Proceedings of the 5th WSEAS Int. Conf. on Artificial Intelligence, Knowledge Engineering and Data Bases, Madrid, Spain, February 15-17, 2006 (pp80-84)

DESIGN OF A COMPUTACIONAL MODEL OF ADAPTATIVE COEVOLUTION

Alberto Pérez Alonso
CAP DE VILLAVERDE
CONSEJERÍA DE EDUCACIÓN
Calle Fernando Ortiz, 2
SPAIN

Abstract: The purpose of the immune system is to protect our bodies from infection. The system works by recognizing the molecular signature of microbes or viruses that attack our bodies, and once identified, eliminating the foreign molecules in a variety of ways.

As in previous evolutionary computation models of the vertebrate immune system , our model is limited to the interaction between B-lymphocytes and antigens.

To represent these molecules using a genetic algorithm, one possibility is to make no distinction between their genotype and their phenotype, that is, simply to represent both antigens and antibodies as binary strings.

Key-Words: Markov´s chains, Boltzmann tournament selection, immune system, antigen, antibody, simulated annealing.

Summary

Here we take a different aproach in which we use a simple model of one of the recognition processes occurring within the vertebrate immune system. The motivation behind this approach is that the immune system has a highly devoloped ability to discriminate between self and non-self, that is, to distinguish between the vast array of molecules that are an integral part of our bodies and foreign molecules.

The purpose of the immune system is to protect our bodies from infection. The system works by recognizing the molecular signature of microbes or viruses that attack our bodies, and once identified, eliminating the foreign molecules in a variety of ways. The immune system consist of two interrelated components: an innate defense component and an adaptive component. Here we will focus on the adaptive component, which is responsible for acquired inmunity.

We call antigens the molecules capable of stimulating an acquired immune response. When the system is working properly, only foreign antigens will produce an immune response.

There are a number of ways antigens are recognized, depending on whether the foreign molecule is inside or outside one of our body´s own cells. It is the job of antibodies to recognize antigens that are located in our body fluid outside the cell boundary. Recognition by a B-lymphocyte occurs when one of its antibodies comes into contact with an antigen of complementary shape. If the B-lymphocyte recognizes an antigen, it develops into a plasma cell and begins excreting large quantities of the antibody. The antibody, now circulating freely in the serum, coats foreign molecules of like type and flags them for destruction. The flagged molecules may be consumed, for example, by scavenger cells such as macrophages.

In addition, the activated B-lymphocytes enter a phase of hypermutation. The effect is to create offspring that produces antibodies with an even greater affinity to bind to the specific type of foreign molecule under attack.

The antibody molecules are composed of two pairs of protein chains: the so-called heavy chains and light chains- The heavy chains are constructed from four families of genes called variable, diversity, joining and constant. While each of these gene families has a number of members, only one gene from each family is used in constructing the protein. The chosen gene from each family is not determined until the antibody is being formed, thereby enabling a few hundred genes to create thousands of different heavy chain types through combination. Similarly, the light chains are constructed from variable, joining and constant families. Because it is the specific combination of light and heavy chains that determines what form of antigen the antibody will recognize, the potential coverage is around 100 million distinct foreign molecules (cf. Forrest and Javornik 1993) [1]

Another type of white blood cell -called a T-lymphocyte- is able to recognize foreign molecules, such as viruses, that take up residence within the body of our own cells. This is a more complex recognition process in which protein fragments (peptides) of the invader are carried to the surface of the cell in which they are hiding by the molecule major histocompatibility complex (MHC). The T-lymphocytes display receptors on their surface that are sensitive to a specific peptide-MHC complex, and they are constructed and function similarly to the receptors on the surface of the B-cells. Therefore, once the foreign peptides are transported outside the cell membrane by MHC, the T-lymphocytes are able to recognize them and launch an attack. The specific nature of the attack depends on whether the peptide-MHC complex is recognized by a helper T-cell or a killer T-cell. The killer T-cells respond to the so-called class I MHC by attaching themselves to the infected cell and attacking it directly. The helper T-cells respond to the so-called class II MHC, which is only produced by macrophage cells, by sending out a chemical messenger called cytokines that stimulates the macrophage to destroy the parasite hiding within it. Helper T-cells also play a role in the stimulation of B-lymphocytes to begin secreting antibodies.

As in previous evolutionary computation models of the vertebrate immune system (cf. Forrest and Perelson 1990) [2], our model is limited to the interaction between B-lymphocytes and antigens. It evolves these entities with a coevolutionary genetic algorithm that uses binary strings to represent their genetic codes.

In biological system, antibodies and antigens are folded into complex three-dimensional shapes. The closer the complementary match between their shapes, the stronger the binding forces will be between them.

To represent these molecules using a genetic algorithm, one possibility is to make no distinction between their genotype and their phenotype, that is, simply to represent both antigens and antibodies as binary strings. Given this type of representation, the binding force between a particular antibody and antigen can be computed simply as a function of the similarity between their sequences of ones and zeros.

However, we use a slightly more complex schema representation for antibody phenotypes to enable some regions of the receptor protein chains to be ignored in determining their final geometric shape. This gives us the ability to model a range of antibodies form specialist, which can only bind to a specific antigen, to more general antibodies, which can bind to whole families of antigens that share common characteristics. In addiction to antibody genes, each B-lymphocyte has a “threshold gene” that represents the binding strength required to initiate an immune response.

We produce an antibody -represented as a trinary schema- from a binary pattern and mask gene. A mask bit of one generates a schema value equal to the corresponding pattern bit, while a mask bit of zero produces a “don´t care” schema value. The length of the pattern and mask genes depends on the complexity of the antigens the antibody must recognize. The genetic operator used is similar to selection Boltzmann tournament with simulated annealing.

In immune system (IS), the composition of population for a fixed generation depends on likelihood of earlier generations. IS are shaped efficiently by Markov`s chains.

A finite Markov´s chain (cf Lial & Miller, 1989) [3] is a process sequence, with finite result (states) in every process, and every state depends on the previous state. A Markov´s chain is depicted by a transition probability matrix that expresses likelihood between states. The transition matrix remains constant between processes. If the transition matrix reach the *k* process it generates a matrix with a likelihood of every result after *k* process, where the rows which leades to the result in matrix specific initial states and the columns generate final states. Every value in the result matrix will contain the likelihood of *k* iterates of an inicial state to a final state.

The real-valued activation threshold of the B-lymphocyte, in the range (0,1), is produced from an 8-bit threshold gene. There is no distinction made between the genotype and phenotype of an antigen.

In our model of coevolution, each species represents a population of B-lymphocytes in one of three emergent phases of development. During the first phase -which begins immediately after the species is created and continues until some of its B-cells are activated by antigens- no cell has a selective advantage over any other so they are all reproducing at a uniformly slow rate. Once some of the B-cells are activated by antigens, these cells begin rapidly reproducing -marking the beginning of the second phase. This is also a time when large changes in fitness occur as the crossover operator splices pieces of various successful B-cells together; in this phase used simulated annealing. Eventually, the population will converge to slight variations of the most highly fit B-cell and enter a third phase. Mutation is the dominant genetic operator during this third phase, and it will produce relatively slight changes in cell fitness.

The B-cell development phases of our model differ somewhat in nature from those (cf. Holland, 1975) [4] Recall from the immune system overview that when an actual B-cell is activated, it enters a state of hypermutation called clonal selection. However, in both nature and our model the activation of a B-cell marks the beginning of a period of rapid change.

Evolutionary stagnation generally occurs after the most recently created species has entered its third phase. In the context of this model, problem descomposition consists of determining how many B-cells are required to cover a set of antigens, and which antigens will be recognized by which B-cells.

An example of application they are the problems of Optimization. Their main interest and maybe their great difficulty comes in and of itself determined to be, in its great majority, real problems that try to give answer to necessities so much economic as social, etc. (cf. Koza, 1992) [5]

A type of problems of optimization that have a special interest belongs to Combinatory Optimization.

The example proposed in this model is based on the problem of the traveling salesman problem:

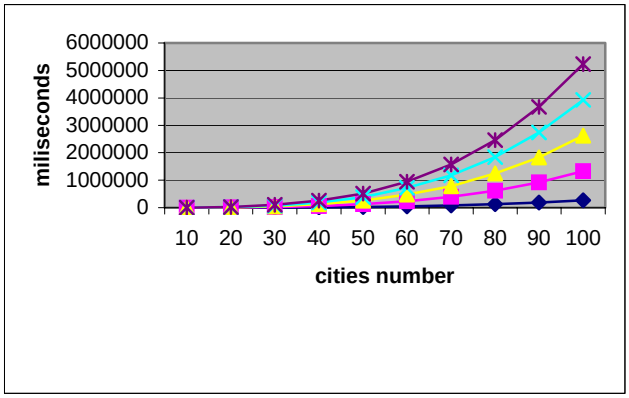
$$\text{Min} \sum_{i \in N} \sum_{j \in N} c_{ij} x_{ij}$$

$$\sum_{j \in N} x_{ij} = 1, \quad \forall i \in N$$

$$\sum_{i \in N} x_{ij} = 1, \quad \forall j \in N$$

$$\sum_{i \in S} \sum_{j \in S} x_{ij} \leq |S|-1, \ \forall S \subset N, S \neq \varnothing$$

$$x_{ij} \in \{0, 1\} \quad , \forall i, j \in N$$



References:

[1] Forrest, S., B. Javornik, R. E. Smith, and A. S. Perelson (1993). Using genetic algorithms to explore pattern recognition in the immune system. *Evolutionary Computation 1 (3)*, 191-211.

[2] Forrest, S. and A. S. Perelson (1990). Genetic algorithms and the inmune system. In H. P. Schewefel and R. Mäner (Eds.), *Parallel Problem Solving from Nature*, pp. 320-325. Springer-Verlag.

[3] Lial, M. L. and C. D. Miller. (1989). *Finite mathematics.* Glenview, IL: Scott, Foresman.

[4] Holland, J. H. (1975). *Adaptation in Natural and Artificial System.* University of Michigan Press.

[5] Koza, J. R. (1992). *Genetic Programming. On the programming of computers by means of natural selection.* The MIT Press.