Modelling the Immune System

Catherine Beauchemin, Department of Physics, University of Alberta

December 20, 2002

Contents

1 Introduction

Self-organization, or *emergence*, occurs in systems composed of many extremely simple agents which interact through very simple rules to collectively exhibit extremely complex behaviour. In such systems, the intelligence is neither contained in the individual agents, nor in a centralized information structure. Rather, the intelligence or complexity of those systems arises from the multiple interactions between the simple agents. A good example of such systems is droplets accumulating slowly with simple rules giving rise to the intricate patterns observed in snowflakes. The immune system is also a perfect example of such systems. It is made of many types of agents, for instance B cells, T cells, etc., which interact directly, e.g. a B cell binding with an antigen, or through their environment, e.g. an activated T cell releasing cytokines that trigger the division of neighbouring B and T cells.

Self-organizing systems (SOSs) are most often modelled using cellular automata (CA) as opposed to differential equations. What makes differential equations sometimes preferable

is that a lot is known about their behaviour. For instance, given a set of parameters and a particular set of differential equations, one can often predict the behaviour of the system, and tell for instance, for which range of the parameters the system is in a stable or an oscillatory state and for which value of the parameters the system will bifurcate from one state to the other (see [22] for a good introduction to the theory of dynamical systems and differential equations). However, modelling SOSs using differential equations is rarely an advantageous approach as the set differential equations necessary to represent such systems is often too complex for an analytical solution to be solved, either because the set of equations is too large or because pathological nonlinearities or delays render the system analytically unsolvable. In such a case, one has to resort to further simplifications of the studied models and those simplifications are often of mathematical nature rather than based on the real system. In other words, simplifications are made to render the set of equations solvable based on mathematical concerns rather than based on the exploitation of redundancies or other such aspects that are inherent to the nature of the real system. For cases such that the differential equations can no longer be solved analytically, CA are a preferable modelling approach. The penalty for using the latter method is that CA are a fairly new method and although they provide an easy way to model SOSs, they give very little information on the landscape of the global dynamics in the parameter space. This information has to be painstakingly extracted through averaging many simulations over the whole parameter space. On the other hand, CA are extremely simple to implement and adding nonlinearities, delays or additional compartments does not introduce any new difficulties in solving the model. Further, using a CA it is possible to represent the components and processes of interest in biological language so that the approximations made to simplify the simulations are usually more biological in character than mathematical, yielding a much more faithful model. Finally, CA by nature emphasize the local dynamics of a system rather than representing averages of its dynamics as is the case with differential equations. This allows one to witness the subtle rules at play in the model and it is a good starting point to understanding the impact of local interactions on the resulting global dynamics of the complex system. For these reasons, this work focuses on efforts to model the immune system using CA rather than differential equations.

In Section 2, basic concepts of immunology are introduced for the benefit of the uninitiated reader. In Section 3, the advantages of choosing CA over differential equations to model the immune system are discussed. In Section 4, a particular class of CA models of the immune system, the bit-sting models, are presented and the particular case of the Celada-Seiden bit-string model is described. In Section 5, other interesting efforts of particular interest to CA modelling of the immune system are reviewed and criticized. Finally, in Section 6, suggestions for the investigation of possible weaknesses in current bit-string model implementations as well as plans for possible applications of such model to specific immunological problems are made.

2 The Immune System, Briefly

In order to be able to model any SOS adequately, one has to learn what constitutes the system of interest and what rules govern the movements and actions of these constituents. In the first step of this project, the basics of immunology had to be learned. Here, a brief overview of the crucial components of the immune system will be introduced. This is a good chance for the uninitiated reader to collect some of the immunological knowledge necessary to understand the models that will be introduced later. For a great overview of the immune system from the point of view of a physicist, see [18] and for an in-depth coverage of the subject, see [10].

Antigens come in all shapes and sizes among which are foreign molecules, viruses, fungi and parasites. To the immune system, the antigen is the enemy to be recognized and destroyed. To fight any antigen efficiently, the vast army of cells and structures that constitute the immune system is divided in two classes: the *nonspecific* (or innate) system and the *specific* (or acquired, adaptive) system. The nonspecific system consists of anatomical barriers (skin, eyebrows, etc.), secretions (saliva, tears, etc.) and phagocytic cells such as macrophages, neutrophils and natural killer (NK) cells. As its name suggests, the defence mechanisms it carries are not specific to a particular antigen. In general, most of the microorganisms encountered by a healthy individual are readily cleared within a few days by the nonspecific defence mechanisms of the nonspecific system even before the specific immune system gets activated. It provides the first line of defence against any exposure to an antigen and will often serve as a backup force while the specific immune system builds up its specificity during a primary invasion, i.e. a first encounter of the system with a given pathogen. However, when an invading microorganism eludes the nonspecific system or is not cleared by it, the specific immune response of the specific system is triggered.

The specific immune system is mainly composed of lymphocytes (such as B and T cells), antibody molecules and other molecules produced by the lymphocytes. As its name suggests, this system has to be tailored to the specific intruder in order to be able to fight it efficiently. The specificity comes from the pattern recognition capabilities of the system. Both B and T cells possess receptor molecules on their surfaces that can recognize antigen or more specifically the antigen's recognizable part called the epitope. The B cell receptor is an antibody (immunoglobin receptor) embedded in the membrane of the cell while the T cell receptor is simply called a T-cell receptor (TCR). Figure 1 presents a schematic view of the B and T cells' receptor. Recognition of an epitope by a receptor occurs at a molecular level and is based on the length of the complementary regions between the strings typically made of amino acid[1] that constitute the receptor and the epitope. The particular amino acid string that composes a receptor determines the receptor's idiotype. Immunologists sometimes refer to the idiotype of a receptor as the "shape" of the receptor. It is important to point out that all the receptors on the surface of a given lymphocyte are made of the same amino acid chain

[1]B cell receptors are mostly made of amino acids. However carbohydrates, lipids and even nucleic acids can serve as B cell epitopes. T cell epitopes are almost always made of peptides (i.e. a string of amino acids). For the purpose of our discussion, we will assume that the building blocks of all receptors and epitopes are amino acids.

and hence are said to all be of the same idiotype. Receptors are constructed by a complex genetic process that insures that the receptors expressed on different lymphocytes have a different randomly chosen idiotype hence maximizing the chance that any random antigen presented to the system can be recognized given that the repertoire of receptors is complete. For a clever calculation and discussion on the completeness of the immune repertoire, see [18]. The B cell receptor interacts with epitopes present on intact antigen molecules that may be soluble or bound to a surface. But since the function of a T cell is to kill or stimulate other cells, it needs to be able to recognize that it is interacting with a cell rather than a soluble molecule. To accomplish this, T-cell receptors are designed to only recognize antigen when they are bound to a cell surface molecule called a major histocompatibility complex (MHC). MHC molecules come in two classes: class I molecules which specialize in presenting proteins synthesized within the cells and are found on every cell (referred to as target cells in this context) and class II molecules which specialize in presenting fragments of molecules picked up from the environment and are found only on professional antigen-presenting cells (APCs) like B cells, macrophages and dendritic cells. Both classes of MHC molecules bind peptides and present them to T cells. MHC class I typically presents peptides to cytotoxic T cells (Tc or CTL for cytotoxic T lymphocyte) which then destroy the cell if it can bind the peptide-MHC class I complex. MHC class II typically presents peptides to T helper cells (Th) which will then secrete cytokines, proteins that regulate the intensity and duration of the immune response by exerting a variety of effects on lymphocytes and other immune cells. For example, once a B cell has successfully recognized an antigen and has been successfully bound by a Th cell, co-stimulation will occur along with the release of cytokines which will initiate the division and differentiation of both the B and Th cells into effector and memory cells. In the case of B cells, the effector cells are called plasma B cells and secrete large amounts of antibody of their idiotype. The antibodies then serve as tags with which the immune system labels cells and molecules as foreign. An antibody-antigen complex or a foreign cell with antibodies attached to it is quickly eaten by large phagocytic cells such as macrophages. Note that immune responses can be divided into two branches: the humoral and the cell-mediated response. The humoral response's main effectors are the antibodies and such responses comprise the interaction of B cells with antigens and their subsequent proliferation and differentiation into plasma and memory B cells, with the former producing great amounts of antibodies capable of binding an antigen to tag it for removal. The cell-mediated response's main effectors are T cells, both Th cells and CTLs, with the former secreting cytokines enabling various phagocytic cells to phagocytose and kill microorganisms more effectively and the latter killing altered (e.g. virus-infected) self-cells.

The immune system constitutes only a few percent of the total cells in the body. During an immune response, the lymph nodes may swell to allow some increase in lymphocyte populations but that percentage cannot increase very much before affecting other bodily functions. Because of this size constraint on the immune system, specific strategies had to be adopted. Those include clonal selection, learning and memory as well as self-nonself discrimination. On the one hand, diversity of the receptors is necessary in order to maximize the chances of the system to be able to recognize any random antigen presented to it. On

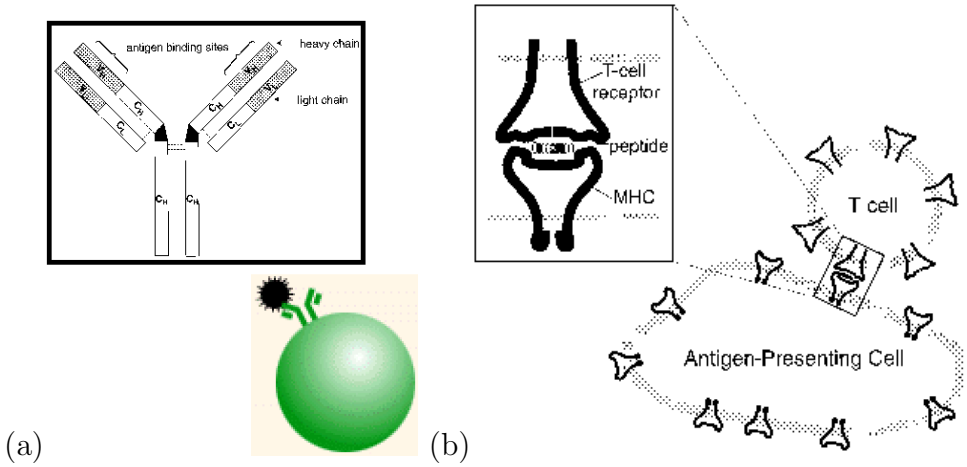

Figure 1: The schematic of the B (a) and T (b) cells' receptor.

the other hand, affinity, i.e. the strength of the binding between two interacting agents of the specific immune system — a function of the compatibility between the cells' epitope and receptor — is also an important asset since the higher the affinity, the stronger the response. Diversity versus affinity is then the complex balance that the immune system has to always carefully maintain. One way the system accomplishes this is through clonal selection, the process by which only those cells that recognize the antigen are allowed to differentiate thus being selected against those which do not. Clonal selection operates on both B and T cells, resulting in affinity maturation, i.e. the increase in average antibody affinity. Lymphocytes fight among themselves to bind an antigen resulting in the natural selection of clones (lymphocytes of a given idiotype) with a higher affinity for the antigen. For example, if two T cells can bind an antigen presented by an APC, the one that can bind it better will generate a stronger response resulting in a larger clonal population carrying its idiotype. Note that when activated lymphocytes differentiate, point mutations (a mistake in the amino acid string) may be introduced in the receptor type of its clones such that their idiotype is slightly changed. Point mutation can result in the creation of higher affinity clones and can be seen as the perturbation used by the immune system to explore the affinity landscape to find higher affinity clones.

The specific system has also developed a capacity for learning and memory. Learning occurs through the process of clonal selection during which the lymphocytes that have proven themselves to be valuable by having recognized antigens see their population size increase. This introduces a bias of the repertoire from random towards a repertoire that more clearly reflects the actual antigenic environment. Upon *primary invasion*, the adaptive immune system may take days to respond, and as a result the attack will mainly be conducted by the nonspecific system and by *low affinity* cells of the specific system. But upon *secondary invasion*, the immune system's response to the antigen has a larger amplitude and is faster than that of the primary invasion, suggesting that the system maintains some form of memory

of the antigens it encounters. Memory, like many other topics in immunology, is still not fully understood but it is likely the result of a combination of factors. It can be due to the fact that upon secondary infection with the same antigen, the repertoire that has been biased through affinity maturation during the first invasion now contains larger populations of high affinity clones, thus avoiding the delays seen in the primary invasion due to the fact that the cell population had to be enlarged before substantial amounts of antibody could be secreted. Here, as in a typical predator-prey situation, the size of the lymphocyte subpopulation with high affinity for a specific antigen relative to the size of the antigen population is crucial in determining the outcome of infection. Memory can also be attributed to differences between naive cells and cells that have already encountered the antigen. For instance, memory cells may be easier to trigger than naive cells. Cells that have never been triggered are small and contain little cytoplasm. When triggered they make at least 50 new proteins. One would expect many of these molecules to remain in the cell, thus making subsequent triggering events easier and faster [18].

While being able to recognize any random antigen that is presented to the immune system is a highly desirable feature, it does not come without a price. Self-nonself discrimination is perhaps the most formidable challenge the immune system has to face. It has to be able to recognize as many antigen as possible while constantly working to avoid autoimmunity, i.e. an immune response against self. For this, the immune system has developed a "training period" for immature lymphocytes. New lymphocytes are created in the bone marrow [2]. B cells are immediately put in circulation while the T cells migrate to the thymus to begin what the Immunologists call their "thymus education". There, they will encounter self as well and nonself cells and will be selected negatively for their ability to recognize self peptides presented by self MHC and positively for their ability to recognize nonself peptides presented by self MHC. It is also in the thymus that the T cells will evolve to typically become either T helper cells (Th) or cytotoxic T cells (Tc) by respectively expressing CD4 or CD8 membrane molecules on their surface. As for the B cells, since they typically require to be bound by Th cells in order to initiate differentiation into antibody producing plasma cells and memory cells, they are in a sense selected through Th cells selection since the unsuccessful B cells, i.e. B cells that are not bound by any Th cells, will eventually die without ever being activated. But nothing is perfect and the threat of an autoimmune disease is never fully eliminated.

3 Modelling approach

Computational modelling is of growing interest in a number of fields including biology, mathematics, physics and computer science. The growing attraction to such modelling techniques is clear: computers are getting evermore efficient and fast while the systems we are dealing with are getting evermore complex and diversified. Although computational biology has lately mainly concentrated on data-mining notably in the context of genomic research, bi-

[2]Lymphocytes are also created during an immune response when cytokines initiate cell division and differentiation. However, since those are created from successful lymphocytes clones, they are very unlikely to recognize self.

ological simulations have always been around and have recently been gaining momentum. Here, we will explore a few of the efforts made to model the immune system using cellular automata.

We have to distinguish between two types of computer modelling, namely differential equation models and cellular automaton (CA) models. As mentioned above, the former is often a preferable way to model a complex system since more is known about differential equations. Given a set of differential equations, it is often possible to deduce the global dynamics of the system for a any set of parameters and the parameters can in turn be tuned to drive the system into the desired global dynamics. However, when the system to be modelled is complex, which is often the case, the differential equations will likely contain nonlinearities and time-delays. Such differential equations are rarely analytically solvable and one can either choose to simplify the equations or try to map the behaviour of the system on its entire parameter space through painstaking exploration with the help of computer simulations. In the former case, one attempts to simplify the equations to render them solvable. The simplifications made are of mathematical concern and will most often distort the dynamics away from that of the real system. In the latter case, the effort required is great and is rarely rewarding as one does not have the benefits of using differential equations knowledge, namely being able to obtain the entire landscape of the dynamics in the parameter space that comes from analytical solutions, while still having to sacrifice the real nature of the system in order to translate it in the form of differential equations. Immunology gives one further reason to prefer cellular automata, that is the overwhelming number of populations involved and their continuously changing numbers. The immune system has a continuously evolving population of cells and although some of the cell populations are nonspecific and thus are always existent, the specificity of the lymphocytes constitutes a definite challenge for the differential equation modeller. For the model to include idiotypes, it would need a differential equation for every idiotype of every lymphocytes with corresponding parameters such as affinity with the antigen. Further, every idiotype would probably also need a different variable to differentiate between the dynamics of the memory cell, the effector cell and, in the case of a B cell, the antibody population for that idiotype. The idiotype populations are also bound to change drastically with the constant extinction of "bad" idiotypes and the creation of new idiotypes from random generation in the bone marrow as well as through errors occurring during triggered cell division of favoured idiotypes. Of course, if one is to model the immune system using differential equations, common sense dictates that one has to first simplify the dynamics to no more than a few equations. This is not always a bad approach as it will, for a valid model, give an averaged estimate of the dynamics. For addressing very specific questions, it is possible to make valid simplifications in the realm of the question addressed and for such cases a differential equation model is likely the best way to come about a better understanding of the dynamics at play.

But when one's ultimate goal is to understand the full behaviour of the system, in particular the characteristics of the local rules that give rise to the formidable array of complex behaviours observed in the immune system, it is necessary to resort to a more appropriate tool, namely cellular automata. CA were originally introduced by John von Neumann and

Stanislas Ulam under the name of "cellular spaces" as possible idealization of biological systems. They sought to show that biological processes such as the reproduction and evolution of organized forms could be modelled by simple cells following local rules for changing a cell parameter with time [11]. Traditional CA usually consist of a regular uniform N-dimensional grid that can either be finite or infinite (periodic boundaries) in extent. The grid contains a discrete variable (or "cell") at each site that can assume m possible discrete values. The state of a CA is completely specified by the values of all variables at each site. The CA evolves in discrete space with discrete time steps with the value of a variable at a given site being affected by the values of variables at sites in its "neighbourhood" at the previous time step. The "neighbourhood" of a site can be defined in numerous ways and can be as extended as one wishes. At a new time step, the variables at all sites are updated based on their own value and that of their defined "neighbourhood" at the preceding time step according to a definite set of local rules. Lattice gas simulations are an extension of CA and are so named because of their original use in modelling fluids as particles which are restricted to moving between discrete positions in the CA lattice. In a lattice gas simulation, the sites are only the discretized space in which the variables are allowed to move and evolve. Cellular automata and in particular lattice gas models are good candidates for studying SOSs as they themselves are SOSs and hence constitute the natural way to represent such systems. A well known feature of CA computations is their ability to produce surprisingly complex behaviour from very simple rules. A CA can produce a range of periodic, chaotic and generally very complex behaviours with intricate spatial and temporal patterns [23]. Nonlinearities or time-delays are not intrinsically difficult to treat since they are either a consequence of the dynamics rather than a cause or are otherwise trivial to include in the context of the CA local rules. Another great advantage of CA is that its computational structure is inherently parallel and it can therefore be run on a parallel computer very efficiently.

The earliest CA of the immune system neglected the microscopic details of the immune system's behaviour. Typically, a set of characteristic states like "in rest" or "infected" would be represented by a 1-D CA and the evolution of this system would be based on simplistic rules that defined how the CA switches from one of those states to another (see [18], Section V for a brief overview of the most prevalent models of this type). The first models to really employ CA, more precisely lattice gas simulations, to model the various components of the immune system were born in the early 90's, the most outstanding of which are the bit-string models and they are the subject of the next section.

4 Cellular Automaton Bit-string Models

As discussed above, in a real immune system, binding between for example a B-cell receptor and an antigen's B-cell epitope will takes place through the matching of the amino acids string composing both the receptor and the epitope, and binding will occur only if a sufficient match between both amino acid strings is met. It is possible to represent the amino acid string using an alphabet of m characters where each character corresponds to a given amino acid. Then, any of a number of string-matching algorithms can be used to determine the

degree of complementarity. For a sting of length N, the repertoire size is given by m^N such that the greater m is, the larger the repertoire size will be. This clever idea for representing a receptor's or epitope's idiotype was first introduced by Farmer et al. [7] as a way to perform clever calculations for determining molecular complementarity and predict the optimal size of an epitope — see [18] for a short overview of this calculation. But the greatest impact of the string representation introduced by Farmer was in lattice gas simulations of the immune system. The typical lattice gas simulations chose to build the epitope and receptor strings using an alphabet containing only two characters, namely 0 and 1, and those models were consequently referred to as "bit-string" models. Since the receptors and epitopes of a bit-string model are represented by a string of bits, i.e. a string of 0s and 1s, those models have a repertoire size of 2^N, where N is the length of the bit string.

The idea of using bit-strings to model the receptors and epitopes was first applied to cellular automaton modelling by Seiden and Celada [5, 19]. They called their model `IMMSIM`. More accurately, `IMMSIM23` is the original version of `IMMSIM`. It was written in the `APL2` language by Philip E. Seiden and requires the `IBM APL2` runtime environment. In its original version, `IMMSIM` contains B-cells, T-cells, and antigen presenting cells (APC) as well as antigen (Ag) and antibody (Ab) molecules and is concerned with the humoral response only [19]. The later version of the model, `IMMSIM3` developed by Seiden in the `APL2` language as well, was made to include Th and Tc cells as well as epithelial cells in order to add the cellular response to complement the already existing humoral response [3, 4]. Although the general idea for the code was good, the fact that the code was written in the `APL2` language was beginning to be a real nuisance. As Bernaschi and Castiglione point out in [1],

> "The original implementation of the model made use of `APL2` that is an interpreted language with no "explicit" dynamic memory allocation capability. This choice along with the intrinsic complexity of the model prevented the authors from running any but "relatively" small-scale simulations."

To remedy this situation, Bernaschi and Castiglione came up with a parallel coded version of the latest `IMMSIM` entirely written in the `C` language that they called `ParImm` [1] and later `CIMMSIM`. Finally, a tutorial `C++` version of IMMSIM was developed by Steven Kleinstein based on the C version developed by Filippo Castiglione [13].

In bit-string models, much like in the real immune system, the immune components are characterized by their receptors. Here, B cells are composed of a B-cell receptor and a MHC class II molecule. T cells only have a T-cell receptor and antigen presenting cells (APCs) only have a MHC class II molecule. Antibody molecules are typically made up of a single B-cell receptor. Finally, antigen molecules are made of a number of B-cell and T-cell epitopes, although in most application of bit-string models so far, the antigen has been limited to one B-cell epitope and one T-cell epitope. Figure 2 presents the schematics of those cells and molecules. In a real immune system, the B-cell and T-cell receptors are physiologically different, but because all receptors are implemented the same way in bit-string models, the distinction between various types of receptors is made by restricting what can bind with what through simple binding rules. The various binding rules have been designed as follows. An antigen's B-cell epitope can be bound by an antibody's B-cell receptor depending on

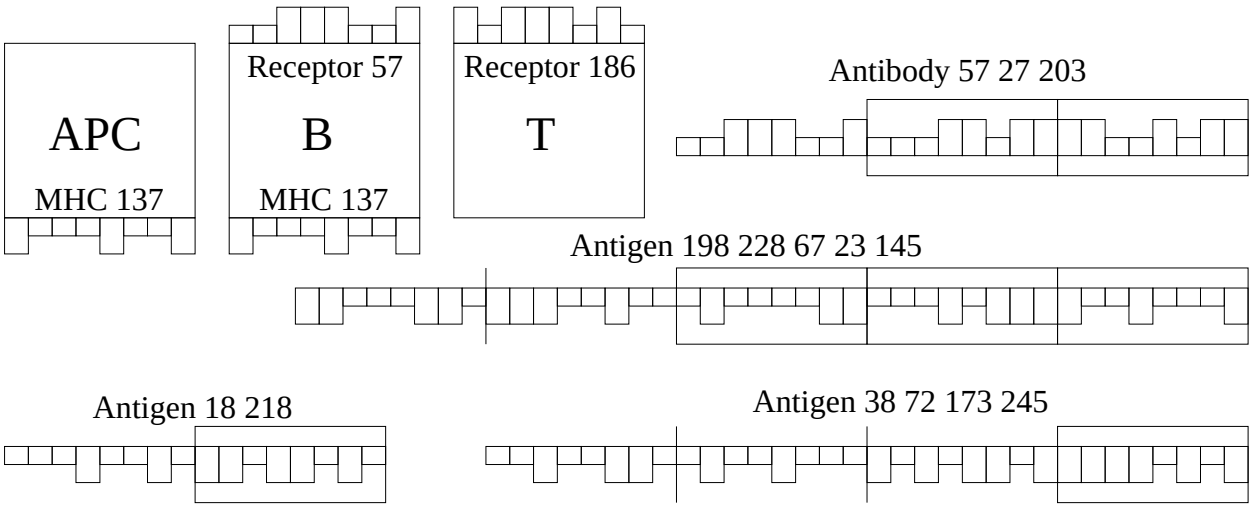

Figure 2: The schematic of the B-cell, T-cell and APC as well as examples of that of the antibody and antigen molecules. On antibody, the B-cell epitope is shown exposed while T-cell self and non-self epitopes are boxed. On antigen, B-cell epitopes are shown exposed while T-cell epitopes are boxed.

the probability of interaction in which case an antibody-antigen complex forms and it is removed from the simulation. The antigen's B-cell epitope can also be bound by a B-cell receptor, in which case the antigen's T-cell epitope is then broken in two and whichever half matches the right half of the B-cell MHC II molecule best will bind it and the other half of the antigen's T-cell epitope will be presented to encountered T-cells along with the left half of the B-cell's MHC II molecule. Figure 3 illustrates this binding procedure. APCs will bind any antigen with a fixed probability (e.g. 0.002 in [3]) and will present the antigen's T-cell epitope on their MHC II molecule to encountered T-cells in the same fashion B cells do. Finally, when a T-cell is presented with an antigen's T-cell epitope-MHC II complex attached to a B-cell or an APC, depending on the probability of interaction, the T-cell will bind with the presenting APC or B-cell and will be activated and if bound with a B-cell, the latter will also be activated. Activation results in differentiation at a constant rate over a fixed number of time steps. In the real immune system, an individual will have less that 10 different clones of MHC class II molecules. In the Seiden-Celada model, the simulations usually consider only 1 or 2 [3].

In most bit-string models, notably in the Celada-Seiden model [19], which is the model we are interested in, the function used to define a "match" between a pair of bits is simply an exclusive OR function (`XOR`), where 0 matches 1 and 1 matches 0. The `XOR` function is presented in Figure 4. From the number of bit mismatches for a given pair of strings, the probability of interaction is typically calculated as follows, although it does change from models to models depending on the question addressed and the author's preferences. Interaction probability is characterized by 3 parameters, namely the minimum match, the

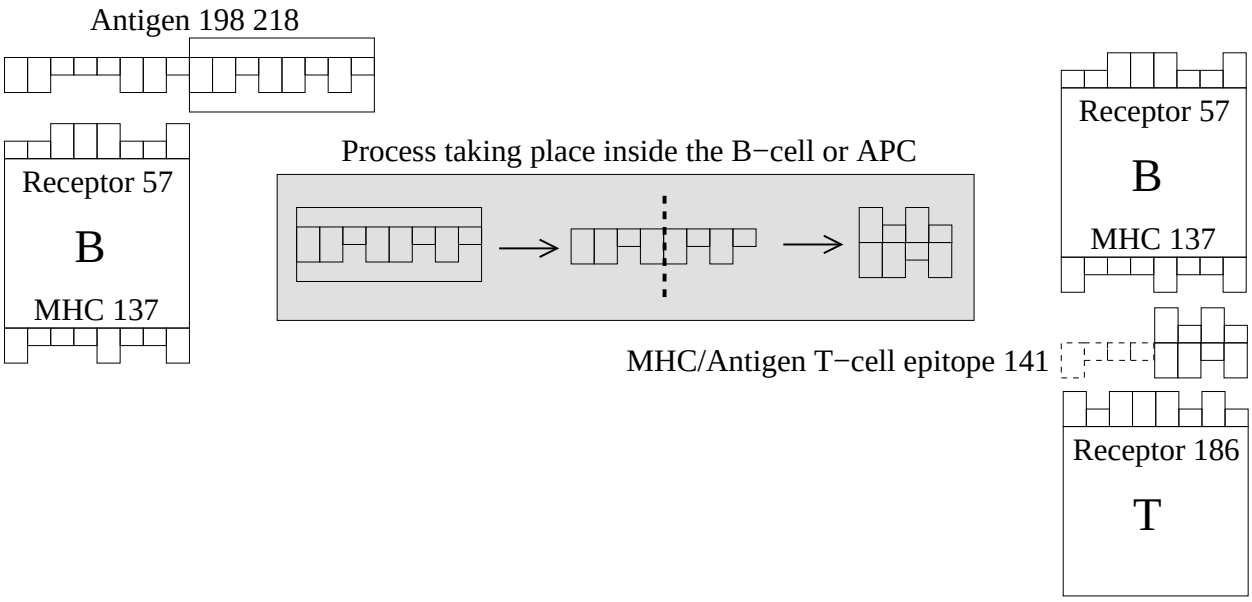

Figure 3: Schematic of the algorithm for the binding of an antigen with a B-cell and the subsequent presentation of the MHC/Antigen T-cell epitope to a T-cell.

0	XOR	0	=	0
0	XOR	1	=	1
1	XOR	0	=	1
1	XOR	1	=	0

	0	0	1	0	1	0	0	0	1	0	0	1	0	1	0
XOR	1	0	0	1	0	1	0	1	0	0	1	0	1	0	1
	1	0	1	1	1	1	0	1	1	0	1	1	1	1	1

Figure 4: On the left, the definition of the XOR rule. On the right, the XOR rule applied to a pair of 10-bit strings resulting in a 3-bit mismatch (3 null characters) between the two stings.

minimum affinity and the affinity increase. The minimum match is the threshold of matching bits below which the interaction will not take place. The minimum affinity is the probability of interaction when the number of matching bits is exactly the minimum match. Finally, the affinity increase is the increase in affinity between an n-bit match and an $n+1$-bit match and can be defined in a number of ways. In [13], the affinity increase is defined as

$$\frac{\text{Prob. for } n+1 \text{ mismatch}}{\text{Prob. for } n \text{ mismatch}} = \text{affinity increase} \times \frac{\binom{N}{n}}{\binom{N}{n+1}}, \tag{1}$$

where

$$\binom{N}{n} = \frac{N!}{n!(N-n)!}, \tag{2}$$

and N is the chosen length of the bit-strings. To give an idea of the computational scale of the simulation, it is interesting to mention that most simulations were run on relatively small grids, typically a 15x15 hexagonal grid[3] and that the size of the bit-string, N, is typically chosen to be 8 yielding a clonal diversity of $2^8 = 256$. The time scale of a simulation is not always specified, but in [13](p.74) it is suggested that the relationship between a time step of the model and real time is such that 1 time step in the model corresponds to 1 B-cell division. It is however impossible to tell whether this is the standard way to define the time scale or if other authors have chosen different conventions.

At this point, it is useful to describe what a time step corresponds to. This is slightly difficult because this may also vary between articles and it is not always specified. Instead, we will give an example by describing a time step as defined in [3]. Interactions only occur between entities that are allowed to interact and are located in the same cell. Since at any given step, there will likely be more than one possible interaction for a given entity, all possible interactions are considered and their interaction probability is calculated and compared with a random number to determine whether the interaction is successful. An entity can have at most one successful interaction on any one time step such that only one of the successful interaction can take place. One can chose the interaction that had the highest probability or chose one of the successful interactions randomly. Once all interactions have been determined, they are allowed to take place. Then, entities are allowed to die, stimulated cells divide, new cells are born, and antibodies are generated. Finally, the entities diffuse to a randomly chosen neighbouring cell and that concludes one time step. From what the articles seem to say, cells do not die after a given number of time step, but rather die with a probability defined by their half life. Cell division will occur in B cells and T cells that have been activated. B cells can divide into memory cells with a longer half life and plasma cells that have a shorter half life but will generate great amounts of antibodies of the same idiotype as the B cell they originated from. When new B and T cells are "born" they are

[3]Note that in most articles, the authors mistakingly claim they were using a triangular grid, namely in [6, 13, 19]

added at random location on the grid and newly born T cells will have to undergo thymus selection. Thymus selection in the simulations is done through the positive selection of T cells who's receptor's left half recognizes MHC II molecules' left half with at most 1 mismatch and right half recognizes antigen peptides and through negative selection of T cells who's receptor's left half does not recognize the left half of the bare MHC II molecules. This results in the B cell repertoire being complete, i.e. 2^N, while the T cell repertoire starts off complete but is then tailored through the thymus selection process. The removal of antigen can be the result of various scenarios. When an antigen-antibody complex forms, both entities are removed from the simulation. Also, when an antigen has been recognized by a B cell and successfully presented to a T cell, resulting in the B and T cells being activated, the antigen is removed and the B and T cells are made to divide. The ratio of plasma B-cells to memory B-cells resulting from a division is a parameter of the simulation, typically set to 0.5. B cells require the help of Th cells to activate. Finally, when an APC presenting an antigen peptide on its MHC II molecule is bound by a T cell, only the T cell divides and the antigen is removed.

The immune system simulations described above consider only the humoral branch of the immune system (T cell is to be understood as T helper cell) and the antigen is not considered infective. Infection had to be included in order to represent the cell-mediated branch since this branch's major contribution is through cytotoxic T cells (CTL) recognizing infected cells. Using the description of the early implementation of a model including the cell-mediated branch described in [3], it appears the antigen can infect B cells and APCs (such that no target cells had to be added) and infection is described by three parameters: the probability of infection when it encounters a B cell or APC, the rate of replication of the antigen inside the infected a cell and the maximum unit of antigen an infected cell can contain before exploding and releasing all the antigens it contains into the simulation. At this early stage, CTLs are non-specific cells, i.e. they have no receptor, and they interact with any infected cell with a fixed probability and destroy the cell. In a later paper [4], Celada and Seiden describe a more complex implementation of the cellular and humoral branch which adds to the simulations epithelial target cells, danger signals and cytokines, Th_0 cells that evolve into Th_1 and Th_2 cells, MHC class I molecules along with specific CTLs. However, their description of the structures and functions of the new entities is extremely brief at best and it is impossible to describe their implementation without having to read through their original coding which time does not permit at this point.

5 Other Models

The Celada-Seiden model is a very interesting model as it is attempting to create a general immune system simulator rather than only a simplified toy model for answering a particular problem. The various authors who have used the Celada-Seiden model claim that the results obtained compared well with experimental data. The model has already been used to investigate various immunological questions among which the process leading to affinity maturation and hypermutation in the humoral branch of the specific immune system [6], the

rheumatoid factor paradox [21], the transition process between immune and disease states and the contribution from the humoral only, cellular only, and humoral and cellular branches of the immune system [3, 4], the impact of vaccine efficiency on the humoral branch, the cellular branch and both branches [14], and the dynamics of HIV infection with respect to the selection of escape mutants from immune recognition [2]. Those results were obtained by the `IMMSIM` group based at Princeton University and their collaborators (Bernaschi, Castiglione, etc.).

Another very interesting group is coming up with interesting results. The group seems to be based at the University of New Mexico and is mainly composed of Stephanie Forrest, Ron Hightower, Steven A. Hofmeyr, Derek J. Smith and co-author a lot of their publications with Alan S. Perelson. This group also uses `IMMSIM` to simulate the immune system but it is not their only focus as they also perform a lot of calculations about immune diversity. They have published very interesting papers using the bit-string representation where they explore evolution and somatic learning in V-region genes [17], and the use of genetic algorithms to explore pattern recognition in the immune system [9]. In [17], they find that their 512-bits genome can encode information about 32 64-bits antigens, i.e. about 2048 bits. This organization of the resources by the immune system suggests that the system has the ability to recognize pattern in the antigens and exploits this to maximize its recognition while minimizing its resources. In [9], they explore this further using a genetic algorithm and find that the immune system is impressively good at recognizing patterns. Challenging the pattern recognition capabilities of the immune system, they found that the genetic algorithm together with clonal selection could evolve an antibody type that matched multiple antigens through the identification of common schemes among antigens. They also found that when the genetic algorithm was faced with 2 or 3 different antigens, i.e. multiple peaks in the antigen landscape, the genetic algorithm can discover all peaks and similarity among antigens does not prevent distinction by antibodies, no matter how similar the antigens are. Finally, they found that the number of antibodies exposed to any antigen will greatly affect the capacity of the immune system to recognize patterns. But perhaps even more interesting is the group's attempt at constructing a computer immune system to improve computer security. By trying to engineer a protective system that operates successfully in an environment resembling, in constraints and challenges, the real immune system they address the questions of what is really needed for efficient protection of the system and what it is needed for [8]. As the forces of natural selection don't guarantee perfect or minimal solutions but rather solutions that work well enough to increase survivability, it is interesting to ask whether the human immune system is optimal. Their attempt at building a computer immune system is the first step towards answering this question. In another work [20], using comparison with computer memory, they established that immunological memory is in the same class of associative memories as Kanerva's Sparse Distributed Memory. This class of memories is known to be robust and is characterized by sparse sampling of a huge input space by recognition units (B and T cells in the immune system) and distribution of the memory among many independent agents (B and T memory cell population in the immune system), traits characteristic of SOSs. Overall, the work performed by this group is much

needed and their original approach will likely lead to groundbreaking results.

In [15], Meier-Schellersheim and Mack present their own model, `SIMMUNE`. The model is run on a 3-D grid and can be made to comprise different compartments (e.g. lymph nodes, thymus, etc.). It lets the user define the properties of the compartments within the simulation such as the dimensions of the compartment, the diffusion rates of molecules and cells within the compartments and the initial concentrations of the different types of agents can be given. Further, the exchange of agents between the different compartments can be regulated, e.g. which kind of agents are allowed to pass from one compartment to another and at which rate. Interestingly, this model incorporates the ability of B-cells to respond to certain kind of antigens (e.g. polysaccharides as they appear of the membranes of bacteria) directly, i.e. without the assistance of T helper cells. Although this could be added to `IMMSIM` without much difficulty, it has not yet been done. One major drawback of the `SIMMUNE` model is that it is impossible to find any kind of precise description of its implementation. In [15], the model is defined in very general terms which makes criticizing its functionality very difficult. What makes this project worth mentioning is that the authors suggest a method to investigate how "context adaptative behavior" of the immune system might emerge from local cell-cell and cell-molecule interactions. It suggests starting from single-cell behaviour and entering the scale of cell-cell cooperations by looking for correlations between cell actions, which takes one to the next scale. The authors suggest that repeated application of this process might lead to a scale that directly describes the macroscopical behaviour of the immune system. This approach should be explored as a possible way to systematically determine the global dynamics of a system from its local rules.

Finally, it may be worth mentioning the work of Rita Maria Zorzenon dos Santos et al. who have developed simplistic CA toy models of the immune system to answer specific questions. In [24], Zorzenon dos Santos and Bernardes use a very simple cellular automaton of immunization and aging in the context of immune network. The cellular automaton model is such that rather than having different types of cells moving around on the grid (lattice gas model), every site of the CA grid in a d-dimensional space corresponds to the shape of a molecule (associated with a B cell or antibody) and can be in any of 3 states: $B_i = 0$ for low (or nonactivated), $B_i = 1$ for intermediate, or $B_i = 2$ for high (activated). Site i interacts with $2d+1$ sites: i.e. its mirror image, $-i$ and the nearest neighbours of the mirror image (representing imperfect but high affinity matches). The authors then define a field, h_i that represents the concentration of complementary receptors as $h_i(t) = \sum_{2d+1} B_i(t)$, which is the sum of the state of the site itself and its $2d$ neighbours. The rule for the evolution of the system is then defined as

$$B_i(t+1) = \begin{cases} B_i(t) + 1 & \text{if } \theta_1 \leq h_i(t) \leq \theta_2, \\ B_i(t) - 1 & \text{otherwise,} \end{cases} \tag{3}$$

but no change is made if it would lead to $B_i = -1$ or 3. This simplistic model has a transition region, between a stable and a chaotic region, where the authors have observed emergent complex behaviour which they claim is appropriate to describe a self-regulated multiconnected network and possibly provides the much needed connections between the

available experimental results and the Jerne theory [12]. Continuing with the model, they move on to simulating antigen presentation by setting a randomly chosen region of the receptor population in the $B = 0$ state to the $B = 2$ state and find that the "older" the system is (i.e. the more time steps elapsed before the first antigen presentation), the more rapidly it saturates and hence the less intense is its response. By taking a closer look at what they are observing, it appears the results obtained are really an artifact of the rules defined rather than a real phenomenon. For instance, in the context of an aging system, after more time has elapsed, the system will be "noisier" and is therefore unlikely to notice and hence react to any new perturbation like the introduction of a new antigen, a phenomenon the authors confused with the loss of plasticity as a system ages. In [25], Zorzenon dos Santos and Coutinho suggest an extremely simple CA to model the dynamics of HIV. In this model, every site in the 2-D CA grid represents a target cell for the HIV. Each cell can be in any of four states, namely healthy, infected-A1, infected-A2 or dead. Note that the absence of cell is not a possible state. The states infected-A1 and infected-A2 correspond respectively to an infected cell that is free to spread the infection and an infected cell in its final stage before apoptosis such that it can only infect a healthy cell if it is present in sufficient concentration. The rules for the evolution of the CA are the following. A healthy cell becomes infected-A1 in the next time step if any of its 8 neighbours are infected-A1 or if at least $2 < R < 8$ of its neighbours are infected-A2. An infected-A1 cell becomes infected-A2 after τ time steps. An infected-A2 cell becomes a dead cell in the next time step. A dead cell is replaced by a healthy cell with probability p_{repl} or otherwise remains dead in the next time step. Finally, any new healthy cell can instead be created as an infected-A1 cell with probability p_{infec}, such that the rate at which dead cells are replaced by infected-A1 cells is $p_{\text{newinfec}} = p_{\text{repl}} \times p_{\text{infec}}$. This last rule is suppose to simulate the introduction of new infected cells in the system, either coming from other compartments of the immune system or resulting from the activation of the latent infected cells. Even more so than in [24], the model here is much too simple to truly represent the dynamics of HIV and yields extremely misleading results. In Figure 5, the averaged number of healthy, infected and dead cells over the course of a simulation are presented. We can clearly see that the model seems to have three different phases that one can be tempted to associate with the three-phase dynamics of HIV, namely the primary infection, the asymptomatic latency period followed by the onset AIDS. In Figure 6, screenshots of the model during the various phases of the infection are presented. One can see that the model has a natural tendency to form pattern as a direct consequence of the chosen rules and the reintroduction of infected cells. Those patterns would not form if the cells were allowed to move from site to site like it is the case in vivo rather than have every site representing a cell. It is likely that the typical wait time before the patterns seen in Figure 6(d) form, although this is not said explicitly in the article, has been used by the authors to "calibrate" their model to real time. About those patterns, the authors say

> "These growing structures of infected cells may be associated with syncytia, an aggregation of infected cells observed experimentally, and according to our results they would be responsible for the depletion of T cells leading to AIDS. These

result actually corroborate some previous suggestions that syncytia could be responsible for the permanence of the virus in the system, based on the analysis of the similarities between HIV infection and other diseases."

It appears that even the authors were mislead by the results they obtained. The permanence of the infection is due to the fact that infected cells are continuously being added at a rate p_{newinfec} per dead cell per time step. As for the patterns, they form naturally as an artifact of the rules and the density of cells and can be mistaken for the onset of AIDS as the patterns are such that the healthy cells are few while infected cells are numerous (see Figure 6(d)) which is consistent with AIDS. In fact, their three-phase dynamics correspond to the following sequence of events. Initially, the grid of healthy cells is implanted with 5% of infected-A1 cells (Figure 6(a)). As the "squares" of infection grow, they merge and wipe out the grid leaving mainly healthy cells with the occasional infected cells triggering "squares" of infection (Figure 6(b) and (c)). Finally, the randomly introduced infected cells and their resulting "squares" of infection invariably lead to the formation of the more complex square patterns. This hypothesis is further strenghtened by the results presented in Figure 5 or more precisely by the error bars on those results. The first two phases have very little error bars while the third phase has large error bars that slowly attenuate. This is a direct consequence of the fact that the patterns corresponding to the third phase are triggered by very particular configurations of the cell populations that are randomly generated when new infected cells are added. This results in the third phase occurring invariably but at very variable times explaining the large error bars associated with the initiation of the patterns and a later attenuation of the error bars associated with the ever decreasing probability that the patterns would not yet have formed. Thus, one can expect that the magnitude of the error bars probably follows a Poisson distribution. The work of Zorzenon dos Santos et al. has been mentioned here as a reminder that one has to always be very cautious about the way one implements a model and has to continuously make sure that the results obtained are not an artifact of the way the system has been implemented. In this case, the fact that the authors chose to have each site represent a cell and have no cell diffusion led to patters that are not to be expected in the real system.

6 Pushing the Envelope with Bit-string Models

A lot has already been done using the Celada-Seiden (CS) model and results obtained so far with various versions of the model confirm that CS model is a good approach to simulating the immune system. It is also clear that the model offers endless possibilities of application and a lot of projects can be derived from this. In this section, an overview of possible projects involving CS types of models is presented. The goal is to draw tentative plans for the type of projects that can be carried out to fulfil the requirements of a PhD degree in Physics.

The first step will be to implement a bit-string model based on the latest version of the CS model. I have already implemented a visual interface program along with a message passing library to simulate multi-agent systems. I called it `MASyV` for **M**ulti-**A**gent **Sy**stem **V**isualization. I have successfully written several client simulations using `MASyV` among which

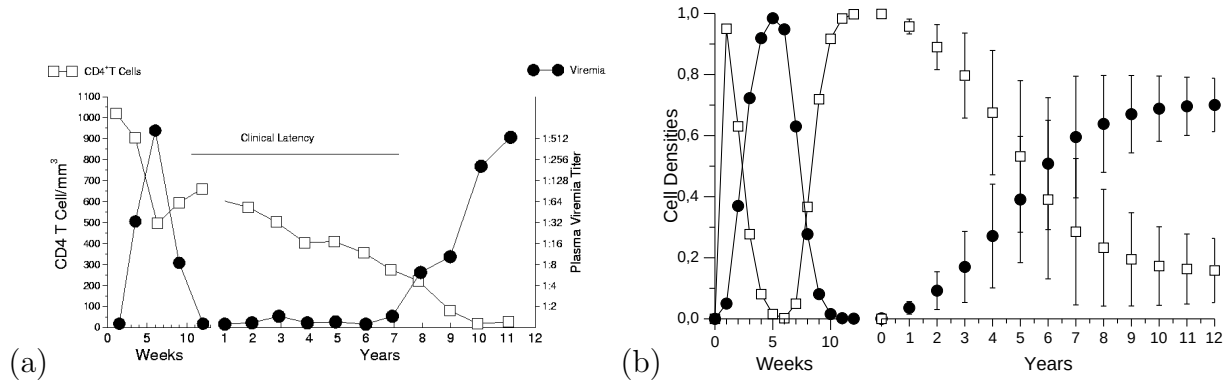

Figure 5: (a) Real experimental data. (b) The simulation's result of the evolution of the density of healthy (open squares), infected (full circles) and dead (open triangles) cells. From [25].

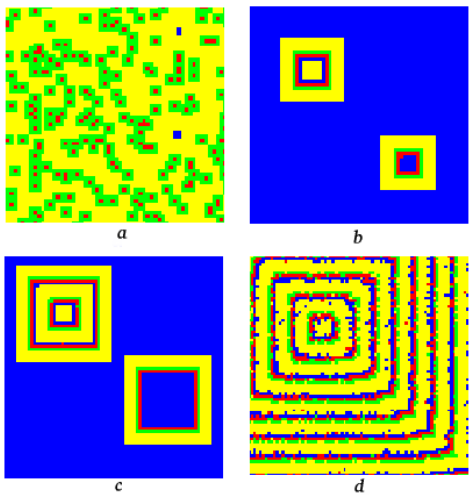

Figure 6: Screenshots from the CA simulation of HIV infection at time step 5 (a), 18 (b), 25 (c) and 200 (d) weeks where each time step correspond to 1 week. From [25].

are simulations of the early formation of trail networks by ants, the Ising spin model and the organization of electrons withing the double-well potential of a grid of tubulin dimers (a microtubule). I plan to implement a `C` version of the CS model as a client for `MASyV`. The model would not originally be parallelized but development would rather be focused on implementing all the behaviours that are currently part of the latest CS model. A second step would be made to parallelized the code to allow for larger scale simulations.

A lot of research teams are currently working at developing and improving the CS model, groups namely headed by Celada at Princeton and Bernaschi in Italy. As a single researcher, it would be unrealistic to try to compete with those big research teams. The strategy would then be to undertake projects that are unlikely to be attempted by them but that are still of interest. Most people currently working on `IMMSIM` are immunologists or computer scientists and the main focus of all research groups so far has been on applications of the model to current immunological questions or on the improvement of the model as immunological knowledge is improved by new discoveries. However, there is now a great need for investigation into possible scaling laws in order for this model to be properly compared to the real immune system such that it can address valid questions in a quantitative manner. As we have mentioned previously, the immune system is restrained in size and of course the size of an organism determines the maximum number of its lymphocytes. Mice have of the order of 10^8 lymphocytes while humans have of the order of 10^{12} lymphocytes. With the computer capacities of today, this number is still quite big to tackle. For this reason, one has to pay constant attention to whether the results seen are artifacts due to the restricted size of the simulation. Thus, once a working model has been achieved, it will be important to develop scaling laws, i.e. quantitative parallels between the real immune system and the simulations. For instance, the CS model uses bits to represent amino acids and the repertoires considered are complete, i.e. the number of possible bit-strings is 2^N where N is the length of a bit-string. The real immune system typically, although not always, uses amino acids rather than bits as building blocks for the epitope and receptor strings. A careful study of the amino acid strings' assembly process in the bone marrow will allow one to draw quantitative parallels between the bit-strings and the amino acid strings. One may find that bits are not sufficient and an alphabet of at least 5 characters is required to give sufficiently close behaviour. There are also other simplifications made in implementing the CS model that should be reconsidered and analyzed for possible impact. For instance, the model assumes a threshold for the number of bit match between two bit-string below which no interaction can occur. Further, matching between a pair of bits is all or none, i.e. there is no partial match. It will be interesting to measure the consequences of those choices. Also, it would be interesting to look closer at the ratio of plasma to memory B cells produced when a B cell gets activated as it appears the model implements the ratio as a fixed constant (see [13] p.74). Whether using a binomial distribution to calculate the probability of interaction as a function of the number of bits matching constitutes a good approximation should also be investigated. There are also reasons to suspect (see [13] p.75) that cell death occurs with a probability defined by the inverse of the half life rather than as a function of the "age" of the cell, a factor that could have an impact on the global dynamics. Another important

choice in the model is the fact that any given entity is allowed to have at most one interaction per time step. To do this, the model will usually select randomly which interaction takes place among all the interactions that have been found to be successful. This choice is questionable and it would be interesting to look into the impact of it on the dynamics by considering different scenarios and comparing with experimental data when possible. Those little projects each constitute reasonable amounts of work, address important problems and will all yield publishable results.

There are also a few aspects of the immune system that could be investigated under a new perspective. In many aspects of the immune system, there is competition between modes of functioning or between entities. A good example is the constant competition between the humoral and cellular branch of the specific immune system. In effect, there are two types of T helper cells, namely Th_1 and Th_2 which respectively contribute to the cellular and humoral branches of the specific system through emission of different patterns of cytokines [3]. The cytokines produced by B-Th_1 binding down-regulate CTL division while those produced by CTL activation through the intervention of Th_2 down-regulate B-cell division; this phenomenon is known as cross-regulation. A lot of work has been done on predator-prey models and competition models where two species compete for the same resource. One cannot help noticing the striking parallel between such competition over resources and the constant competition between the cellular and humoral branches for the antigen. There would be a lot of potential in studying this delicate balance in the light of what is known about competition strategies. This would be another interesting and reasonable project that can be undertaken in a reasonable amount of time and guarantees results.

Finally, at this point, I will have learned a lot about the immune system and, if time permits, I will be ready to start using the model to answer specific problems. Through close collaboration with experimentalists, brand new hypothesis could be tested with the model that could then be confirmed in the laboratories through experiments that will have been shaped to test specific key results provided by the model. This initiative would also present very little risk of replication of work as the tested hypothesis would come from very particular research undertaken by our collaborators at the department of Pharmacy.

7 Conclusion

What has been demonstrated here is the ingenuity and applicability of the bit-string model, and in particular the Celada-Seiden general immune system simulator. While Physicists will probably not contribute as much to drug design and new discoveries which is something immunologists do very efficiently through experimentation, their major contribution will probably come as a better understanding of the general behaviour of the immune system with the ultimate goal to apply the immune system's strategies to solving problems such as computer security. By looking at the big picture rather than more specific problems, one can start to use the mechanisms of the system to one's advantage. For instance, how to drive the system to a more humoral than cell-mediated response and how to design drugs to do it. In a time of collaboration between various field of research Immunology perhaps more than

any science will benefit largely from this new initiative and other fields such as Computer Science and the study of self-organizing behaviour will also gain from this experience. It has been shown that cellular automaton modelling is a powerful tool, but just like any other tool, one must be cautious when using it and when interpreting the results as a faulty model will lead to misconceptions about the nature of the dynamics. But overall, modelling offers a cheap and simple way to test hypothesis that would perhaps not be given a chance because of the costs usually involved in carrying out experiments. It also allows to test aspects that could not necessarily be tested in vivo, e.g. looking at the effect of only the cell-mediated or only the humoral branch by simply turning off one branch or the other would be hard and costly if at all possible in vivo.

Furthermore, knowledge gained about the strategies used by the immune system will benefit much more than just Immunology. The network of entities that constitute the immune system is perhaps as intricate as the neural networks that make up our nervous system. However, investigating the immune system is more straightforward as we know what its ultimate role is, i.e. to regulate and protect self against nonself. Once the role of the system is known, it is easier to interpret results and hypothesize strategies based on the need to always better perform its role. The role of the nervous system is not so simple as cognition is poorly understood and intelligence is still awaiting a proper definition. Understanding the immune system's learning and memory capabilities may well be the first step towards a better definition of cognition and intelligence and will likely lead to a better understanding of distributed intelligence systems such as our nervous system. Even though Immunology has developed itself faster than most sciences, mainly due to the fact that it is a recent science and as such has benefited from more modern techniques since its foundation, it is still in its infancy. A lot is known about the immune system, but even more remains to be understood. Given the nature of the immune system, it is certain that modelling efforts will be of great help in shaping the face of modern Immunology.

References

[1] M. Bernaschi and F. Castiglione. Design and implementation of an immune system simulator. *Comput. Biol. Med.*, 31(5):303–331, 2001.

[2] M. Bernaschi and F. Castiglione. Selection of escape mutants from immune recognition during hiv. *Immunol. Cell Biol.*, 80:307–313, 2002. Special issue devoted to Immunoinformatics.

[3] M. Bezzi, F. Celada, S. Ruffo, and P. E. Seiden. The transition between immune and disease states in a cellular automaton model of clonal immune response. *Physica A*, 245(1–2):145–163, 15 October 1997.

[4] F. Celada and P. E. Seiden. Modeling immune cognition. *Proceedings of the 1998 IEEE International Conference on Systems, Man and Cybernetics, San Diego, CA*, pages 3787–3792, 11–14 October 1998. IEEE Catalog Number 98CH36218.

[5] F. Celada and P. E. Seiden. A computer model of cellular interactions in the immune system. *Immunol. Today*, 13(2):56–62, February 1992.

[6] F. Celada and P. E. Seiden. Affinity maturation and hypermutation in a simulation of the humoral response. *Eur. J. Immunol.*, 26(6):1350–1358, June 1996.

[7] J. D. Farmer, N. H. Packard, and A. S. Perelson. The immune system, adaptation, and machine learning. *Physica D*, 22:187–204, 1986.

[8] S. Forrest and S. Hofmeyr. Engineering an immune system. *Graft*, 4(5):5–9, 2001.

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

[10] R. A. Goldsby, T. J. Kindt, and B. A. Osborne. *Kuby Immunology* 4^{th} *Edition*. W. H. Freeman and Company, New York, 2000.

[11] C. E. Hecht. *Statistical Thermodynamics and Kinetic Theory*. W.H. Freeman and Company, New York, 1990.

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

[13] S. H. Kleinstein and P. E. Seiden. Simulating the immune system. *Comput. Sci. Eng.*, pages 69–77, July/August 2000.

[14] B. Kohler, R. Puzone, P. E. Seiden, and F. Celada. A systematic approach to vaccine complexity using an automaton model of the cellular and humoral immune system. i. viral characteristics and polarized responses. *Vaccine*, 19:862–876, 2000.

[15] M. Meier-Schellersheim and G. Mack. SIMMUNE, a tool for simulating and analyzing immune system behavior. arXiv:cs.MA/9903017, 28 March 1999.

[16] A. S. Perelson. Modelling viral and immune system dynamics. *Nature Rev. Immunol.*, 2(1):28–36, 2002.

[17] A. S. Perelson, R. Hightower, and S. Forrest. Evolution and somatic learning in V-region genes. *Res. Immunol.*, 147:202–208, 1996.

[18] A. S. Perelson and G. Weisbuch. Immunology for physicists. *Rev. Mod. Phys.*, 69(4):1219–1267, Oct. 1997.

[19] P. E. Seiden and F. Celada. A model for simulating cognate recognition and response in the immune system. *J. Theor. Biol.*, 158(3):329–357, October 1992.

[20] D. J. Smith, S. Forrest, and A. S. Perelson. Immunological memory is associative. In *Workshop Notes, Workshop 4: Immunity Based Systems*, pages 62–70, Kyoto, Japan, 1996. Intnl. Conf. on Multiagent Systems.

[21] J. J. Stewart, H. Agosto, S. Litwin, J. D. Welsh, M. Shlomchik, M. Weigert, and P. E. Seiden. A solution to the rheumatoid factor paradox: Pathologic rheumatoid factors can be tolerized by competition with natural rheumatoid factors. *J. Immunol.*, 159(4):1728–1738, 1997.

[22] S. H. Strogatz. *Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry, and Engineering*. Perseus Books, Cambridge, MA, 1994.

[23] S. Wolfram. Cellular automata as models of complexity. *Nature*, 311, 1984. available at: `http://www.stephenwolfram.com/publications/articles/ca/84-cellular/`.

[24] R. M. Zorzenon dos Santos and A. T. Bernardes. Immunization and aging: A learning process in the immune network. *Phys. Rev. Lett.*, 81(14):3034–3037, 5 October 1998.

[25] R. M. Zorzenon dos Santos and S. Coutinho. Dynamics of HIV infection: A cellular automata approach. *Phys. Rev. Lett.*, 87(16), 15 October 2001.

Question on articles

Seiden, 1992[19]

- What is the meaning of clonotypic (p.329)
- What is the humoral immune system (p.329)
- What is the difference between in vitro and in vivo (p.329)
- What are the rheumatoid factor B-cells (p.331)
- Explain highlighted part 2.2 The Molecules (p.332)
- "and A cells that have bound to the antigen by means of an epitope". This does not match the description of A-cells of Fig. 1 (p.332)
- "It is not necessary... A-cell receptors"? What about A-cell receptors? (p.333)
- What do you think of the way antigen's peptides are broken up in the model? (p.336)
- How smart is the choice of hierarchy of binding? (p.339)
- I don't understand the experimental explaination at the bottom of p.349
- What is anergy (p.352 top)
- I don't understand auto-immunity! 4.4 more about anergy? (p.352)

Celada, 1996[6]

- How realistic is the point mutation suggested? (p.1352)
- Does a cell really stop multiplying if its mutation does not bind the antigen? (p.1352)
- Explain how mutations can come about and what are silent mutations? (p.1352)
- Should mutation of the framework of the cell cause death? (p.1352)
- For Jack: discuss/explain the probability of mutation (p.1352)
- For John: clarify the following paragraph (p.1352)
- How full of gaps it he naive system? (p.1357)

Kohler, 2000[14]

- What is polarization? (p.2)
- More on cytopathic viruses and other types of viruses (p.3)
- Is Fig.1 immune attack scenario good? (p.4)
- What are epithelial cells? viremia? (p.5)

Stewart, 1997[21]

- What are RF cells? (p.1728)
- I don't understand 2nd paragraph (p.1728)
- What is peripheral tolerance (p.1728)
- Are pathogenic RF produced by our system? (p.1728)
- NOTE: unfinished reading.

Perelson, 2002[16]

- What is a chronic disease? What are the other types of diseases? (p.28)
- Can you get non-chronic HIV? (p.28)
- What is meant by "Plasma level of HIV"? (p.28)
- What are reverse transcriptase inhibitors and protease inhibitors? (p.30)
- How come they found data to fit both models? (p.30)
- I would like to know more about MHC-tetramers. (p.32)

Kleinstein/Seiden, 2000[13]

- Explain how B-cells/antibody receptors are made from DNA (p.70).
- What, in IMMSIM, differentiates naive from memory cells?
- What are somatic mutations? (p.74)
- NOTE: 1 time step = 1 B-cell division!!! (p.74)

Bezzi, 1997[3]

- Definition of humoral vs. cellular? (p.3)
- In terms of terminology, is epitope or antigenic determinants better? (p.4)
- Again, what is physiologically different in memory cells? (p.4-5)
- Is it true that lymphocytes (what are they) are not able to begin a response without APC? (p.5)
- Note: B-cells have MHC class II molecules (p.5)
- Am I to understand that Th bind MHC II and Tc binds MHC I? (p.5-6)
- How does MHC molecules diversity arise and how is it distributed on the cell population? (p.7)
- In the simulation, APCs are generic, i.e. dendritic cells, macrophages, etc. are all treated as the same thing. Is this reasonable? (p.8)
- Do T-cells actually stimulate APCs? (p.10)
- "The CTL response is activated against intracellular viruses or parasites, so before introducing CTL we need an infection step". I don't understand. (p.11)
- The virus causes only B cells and APC to be infected and they are infected in the same way. How good is that? (p.11)
- How good is the cross-regulation mechanism suggested? How does the real thing work? (p.14)
- Why is competition better than cooperation between humoral and cellular? (p.15)
- Should antibody/antigen complexes dissociating because they were not eliminated by catabolism fast enough be included in the model? How significant is it? (p.16)
- Maybe their oscillating state is just lucky and has nothing to do with real pathologies exhibiting oscillations? (p.16)

Info:

- Cellular internal and humoral external description (p.6)
- MHC is less than 10 different clones in real, usually 1 or 2 in model (p.7)
- T-cells are selected in Thymus, B-cell repertoire is complete (p.7)
- APC bind all antigen but have same MHC as B-cells (p.8)

- Antibodies have a BCR and optionally a peptide (p.8)
- APR bind antigen with fixed probability of approx 0.002 (p.8)
- Virus infects B and APC cells with probability P_i per time step and grows inside cell at rate r such that $V^{n+1} = rV^n$. When $V > V_{\max}$ cells dies and $V_{\max}$ virions are freed (p.11)
- CTL not specific and interact and kill with prob P_k see explaination (p.14)
- cross-regulation: modelling IL and IFN-γ by making (p.15)
  - B-T binding down regulate CTL division
  - CTL activation down regulate B-cell division
- Not includes Th_1 and Th_2 (p.17)

Bernaschi/Castiglione, 2002[2]

- Again, what is anergy? (p.1)
- What are haematopoietic progenitor cells? (p.1)
- FOR ME: How is the CTL receptor represented? (p.3)
- NOTE: Each time step corresponds to 8 hours (p.3)
- NOTE: Check Weinand J. Theor. Biol 1990 p.409–428, Pandey and Stauffer J. Statis. Phys.