selecting a population of artificial agents that are tolerant to self behaviours and are able to cover the entire antigenic space (recognise all non-self behaviours).

Notice that the algorithm starts with a population of agents generated randomly and optimises it over time.

The motivations to use an algorithm inspired in the immune system seem clear for problems as the one we present.

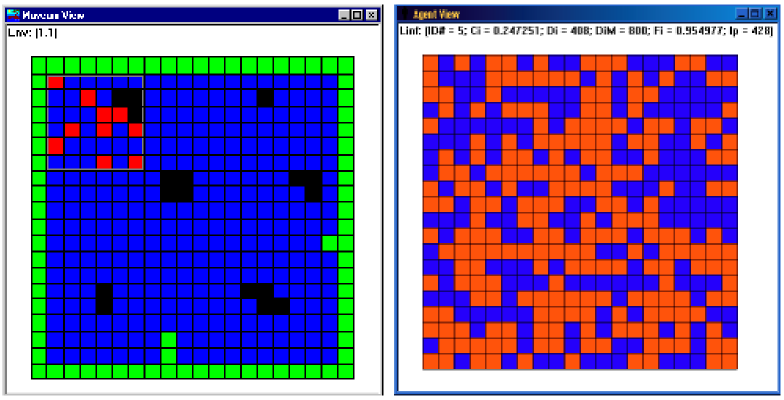
3.2 The experiment

The computer simulation we develop allow us to test the immune algorithm we propose, namely the adequacy of the algorithm select agents that learn and classify patterns.

Coding the environment. The antigenic environment we consider has dimension 20×20. Since there are three relevant subjects: the museum structure, the objects in exhibition, and the visitors, each bit string representing the self is coded as two bits of information per locus, meaning:

00,	an empty space;	01,	a wall;
10,	an object;	11,	a visitor.

To establish a link between the real and artificial environments we represent the self strings as matrices. Figure 1 depicts an environment as shown by the computer simulation.



(a) A self element. (b) An agent.

Fig. 1. Matrix representations of simulation actors.

Coding the artificial agents. Each lymphocyte is coded as a binary string that represents its immune specificity. For the experiment we use binary strings of the same dimension as for the environment. Each position is either 00 or 11. Figure 1 illustrates the matrix representation of an agent.

Coding the guided tour. The behavioural pattern that we learn is "a guided tour" to a museum room. Besides the representation of the room and the visitors, it is important to learn how the tour proceeds, *i.e.*, it is necessary to define the set of rules that govern the guided tour. The aim of the training phase is to select the agents that are able to learn this behaviour.

A guided tour consists in the movement of a group of people around a museum room (place where the exhibition is taking place). The group is fixed at the beginning of the simulation and moves always in the same way, avoiding possible obstacles. The group is always suited in a delimited area (listening to the explanations of the guide). In the present case, there are 10 visitors delimited by a 6×6 square.

Computing the dynamic equation. The non-linear, autonomous dynamic system referred in section 2 is not integrable by algebraic means. Therefore, we compute a numeric approximation using Euler's method that suffices for the study of its dynamics. The concentration of each lymphocyte species is approximated in each simulation step, using

$$x_i(t+1) = x_i(t) + hf(\mathbf{x}(t)).$$

Since we can take a discrete simulation time, and the concentration of all species are computed at each instant, we let $h = 1$. Hence, the approximate variation of the concentration of a lymphocyte species i between time t and $t+1$ is given by

$$x_i(t+1) = x_i(t) + x_i(t)\left(k_2 F_i - k_3 x_i(t) + k_1 \sum_{j \neq i} m_{ij} x_j(t)\right).$$

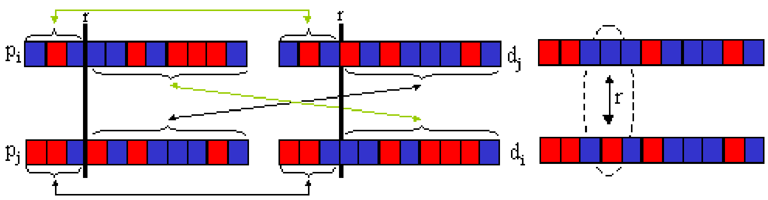
Generation of new lymphocyte species. The generation of new species sets three attributes: (1) the bit string that represents the receptor's molecules; (2) the initial concentration; and (3) the activation threshold.

The process inspired in the bone marrow produces agents with a random generated receptor bit string (the probability of the fragment 00 is equal to the probability of the fragment 11). The initial concentration and the activation threshold are set to the simulations parameters C_{in} and D_{in}, respectively, fixed along the simulation. The results we present were computed with $C_{in} = 0.2$ and $D_{in} = 408$.

The production of agents influenced by clonal selection is based on genetic algorithms, as described in [Gol89], identified as the *simple genetic algorithm*. The algorithm may be decomposed in two major parts: (1) random selection of the well-fitted agents and (2) production of new agents using genetic operators, namely crossover and mutation.

The selection of mates is random, but proportional to the concentration of each species. Therefore, dominating species are likely to be selected.

The operations of crossover and mutation are standard. For the former, we randomly pick a cutting position and generate two offsprings as illustrated in figure 2(a). The concentration of each offspring corresponds to that of the (genetically) closest parent, multiplied by its fitness coefficient. Thus, the offspring's concentration reflects the adequacy of its parents to the antigenic environment. Similarly, the activation threshold is computed from the activation threshold of the (genetically) closest parent incremented with the



(a) Genetic crossover schema.

(b) Genetic mutation.

Fig. 2. Genetic operators used for the production of lymphocyte species.

simulation parameter D_d. In the simulation we use $D_d = 4$. As for mutation (figure 2(b)), we apply it simultaneously with crossover. When cloning the parent strings we make some “mistakes” and copy a 00 as 11 or vice-versa. The probability of a mutation that we use is 0.02, which corresponds to around eight fragments per crossover. Mutation is needed to bring diversity to the population.

Recruitment tests. We run a *recruitment test* to evaluate the plausibility of brand new agents either in what concerns the recognition of self material or its acceptance by the idiotypic network. We perform a test similar to what happens in the thymus (*c.f.* [BS93]). So, every time a new agent is generated we test it against self elements. If the agent recognises some self material, it fails the test and is eliminated. The inclusion of this test improved the results we obtained.

Elimination of agents. The elimination of an agent happens when it reveals auto-immune or when its concentration falls below some minimum threshold (defined by $C_{min} > 0$). In any case, the suppression of some species from the idiotypic network removes all the interaction both with the antigenic environment and the idiotypic network.

3.3 The immune algorithm

The immune algorithm describes the interplay between the dynamics and meta-dynamics of the system. The algorithm is inspired by the immune metaphor and combines ideas from the John Holland’s *classifier system* described in [Gol89,Hol75]. Moreover, Farmer *et al* [FPP86] settle a parallel between the immune system and the classifier system.

Figure 3 depicts the immune algorithm we use in the simulation software. There is a generation of an initial population before the iterative process begins, observing the assumptions of the clonal selection and idiotypic network theories.

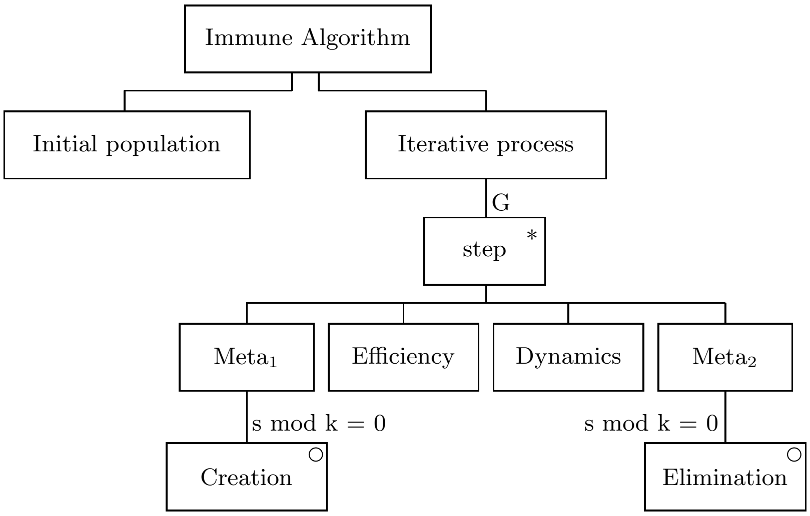


Fig. 3. Michael Jackson's diagram describing the immune algorithm.

The iterative process amounts four major steps. Meta-dynamics is divided into $Meta_1$ and $Meta_2$ steps that perform introduction and elimination of new species, respectively. Notice that the meta-dynamics runs only on each k^{th} step. Processes *efficiency* and *dynamics* carry out, as their names suggest, the updating of efficiency and of the concentration of each species.

This straightforward algorithm is the core of the simulation software we use to test the model. Next section presents the results.

The pattern learning that took place during the training phase is based on a technique called *negative selection* [FPAC95,D'h95,DFH96,D'h96], where the agents that recognise given patterns are discarded. Therefore, the agents are trained not to recognise any self material instead of being trained to recognise non-self material. This technique allows for the learning phase to be performed in an antigen free environment, meaning that, in the case of a surveillance system, the agents are trained in their security task without the presence of intruders.

4 Conclusions

The results we present correspond to a simulation for the parameters described above, with 7500 meta-dynamic steps. The meta-dynamics only takes place when two (artificial) stability conditions are met: (a) the maximal variation in the concentration of the species is inferior to a given threshold (0.001) or (2) after a finite number of tentatives to stabilise the dynamic system.

We discuss the evolution and the adaptation of the immune population during the training phase, and the classification ability of the population.

4.1 Analysis of the immune repertoire

The immune repertoire starts with 150 lymphocyte species and stabilises around 90 species. The existence of an (almost) constant number of individuals in the population reinforces the auto-control of the system. Neither the population grows in an uncontrolled way, nor one species dominates all the other species. The reduction in the number of agents is related to the influence of the idiotypic network. Agents become more and more adapted to the environment and therefore suppress similar ones.

In what concerns the fitness of the repertoire the values are between 94% and 98%. The values do not indicate a clear tendency to optimise as the simulation proceeds. This behaviour is explained by the constant renewal of the population with unfitted agents that are created randomly.

Another interesting indicator to examine is the diversity of the population. The measurements we perform indicate that the population remains diverse during the training phase. This reveals that the genetic operators we use maintain an heterogeneous population during simulation.

So, from a quantitative point of view the results are plausible in what concerns the number of agents, its diversity, and its fitness.

4.2 The cognitive evolution

Now we analyse the model from a qualitative perspective, *i.e.* the capacity of the selected repertoire to classify patterns as self or non-self. The results show the interaction between agents and non-self patterns generated randomly. Figure 4 shows the results of non-self recognition.

As for recognition of self material, the positive reactions are nearly zero. This happens because of the recruitment test we perform. Without this recruitment test some auto-immune agents were incorporated in the idiotypic network and the self recognition was around 10%. The test also added more stability to the number of elements of the population.

To test for non-self recognition we generate random patterns but keep track of the number of relevant positions (11) in the pattern, because we want to cover the entire antigenic space. The model is sensible to the size of the "perturbation" of a non-self element. It seems especially hard to identify non-self elements that posses the same number of relevant positions as the number of visitors used during the training phase. We show the results by perturbation level. Therefore, the response of the system can be measured with accuracy. We consider that the system triggers an alert if at least one agent in the population recognises a pattern.

The results are presented in figure 4. The 3D graphic shows a *time-number of points-recognition rate* relation. Therefore, we can observe, for a selected number of points, the evolution in the recognition of antigens of this class. Also, for a given moment in time, we can check how the immune repertoire covers the antigenic space.

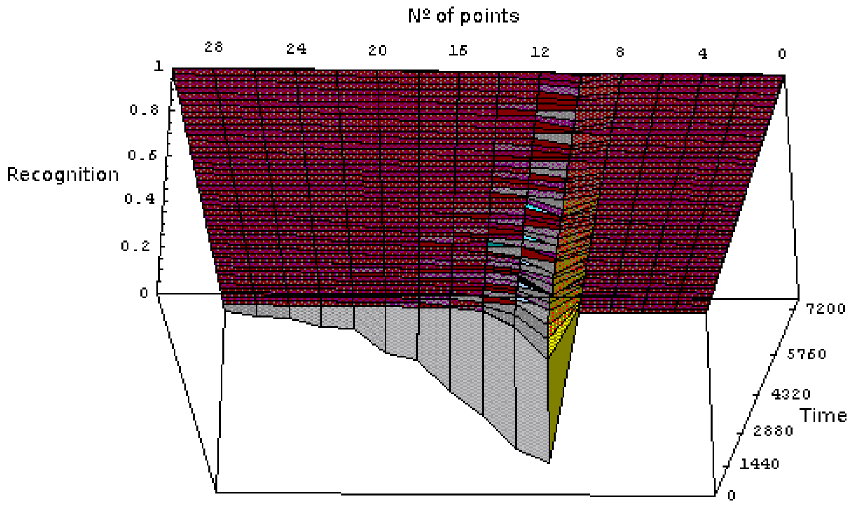


Fig. 4. Non-self recognition rates.

As time passes, the classification of the non-self material increases. The qualitative evolution of the repertoire is interesting in the class [10], which contains the same number of relevant points as the number of visitors. The figure reveals a very good evolution in the selection of the adequate population. Notice that the results for recognition of non-self material are near 100% for all the antigenic space.

As an overall we conclude that the model is suited for NP-hard problems with this kind of configuration. Both the qualitative and the quantitative behaviours allow us to conclude that it is possible to select a stable set of artificial agents that covers an antigenic space, in the present case a museum exhibition, and that we are able to distinguish between normal behaviours (the guided tour to the museum) and any non plausible behaviour.

References

[AN87] Gordon L. Ada e Sir Gustav Nossal. The clonal-selection theory. *Scientific American*, 257(2):50–57, 1987.

[BS93] Hugues Bersini e Gregory Seront. Optimizing with the immune recruitment mechanism. Tech. report TR/IRIDIA/93-6, Institut de Recherches Interdisciplinaires et de Développements en Intelligence Artificielle, 1993.

[Bur59] Frank MacFarlane Burnet. *The Clonal Selection Theory of Acquired Immunity.* Vanderbilt University Press, Nashville, Tennessee, 1959.

[DFH96] Patrik D'haeseleer, Stephanie Forrest e Paul Helman. An immunological approach to change detection: Algorithms, analysis and implications. In *IEEE Symposium on Research in Security and Privacy*, 1996.

[D'h95] Patrik D'haeseleer. A change-detection method inspired by the immune system: Theory, algorithms and techniques. Tech. report CS95-6, Department of Computer Science, University of New Mexico, 1995.

[D'h96] Patrik D'haeseleer. An immunological approach to change detection: Theoretical results. In *The 9th IEEE Computer Security Foundations Workshop*, Dromquinna Manor, County Kerry, Ireland, 1996.

[FPAC95] Stephanie Forrest, Alan S. Perelson, Lawrence Allen e Rajesh Cherukuri. A change-detection algorithm inspired by the immune system. *IEEE Transactions on Software Engineering*, 1995.

[FPP86] J. Doyne Farmer, Norman H. Packard e Alan S. Perelson. The immune system, adaptation, and machine learning. *Physica D*, 22:187–204, 1986.

[Gol89] David E. Goldberg. *Genetic Algorithms in Search, Optimization & Machine Learning*. Addison-Wesley, 1989.

[Hol75] John Holland. *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence*. MIT Press, 4^{th} edition, 1975.

[Jer55] N.K. Jerne. The natural selection theory of antibody formation. *Proceedings of the National Academy of Sciences USA*, 41:849–856, 1955.

[Jer74] N.K. Jerne. Towards a network theory of the immune system. *Ann. Immunol. (Inst. Pasteur)*, 125(C):373–389, 1974.

[Mar00] Francisco Martins. Computation with the Immune System. Masters thesis, Universidade dos Açores, 2000.