International Mathematical Forum, Vol. 7, 2012, no. 3, 109 - 128

Immune System Modelling by Top-Down and Bottom-Up Approaches

Carlo Bianca

Dipartimento di Matematica
Politecnico di Torino
Corso Duca degli Abruzzi 24, 10129, Torino, Italy
carlo.bianca@polito.it

Marzio Pennisi

Dipartimento di Matematica & Informatica
Università di Catania
Viale Andrea Doria 6, 95125, Catania, Italy
mpennisi@dmi.unict.it

Abstract

The biological immune system is a complex adaptive system that constitutes the defence mechanism of higher level organisms to micro organismic threats. There are lots of benefits for building an artificial (mathematical, physical or computational) model of the immune system. Medical researchers can use immune system simulation in drug research or to test hypotheses about the infection process. Given the wide range of uses for immune simulation and the difficulty of the task, it is useful to know what research has been conducted in this area. This paper provides a survey of the literatures in this field comparing and analyzing some of the existing approaches and models.

Mathematics Subject Classification: 34A34, 35Q92, 68T05, 82B31, 92B99

Keywords: ODE, PDE, Cellular Automata, Agents, Models

1 Introduction

Immune system (IS), one of the most fascinating schemes from the point of view of biology, physics, computer science and mathematics, constitutes the fundamental defense mechanism of the vertebrate animals, including human

beings, from invading from the pathogens and harmful foreign substances. Self organization and self adaptation, learning, recognition, memory and so on, are the primitive characteristics of the IS that allow it to be considered one of the most advanced and complex adaptive biological systems according to the following definition [4].

Definition 1.1 *A Complex Adaptive System (CAS) consists of inhomogeneous and adaptive agents (or particles) with the following characteristics:*

- *Agents interact each other and with the outside environment;*
- *The collective behavior cannot be simply inferred from the behavior of its elements;*
- *The alteration of only one agent or one interaction reverberates on the whole system.*

Typical examples of complex adaptive systems, among others, are: the brain, social systems, ecology, insects swarm, crowds. These systems are characterized by a global organization, which emerges from the interacting constitutive particles.

Definition 1.2 *An Emergent Property of a complex adaptive system is a property of the system as a whole which does not exist at the individual elements level.*

The emergent properties of IS included: The ability to distinguish any substance (typically called antigen Ag) and determine whether it is self or nonself; the ability to memorize most previously encountered Ag, which enables it to mount a more effective reaction in any future encounters.

In order to model the IS, researchers have to take care of the initiation of IS. Then all the potential and useful properties of the immune system will be performed on a holistic framework using the mathematical, physical, and computational methods. So far two theories are competing to explain the initiation of IS: the self-nonself theory [25] and the danger theory [30].

The self-nonself theory states that self-nonself recognition is achieved by having every cell display a marker based on the major histocompatibility complex (MHC). Any cell not displaying this marker is treated as nonself and attacked. Despite its successes this theory has several problems: Firstly, self is variable with time. Secondly, it was thought that self reactive cells are removed from the thymus (a process called *negative selection*).

The danger theory states that IS reacts if it receives danger signals no matter what caused it, hence self-nonself discrimination is not required. Again despite several successes and its elegance one thinks that this theory has some problems: if only the danger signals initiate IS then why some organs are

rejected after a long time of being transplanted? Moreover, danger signals in this model do not include starved cells or cells under pressure e.g. within or near a tumor.

Having shown that IS is a complex system a question now arises: how to model the IS? There are not only many theories (see [9, 24, 32, 54]) and mathematical models (see [40, 44, 51]) to explain the immune system process, but also many computer models (among others [8, 11, 12]). Nevertheless, the IS-modelling methods present in the literature can be grouped in the following two approaches:

1. Top-down approach: It solves the problems through a large number of entities. This approach does not emphasize the microscopic entities explicitly, but estimate the behavior in macroscopic level, exemplified by Ordinary Differential Equation (ODE) and Partial Differential Equation (PDE). The ODE and PDE-based models are all population-based, and the spatiality and topology which both depend on individual interactions are, in general, ignored.

2. Bottom-up approach: It is based on the synthesis of a complex from the activities on a lower system level; it emphasizes the microscopic level. This approach requires greater computational power in order to simulate a large number of significant entities in real world. From the model built by this approach, we can observe the interactions between entities specifically and study how they contribute to the emergence of global property. Cellular automata and (Multi)Agent-based methods are the most used bottom-up ones.

This paper critically analyzes the two above approaches for the modelling of the immune system. Meanwhile, it answers the following questions: How many approaches can be used to model the immune system? What are the advantages and disadvantages of the existing models of the immune system?

The rest of the paper is organized as follows. Section 2 overviews the outstanding behavioral aspects of the human IS [46]. Section 3 introduces the mathematical equations-based model and critically analyzes the related problems of this approach. The current state-of-the art in CA and agent-based simulations of the IS are discussed in Section 4 and Section 5 where we focus our discussion on some existing models based on multi-agent. Finally, we conclude the manuscript and present the perspective work in Section 6.

2 The Biological Immune System

The human immune system is a complex set of cells and molecules distributed throughout the body with the ability to memorize the foreign substances, such

as viruses, bacteria, fungi or other parasites that enter the body. While many general foreign antigens' characteristics are intrinsically known and recognized by the IS (by the so called natural "*aspecific*" component of the IS or innate IS), many other antigens (e.g. often mutating viruses) are genetically aprioristically unknown to the IS, as well as the possible location where the infection will take place. Therefore IS has components (the adaptive IS) that are able to tackle unknown challenges, remember them in order to efficiently face following infections of the same kind, covering all the body.

The basic cell of immune system is *lymphocyte* (or white blood cell). There are two main types of lymphocytes that differ in function and type of antigen receptor: *B-lymphocyte* (produced in bone marrow) and *T-lymphocyte* (trained in thymus). T cells, can be divided into different categories according to their function: Tc (*cytotoxic or killers*) cells are able to recognize cells infected by a specific antigen (i.e. containing it) and Th (*helpers*) cells, which assist macrophages and enhance the production of antibodies, stimulating the proliferation of the related B cells (Th cells may also help the activation of Tc cells). There are also *antigen presenting cells* (APC) or *accessory cells* (e.g. *macrophages*, *dendritic cells*), identifying antigens; *natural killer* cells (NK cells), *antibody* molecules and message molecules (*lymphokines*). Tolerance is a well known phenomenon by which the immune system does not respond against too small antigenic stimulations.

2.1 The Recognition Process

When an antigen enters the body, APC do the first recognition step. If the DNA fragment read from APC surface matches the specificity of Th cell, the cell gets activated. That results in triggering the own multiplication process and sending out the signal activating other immunity system elements such as B cells. When a B cell presents antigen to a Th cell, the latter is stimulated to secrete cytokines, which increase B cell proliferation and differentiation. Activated B cell differentiates to plasma cell that secretes thousands of antibody molecules, which are proteins able to bind to a specific antigen, neutralizing its possible harmful effects (see Figure 1). The antibodies have a specific structure that matches with the surface of recognised antigen. When the set of antibodies sticks on the antigen, it causes directly its dead or it marks the alien cell, so the other immune system cells can recognise and destroy it easier. There is also other elimination mechanism, involving the cells natural killers (NK) cells. NK cells recognise the cells that do not present clearly on surface cells activity. This is often the case of the cancer cells hiding their real harmful performance.

It worth stressing that each T lymphocyte is generated with a unique T cell receptor (TCR) that can recognize peptides presented on MHC molecules.

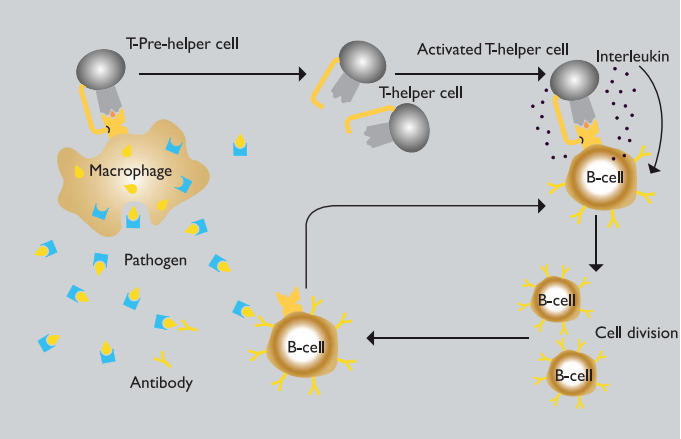

Figure 1: The recognition process of the immune system.

The human T cells are matured in the thymus where all the T cells with a TCR that cannot bind to a self-MHC molecule are killed. This process is called *positive selection*. It ensures that the host does not waste energy on T cells that have no chance of functioning as they are supposed to in the immune system. The other process that occurs in the thymus, as already mentioned in the introduction, is the negative selection where cells that react strongly to self-peptides are killed. This prevents *autoimmunity*.

3 Mathematical Equations-based Models

The most famous top-down approach is the ODE or PDE-based model. These models have been used to understand dynamical systems and this experience has led to many formal methods of analysis as well as an intuitive understanding of how many classes of dynamical systems behave.

The ODE-based approach views the entities (cells, molecules) in the model as homogeneousness and ignores the spatial structure of the biological system in the microscopic scale. The interactions are performed through differential equation based on parameters, populations and subpopulations. The main steps to derive an ODE-based model are the following:

1. definition of the granularity of the model (type of cells);
2. definition of the correlative interactions;
3. formulation of the ordinary differential equations;
4. analysis of the model in order to predict some results;

Although ODE-based approach is relatively easy to construct, there are three main shortfalls: **1)** The complexity of the model grows with increased number of populations, indeed the cells of the IS have many states (*quiescent*, *activated*, *productively infected*, *latently infected*, or *memory state*). This would mean dividing the cell population into more subpopulations, each of which is dedicated to one cell state, modeled by a single differential equation. **2)** It is not suitable for modelling spatial non-uniformity unless PDEs are used. **3)** It fails to correctly describe the phenomena observed by the macroscopic behaviors (system level) which are emergences resulted from the local interaction among entities.

Although models of the IS often involve nonlinearities that make the analytical solution especially difficult, it is sometimes possible to find an analytical solution, steady states, stability conditions and threshold expression to an ODE model for relatively simple systems or when simplifying assumptions are made [42, 44].

Finally, remind that solving a system of coupled differential equations for as many cell types as there are in the IS surpasses the capabilities of any modelling tool. As a result, ODE-based models generally assume homogeneity of entity types so as to limit the number of computable states while compromising on the realism of their predictions. Models that explicitly take spatial nonuniformity into consideration can lead to drastically different simulation results. To consider the spatial non-uniformity using equation-based description, PDE-based models are appropriate; they specify the dynamics with respect space variables in addition to the time dimension. This implies an increased number of coupled equations, making the model computationally more expensive. Moreover the identification of suitable (from the biological point of view) boundary conditions is an other limitation of this approach.

In the literature, there are many ODE-based models. Kuznetsov et al. [27] define a model for effector cells and tumor cells. They predict a threshold above which there is uncontrollable tumor growth, and below which the disease is attenuated with periodic exacerbations occurring every 3-4 months. They also show the model does have stable spirals, but there are no stable closed orbits. DeLisi and Rescigno [14] and Adam [1] consider the populations of IS and tumor cells showing that survival increases if the IS is stimulated and also show in some cases that an increase in effector cells increases the chance of tumor survival. Furthermore, they give a threshold for the chance of uncontrolled tumor growth. Nani and Oğuztöreli [33] developed a model of injection of cultured IS cells that have anti-tumor reactivity into tumor bearing host. Their model incorporates stochasticity effects on the IS-cancer interactions. Results of their model are that success of treatment is dependent on the initial tumor burden. They do not consider sensitivity, bifurcation, or stability analysis of the model. DeBoer et al. [13] describe most of the actors in the

tumor-immune dynamics. They are able to show both tumor regression (with a highly antigenic tumor) and uncontrolled tumor growth (for a low antigenic tumor). Kirschner et al. [26] used several differential equations to demonstrate the infection progress of HIV, interpreted some important characteristics and obtained drug therapy project inspired by both HIV infection and drug specialties. The model built by Nowak et al. [34] indicated that virus diversity speeds up infection progress of HIV. Other interesting and useful ODE-based models can be found in [2] and [52] instead PDE-based models can be find in the reference section of the book [9].

4 Cellular Automata-Based Models

Cellular automata (CA) is a bottom-up approach constituted by a set of identical elements, called *cells*, each of one of which occupies a *node* of a regular, discrete, infinite spatial network (a lattice of n-tuples of integer number), see [15]. Each cell can assume a *state* from a finite set S, and the automaton evolves in discrete time steps, changing the states of all its cells according to a *local rule* δ (or *transition function*) homogeneously applied at every step. The new state of a cell depends on the previous states of a set of cells, which can include the cell itself, and constitutes its *neighborhood*. Two important neighborhoods are considered: the Von Neumann neighborhood (a cell is connected to all the cells at a distance 1 along exactly one of the coordinates and with itself); the Moore neighborhood (a cell is connected to cells at distance at most 1 in each direction, i.e. diagonal connections are allowed). Thus the main characteristic of CA are discreteness and locality. From the repeated synchronous application of the simple local evolution rules, a global behaviors emerges, which can be very complex even in the case of the CA with two states and two neighbors (the so-called *elementary CA*); concepts like chaos or non-ergodicity have been used in CA [55].

CA approach was adopted to model the IS in order to study its adaptation and self regulation. Santos [47] and Hershberg et al. [21] modelled the IS to demonstrate the HIV-immune dynamics in physical space and Shape Space, respectively. Grilo et al. [18] presented a model that could display dynamical changes of both the virus and the immune cells based on the combination of Genetic Algorithm and CA. At the same time, Hershberg and Santos [48] respectively displayed the three-stage model of HIV infection in physical space and Shape Space based on CA. In [7] Beauchemin et al. present a simple two-dimensional CA model for influenza infections. Other CA-based models for the IS can be found in [3, 22, 28, 29, 41].

The advantages of CA lie in its abilities of allowing for spatial structure analysis and emphasizing the emergence from individual local interactions. There are some disadvantages in CA approach. The main limitations of this

model are that one site represents one cell and there is no cell diffusion in the model.

4.1 The Statistical Mechanics Approach

The (equilibrium) statistical mechanics aims to model systems with a huge number of interacting agents troughs a stochastic approach and analyzes the phases the system display by tuning control parameters.

The main advantage of this (canonical) approach is that the system is asked to respect thermodynamics in the following sense: The agents are supposed to interact implicitly defining an interaction energy (an Hamiltonian); this allows to ask for minimum energy and maximum entropy principles to hold, conferring a thermodynamical ground on the successive investigations.

When approaching the immune network modelling, statistical mechanics uses lymphocytes as agents and antibodies (for the B cells) or chitokynes (for the T cells) as messengers: we will briefly review the idea behind this technique by reading within these "*glasses*" the B-cell world. In modelling the B-world, it is possible to use this approach to understand the tolerance property and the self-nonself discrimination. As stated in [24], each antibody must have several idiotopes which are detected by other antibodies. Via this mechanism, an effective network of interacting antibodies is formed, in which antibodies not only detect antigens, but also function as individual internal images of certain antigens and are themselves being detected and acted upon. This can be understood as follows: At a given time a virus starts replication. As a consequence, at high enough concentration, it is found by the proper B-lymphocyte counterpart: let us consider, for simplicity, a virus as a string of information (i.e. 1001001). The complementary B-cell producing the antibody Ig1, which can be thought of as the string 0110110 (the dichotomy of a binary alphabet in strings mirrors the one of the electromagnetic field governing chemical bonds) then will start a clonal expansion and will release high levels of Ig1. As a consequence, after a while, another B-cell will meet 0110110 and, as this string never (macroscopically) existed before, attacks it by releasing the complementary string 1001001, that, actually, is a "copy" (internal image) of the original virus but with no DNA or RNA charge inside. The interplay among these keeps memory of the past infection.

Let us now formalize the interactions taking place within the system. We consider an ensemble of M identical lymphocytes σ_i^α, $\alpha \in \{1, ..., M\}$, all belonging to the i^{th} clone, $i \in \{1, ..., N\}$, and N all different clones. In principle M, the size of available "soldiers" within a given clone, can depend by the clone itself, such that $M \to M_i$. However, for the sake of simplicity, we are going to use the same M for all the clones, at least in equilibrium and in the linear response regime.

If the match among antibodies had to be perfect for recognizing each other, then in order to reproduce all possible antibodies obtained by the L epitopes, the immune system would need $N \sim \mathcal{O}(2^L)$ lymphocytes. Conversely, if we relax the hypothesis of the perfect match, only a fraction of such quantity is retained to manage the repertoire, such that we can define the following scaling among lymphocytes and antibodies:

$$N = f(L)\exp(\gamma L), \tag{1}$$

where $\gamma \in [0, 1]$ encodes for the ratio of the involved lymphocytes (the order of magnitude) and $f(L)$ is a generic rational monomial in L for the fine tuning (as often introduced in complex systems, $f(L) \sim \sqrt{L}$).

Interestingly, a far-from-complete system is consistent with the fact that binding between antigens and antibodies can occur even when the match is not perfect: experimental measurements showed that the affinity among antibody and anti-antibody is of the order of the 65/70 percent or more (but not 100%). Furthermore the experimental existence of more than one antibody responding to a given stimulus (multiple attachment) confirms the statement.

Thus in this approach we can think of each lymphocyte as a binary variable $\sigma_i^\alpha \in \{\pm 1\}$ (where i stands for the i^{th} clone in some ordering and α for the generic element in the i subset) such that when it assumes the value -1, it is quiescent (low level of antibodies secretion) and when it is $+1$ it is firing (high level of antibodies secretion).

To check immune responses we need to introduce the N order parameters m_i as local magnetizations

$$m_i = \frac{1}{M}\sum_{\alpha=1}^{M} \sigma_i^\alpha(t), \tag{2}$$

where i labels the clone and α the lymphocyte inside the clone's family. The vector of all the m_i's is depicted as $\mathbf{m}$ and the global magnetization as the average of all the m_i as $\langle m \rangle = N^{-1}\sum_i^N m_i$.

Now, let us turn to the external field and start with the ideal case of perfect coupling among a given antigen and its lymphocyte counterpart: let us label $\mathbf{h^i}$ the antigen displaying a sharp match with the i^{th} antibody, hence described by the string $\xi_{h^i} = \bar{\xi}_i$. In general, for unitary concentration of the antigen, the coupling with an arbitrary antibody k is h_k^i.

Following classical statistical mechanics, the interaction among the two can be described as

$$H_1^i := H_1 = -\sum_{i,k}^{N} h_k^i m_k, \tag{3}$$

such that if we suppose that at the time t the only applied stimulus is the antigen $\mathbf{h^1}$, all clones but 1, namely $i \in \{2, ..., N\}$, remain quiescent: the interaction term among the system and the stimulus is simply $H_1 = -h_1^1 m_1$.

Note that within this Hamiltonian alone the IS is at rest apart from the clone $i = 1$ which is responding to the external offense and that if we apply contemporary two external antigens $h_1(t), h_2(t)$, the response is the sum of the two responses.

Of course also the generic external input $\tilde{\mathbf{h}}$, stemming from the superposition of L arbitrary elementary stimuli, can be looked as the effect of a string $\tilde{\xi}$ which can written in the idiotype basis such that $\tilde{\xi} = \sum_{i=1}^{L} \lambda_i \, \xi_i$. Moreover, in order to account for the temporal dependence of the antigen concentration we introduce the variable $c(t)$ accounting for its load at the time t, such that, generically, several lymphocytes attack it as commonly seen in the experiments.

As we discussed, it is reasonable to believe that all the immunoglobulins have the same length L, on the other hand this is not obvious for antigens which may arrive from different organisms and places, such that their interactions with the immune system may be different. In a nutshell, let us only remark that APC desegregates the enemies in pieces of "information length" of order L and put them on the proper surface.

So far we introduced the (reductionist) one-body theory, whose "Hamiltonian" is encoded into the expression H_1. If we now take into account a "network" of clones we should include their interaction term H_2. Coherently with H_1 we can think at

$$H_2 = -N^{-1} \sum_{i<j}^{N,N} J_{ij} \, m_i \, m_j. \tag{4}$$

As anticipated, the Hamiltonian is the average of the "energy" inside the system and thermodynamic prescription is that system tries to minimize it. As a consequence, assuming $J_{ij} \geq 0$, the energies are lower when their constituents behave in the same way. For H_2, two generic clones i and j in mutual interactions, namely $J_{ij} > 0$, tend to imitate one another (i.e. if i is quiescent, it tries to make j quiescent as well *suppression*, while if the former is firing it tries to make firing even the latter *stimulation*, and symmetrically j acts on i).

It is natural to assume J_{ij} as the affinity matrix: it encodes how the generic i and j elements are coupled together such that its high positive value stands for an high affinity among the two. The opposite being the zero value, accounting for the missing interaction.

If we consider the more general Hamiltonian $H = H_1 + H_2$ we immediately see that in the case of $J_{ij} = 0$ for all i, j we recover the pure one-body description and the antigen-driven viewpoint alone. Different ratio among the weighted connectivity $w_i = \sum_j J_{ij}$ and h_i will interpolate, time by time, among two limits for each clone i.

As a consequence, the study of the underlying topology of the network built by the general Hamiltonian shaded lights on several phenomena in the

immune behavior [5, 6, 38]: The low dose tolerance is simply the inertia of the spin to flip when properly solicited by the external field (the antigen) because it is coupled with some nearest neighboring that keep it quiescent. Furthermore, if the amount of the weights of latter is enormous (as it can be possible for several nodes because the coupling strength is scale free due to the exponential of the antibody product) the flip is not possible a-priori implicitly offering an explanation to the self-recognition (provided the identification of the more connected nodes with the self-reactive ones).

5 (Multi)Agent-Based Models

Multi-agent method is a bottom-up approach where each agent can not only get local solution by itself, but also get global solution by cooperation with each other through local interactions. Multi-Agent-based Model (MAM) can naturally handle entity heterogeneity and spatial non-uniformity, and suffer less from the issue of directly designed dynamics. Even though specifying agent rules is intuitively straightforward, a complete MAM requires effort to build the basic framework that implements a virtual environment and agent communication channels, which are nontrivial. MAM, on the other hand, can afford many entity types and entity states without significantly affecting computational tractability. To model agent interactions realistically, MAM specifies rules that are dependent on spatial proximity; that is, agents should only interact when they are close to each other.

It is worth precising that MAM is able to exploit the emergence of the complex deterministic macroscopic functions from stochastic microscopic interactions. Through this approach, we can verify the hypothesis on how cells interact with each other. Although MA is a better method to model the complex system, its shortfall is to calculate massive agents, which limits its applications. The urgent task is to create a high effective parallel algorithm or platform to support this approach.

Some applications of this approach are reviewed in what follows. Guo et al. validated the three stages of HIV infection by using the multi-agent immune model [19]. Perrin et al. emphasized the diversity and mutation of HIV virus, which are important to the immune response in the latency [43]. Jacob in [23] presented a swarm-based, three-dimensional model for the human immune system, innate response and adaptive response. The model takes on a strengthening reaction to the previous encountering pathogen, which is the immune memory. Agents are spheres of different sizes and colours that move around randomly in the continuous three-dimensional environment; they interact with each other due to proximity, considering a spherical neighborhood. The fact that agents move in a continuous environment differentiates this model from the usual CA based approach.

Others agent-based simulations of the IS, among others, are the following. AbAIS which uses a hybrid approach that supports the evolving of an heterogeneous population of agents over a CA environment [17]. CAFISS divides the simulation using a rectangular grid, where each division represents a spatial location; but what differentiates CAFISS is the multithreaded asynchronous updating of the simulation [50], where each IS cell instance runs in its own thread, communicating with other cells using events.

C-IMMSIM and PARIMM, are simulator developed in the C programming language [8], with focus on improved efficiency and simulation size and complexity. In these adaptations, the IS response is designed and coded to allow simulations considering millions of cells with a very high degree of complexity.

IMMUNOGRID is a European Union funded project to establish an infrastructure for the simulation of the IS at the molecular, cellular and organ levels [16, 20]. Models included in this project include also SIMTRIPLEX [36, 37] and SIMATHERO [35]. Design of vaccines, immunotherapies and optimization of immunization protocols are some of the applications for this project. Grid technologies are used in order to perform simulations orders of magnitude more complex than current models, with the final objective of matching a real size IS.

SIMMUNE investigates how context adaptive behaviour of the IS might emerge from local cell-cell and cell-molecule interactions [31]. It is based on molecule interactions on a cells surface. Cells do not have states instead they have behaviours that depend on rules based on cellular response to external stimuli, usually external molecule interactions.

SIS [29] is based on a cellular automaton, with descriptive cellular states and rules that define transitions between those states, and aims to provide a larger picture of the IS, including self-nonself discrimination. SIS is capable of performing simulations with large number of cells (in the order of 10^6 to 10^9 cells), with linear correlation between simulation size and time

SENTINEL is a simulation platform [45], where the entities can move from on location to the other, responding to events that occur in the same or nearby locations. This approach consists in the use of specialized engines to manage physical and chemical interactions. As such, agents can move according to chemotaxis, motor capabilities and external forces acting on them. Simulation results were qualitatively consistent with in vivo experiments.

CYCELLS was designed for studying intercellular interactions, allowing to define cell behaviours and molecular properties, as well as having features to represent intracellular infection [53]. CYCELLS uses a hybrid model that represents molecular concentrations continuously and cells discretely. Each type of molecular signal (e.g. cytokines) is given a decay and diffusion rate.

5.1 The Celada Seiden Approach

One of the most notable multi-agent based approaches is represented by the IMMSIM model developed by F. Celada and P. Seiden [11, 49].

IMMSIM reproduces the ab-initio kinetic model that describes the interactions and diffusion of each relevant biological entity. It incorporates enough immunological detail to support studies involving real immunological problems and it developed a general modelling framework that could be used for multiple studies.

Probably the most important feature of this framework arises from the fact that it uses a bottom-up approach working at cellular scale without forgetting to represent fundamental features and phenomena observed at molecular scale, such as receptor binding and antigen processing. Time and space are discrete. Most implementations use a bidimensional lattice with hexagonal geometry with periodic boundary conditions. In this way each site has six identical neighbors and may contain not only one but multiple entities. Both cells and molecules are represented. Cells are modeled as individual agents, with their own life-time, biological behavior, position in the lattice, set of internal states and one or more receptors. Molecules are represented by their concentration per lattice-site, and by their molecular composition in the case of antigens and antibodies.

The first step of the simulation consists of initialization. After cells generation and thymic selection processes, the grid is populated by randomly placing the various cell types in the lattice according to their leukocyte formula. At each time-step all entities that lie in the same lattice-site can probabilistically interact each-others. As a result of an interaction, entities can change their internal status, thus inducing some consequences as the releasing of other entities (i.e. plasma B cells release antibodies), entities duplication, killing of entities or death. After the interaction phase ends, entities can probabilistically move to a lattice site in their neighborhood.

The framework has been formulated focusing on the most basic task of the IS: pattern recognition. Patten recognition is achieved by IS entities with the use of cell receptors. When a receptor and its ligand are able to match each other, there must be some regions of complementarity between the two. The set of characteristics which determines binding among molecules is called the *generalized shape* of a molecule. Antigens are recognized only by receptors that are in a small region in the shape space surrounding their exact complement.

To represent this concept of shape space the computational framework uses bit-strings to model receptors. With binary strings used to model cell receptors and the molecular structures of antibodies and antigens, many binding events can be simulated quickly, making it possible to study large-scale properties of the IS, even if this abstraction only mimics real behavior and mostly ignores the physical properties of receptor/ligand processes. When two entities that

lie in the same lattice are compatible for interaction from a biological point of view, their receptors are compared and the probability that the interaction occurs is defined as function of the Hamming distance between receptors, which is just the number of mismatching bits.

Consider for example a specific interaction between two entities x and y. Focusing on the concept of complementarity between receptors, the binding probability between the two entities is a function of the number of mismatching bits that will decline as the number of mismatching bits decreases. These binding probabilities will affect the cross-reactivity and the polyclonality of the response.

The evaluation of the success or failure of binding as a function of the binding probability $v(m)$ between two strings with hamming distance m depends on the calculation of a random number between 0 and 1. The binding is successful if the number is less than the probability and it fails if it is greater. Then $v(m)$, for two strings x and y having Hamming distance m, has to be maximum when there is a complete complementary between the two strings $(0 \leftrightarrow 1)$, i.e. when the Hamming distance between equal to the bit string length $(m = N)$, while it must be 0 when the receptors are identical, i.e. $m = 0$.

To correctly represent the natural clonal selection process, it is mandatory to allow only the activation of clones that do not differ too much. Thus $v(m)$ has to be set to 0 starting from a threshold value m_c of the Hamming distance, i.e. $v(m) = 0, m < m_c$. It is possible to define $v(m)$ as:

$$v(m) = \begin{cases} v_c^{(m-N)/(m_c-N)} & \text{for} \quad m \geq m_c \\ 0 & \text{for} \quad m < m_c \end{cases} \tag{5}$$

where N is the bit-string length, $v_c \in (0, 1)$ is a free parameter which determines the slope of the function and $m_c \in (N/2, N)$ is the assigned threshold value below which no binding is allowed. In this case m_c can be also imagined as the "size" of the recognition ball we discussed earlier.

Although any Celada-Seiden (CS) inspired automaton usually is more complex than the automata used by mathematicians and cannot be treated with deep analytical techniques commonly used to extract asymptotic behavior, it has some advantages. The automaton is stochastic, so it is possible to estimate the distribution of behaviors exhibited by the entire system, not just the average, in fact determinism is avoided. Spatial distribution represents an intrinsic characteristic of this kind of models, so spatial description can be achieved easily. The automaton is also able to accurately represent many of the biological processes of interest so that the approximations in the model are usually more biological in character than mathematical. Nonlinearities can also be treated easily because they are not intrinsically hard to handle. This also means that it is possible to add complexity or modify the model without introducing any

new difficulties.

The most important drawback of this technique is probably the same of any multi agent technique. Agents are followed individually and this requires a lot of computational resources. Even if the simulation of an entire individual is still far from being reached, simulation of tissues and organs can be however managed by modern laptops and desktop computers thanks to the appearance of modern multi-core powered machines even in the consumer and mainstream computer market.

From this modelling paradigm many biologically accurate models have been successfully used for simulating, for example, cancer [36, 37, 39], HIV [10] and atherosclerosis [35].

6 Conclusions and Perspective

There are several approaches to model the IS or parts of the IS, among which models based on differential equations are probably the most common. These models usually simulate how average concentrations of IS cells and substances change over time, identifying key aspects of the immune response. However, building and managing these equations, as well as changing them to incorporate new aspects, is not trivial nor intuitive, leading to sometimes mathematically sophisticated but biologically useless models or biologically realistic but mathematically intractable models. This approach also yields average behaviours, characteristics or concentrations of the IS components, ignoring the important aspects of the immune response, such as locality of responses and diversity of repertoires. Nonetheless they have had practical use regarding particular aspects of IS modelling. Both ODE and PDE assumes that local fluctuations have been smoothed out; Typically they neglect correlations between movements of different species; They assume instantaneous results of interactions. Most biological systems (including IS) shows delay and do not satisfy the above assumptions.

CA appears more suitable to model the IS but such systems suffer from a main drawback namely the difficulty of obtaining analytical results. The known analytical results about CA-type systems are very few compared to the known results about ODE and PDE.

Agent-based approach is also well suited for modelling the IS. The advantages include the possibility to determine behaviour distribution (and not just the average), rapid insertion of new entities or substances and natural consideration of non-linear interactions between agents. This approach is not without problems of its own: it lacks the formalism provided by differential equations, requires considerable computational power to simulate individual agents, and parameter tuning is not trivial. The success of any model based in these approaches depends on how well these and other problems are solved.

Recently was introduced a new mathematical method, which is between the bottom-up and top-down approaches: the kinetic theory for active particle (KTAP), see [9]. According to this theory, the complex system is divided into *functional subsystems* FS of cells characterized by a specific *biological function* (*activity*). KTAP studies the time course of the FS–*distribution function*; interactions modify the FS–*microscopic state* (*space–velocity–activity*). There are three kinds of interactions: *conservative* (do not modify the number of cells in the FS), *non conservative* (modify the number of cells in the FS), and *transitive* (generate a mutated cell and therefore a new FS). In the modelling of IS, transitive interactions can model the passage from B to plasma cells and for the recognition process, jointed functional subsystems may be defined (APC-Th-B or Th-B). Moreover discrete frameworks of this theory can be derived similar to CA-based model with the advantage that analytical results can be derived, see [9], and the reference there in.

Therefore, the computer simulation and/or the mathematical models can help biological researchers to further understand the mechanisms of the immune system and verify their hypotheses. Furthermore, it can provide some inspirations for biological researchers to develop some new medicines capable of restraining certain disease and to verify the suitability of the medicines for human body.

Acknowledgements

CB was partially supported by the FIRB project-RBID08PP3J-Metodi matematici e relativi strumenti per la modellizzazione e la simulazione della formazione di tumori, competizione con il sistema immunitario, e conseguenti suggerimenti terapeutici. The research was also partially supported by the INDAM-GNFM-Young Researchers Project-prot.n.43-*Mathematical Modelling for the Cancer-Immune System Competition Elicited by a Vaccine* and by University of Catania under PRA grant.

CB would like to thank Adriano Barra for his suggestions and discussions that have improved this paper.

References

[1] J. A. Adam, Effects of vascularization on lymphocyte/tumor cell dynamics: Qualitative features, Math. Comput. Modelling, **23** (1996), 1-10.

[2] C. T. H. Baker, G. A. Bocharov and C. A. H. Paul, Mathematical modelling of the interleukin-2 T-cell system: A comparative study of approaches based on ordinary and delay differential equations, *J. Theor. Med.*, **2** (1997), 117-128.

[3] S. Bandini, Hyper-cellular automata for the simulation of complex biological systems: a model for the immune system, in *Special Issue on Advance in Mathematical Modeling of Biological Processes* (D. K. Ed.), **3** (1996).

[4] Y. Bar-Yam, Dynamics of Complex Systems (Studies in Nonlinearity), Westview Press, 2003.

[5] A. Barra, E. Agliari, A statistical mechanics approach to autopoietic immune networks, J. Stat. Mech., P07004, (2010).

[6] A. Barra, E. Agliari, Stochastic dynamics for idiotypic immune networks, Physica A, **389** (2010), 5903-5911.

[7] C. Beauchemin, J. Samuel, J. Tuszynski, A simple cellular automaton model for influenza A viral infections, J. Theor. Biol., **232** (2005), 223-234.

[8] M. Bernaschi, F. Castiglione, Design and implementation of an immune system simulator, Comp in Biol and Med, **31** (2001), 303-331.

[9] C. Bianca, N. Bellomo, Towards a Mathematical Theory of Complex Biological Systems, World Scientific Publishing, 2011.

[10] F Castiglione, A Users Guide to CimmSim 1.1, in Dynamical Modeling in Biotechnologies, (F. B. and S. R. eds.), World Scientific, 2000.

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

[12] Y. Chenga, D. Ghersia, C. Calcagnoa, et al., A discrete computer model of the immune system reveals competitive interactions between the humoral and cellular branch and between cross-reacting memory and nave responses, *Vaccine*, **27** (2009), 833-845.

[13] R. J. DeBoer, P. Hogeweg, H. F. J. Dullens, at al., Macrophage T Lymphocyte interactions in the anti-tumor immune response: A mathematical model, The Journal of Immunology, **134** (1985), 2748-2758.

[14] C. DeLisi and A. Rescigno, Immune surveillance and neoplasia-I: a minimal mathematical model, Bull. Math. Biol., **39** (1977), 201-221.

[15] M. Delorme, J. Mazoyer (Eds), Cellular Automata, Kluwer Academic Publishers, Dordrecht, 1999.

[16] A. Emerson, E. Rossi, ImmunoGrid - The Virtual Human Immune System Project, Stud Health Technol Inform, **126** (2007), 87-92.

[17] A. Grilo, A. Caetano, A. Rosa, Agent based Artificial Immune System, in Proc. GECCO-01, LBP (2001), 145-151.

[18] A. Grilo, A. Caetano, A. Rosa, Immune system simulation through a complex adaptive system model, in Proc. of the 3rd Workshop on Genetic Algorithms and Artificial Life (GAAL99), Lisbon, Portugal, 1999, 1-2.

[19] Z. Guo, H. K. Han, J. C. Tay, Sufficiency verification of HIV-1 pathogenesis based on multi-agent simulation, in Proc. of the ACM Genetic and Evolutionary Computation Conference 2005 (GECCO'05), Washington D. C., USA, 2005, 305-312.

[20] M. Halling-Brown, F. Pappalardo, N. Rapin, et al., ImmunoGrid: Towards agent-based simulations of the human immune system at a natural scale, Philosophical Transactions A, **368** (2010), 2799-2815.

[21] U. Hershberg, Y. Louzoun, H. Atlan, S. Solomon, HIV time hierarchy: Winning the war while, loosing all the battles, Physica A, **289** (2001), 178-190.

[22] R. Hu and X. Ruan, A simple cellular automaton model for tumor-immunity system, In Proc. of Robotics, Intelligent Systems and Signal Processing, IEEE International Conference, **2** (2003), 1031-1035.

[23] C. Jacob, J. Litorco, L. Lee, Immunity through swarms: Agent-based simulations of the human immune system, Lecture Notes in Computer Science, **3239** (2004), 400-412.

[24] N. K. Jerne, Towards a network theory of the immune system, Ann. Immunol. (Paris), **125C** (1974), 373-389.

[25] H. Jiang and L. Chess, How the immune system achieves self-nonself discrimination during adaptive immunity, Adv Immunol., **102** (2009), 95-133.

[26] D.E. Kirschner, G. F. Webb, A mathematical model of combined drug therapy of HIV infection, J. Theor. Med., **1** (1997), 25-34.

[27] V. A. Kuznetsov, I. A. Makalkin, M. A. Taylor and A. S. Perelson, Nonlinear dynamics of immunogenic tumors: Parameter estimation and global bifurcation analysis, Bull. Math. Biol., **56** (1994), 295-321.

[28] D. Mallet, L. De Pillis, A cellular automata model of tumorimmune system interactions, J. Theor. Biol., **239** (2006), 334-350.

[29] J. Mata and M. Cohn, Manual for the Use of a Cellular Automata Based Synthetic Immune System, Conceptual Immunology Group, The Salk Institute, P.O. Box 85800, San Diego, CA 92186, 2005.

[30] P. Matzinger, The danger model. A renewed sense of self, Science, **296** (2002), 301-305.

[31] M. Meier-Schellersheim, G. Mack, SIMMUNE, a tool for simulating and analyzing immune system behavior, Tech. Rep. Institut für Theoretische Physik, Universität Hamburg, 1999.

[32] R. R. Mohler, C. Bruni, A. Gandolfi, A system approach to immunology, Proceedings of the IEEE, **68** (1980), 964-990.

[33] F. K. Nani and M. N. Oğuztöreli, Modelling and simulation of Rosenberg-type adoptive cellular immunotherapy, IMA Journal of Mathematics Applied in Medicine & Biology, **11** (1994), 107-147.

[34] M. A. Nowak, R. M. Anderson, A. R. Mclean, T. F. Wolfs, J. Goudsmit, R. M. May, Antigenic diversity thresholds and the development od AIDS, Science, **254** (1991), 963-969.

[35] F. Pappalardo, A. Cincotti, S. Motta, M. Pennisi, Agent based modeling of atherosclerosis: a concrete help in personalized treatments, Lecture Notes in Artificial Intelligence, **5755** (2009), 386-396.

[36] F. Pappalardo, P. L. Lollini, F. Castiglione, S. Motta, Modelling and simulation of cancer immunoprevention vaccine, Bioinformatics, **21** (2005), 2891-2897.

[37] A. Palladini, G. Nicoletti, F. Pappalardo, et al., In silico modeling and in vivo efficacy of cancer preventive vaccinations, Cancer Research, **70** (2010), 7755-7763.

[38] G. Parisi, A simple model for the immune network, Proc. Natl. Acad. Sc., **87** (1990), 429-433.

[39] M. Pennisi, F. Pappalardo, A. Palladini, et al., Modeling the competition between lung metastases and the immune system using agents, BMC Bioinformatics, **11** (Suppl 7):S13, (2010).

[40] A. S. Perelson, Immune network theory, Immunol Rev., **10** (1989), 5-36.

[41] A. S. Perelson, Modelling viral and immune system dynamics, Nature Reviews Immunology, **2** (2002), 28-36.

[42] A.S. Perelson and G. Weisbuch, Immunology for Physicists, Rev. Mod. Phys., **69** (1997), 12191268.

[43] D. Perrin, H. J. Ruskin, J. Burns, M. Crane, An agent-based approach to immune modelling, Lecture Notes in Computer Science, **3980** (2006), 612-621.

[44] D. Přiikrylov, M. Jlek, J. Waniewski, Mathematical Modeling of the Immune Response, CRC Press, 1992.

[45] M. Robbins, S. Garrett, Evaluating theories of immunological memory using large-scale simulations, Lecture Notes in Computer Science, **3627** (2005), 193-206.

[46] I. Roitt, Essential Immunology, Blackwell, 1994.

[47] R. Santos, Immune responses: Getting close to experimental results with cellular automata models. In: Annual Reviews of Computational Physics VI, World Scientific Publishing Company, Singapore, (1999), 159-202.

[48] R. Santos, S. Countinho, Dynamics of HIV infection: A cellular automata approach, Phys. Rev. Lett., **87** (2001), 168102-1-168102-4.

[49] P.E. Seiden, F. Celada, A model for simulating cognate recognition and response in the immune system, J. Ther. Biol., **158** (1992), 329-357.

[50] J. Tay, A. Jhavar, CAFISS: a complex adaptive framework for immune system simulation, in Proc. of the 2005 ACM Symposium on Applied Computing, (2005), 158-164.

[51] F. J. Valera and J. Stewart, Dynamics of a class of immune network I. Global stability of idiotype interactions, J. Theor. Biol., **144** (1990), 93-101.

[52] V. Yorama, C. Gillesb, C. Carsonc, Mathematical models of the acute inflammatory response, Current Opinion in Critical Care, **10** (2004), 383-390.

[53] C. Warrender, Modeling intercellular interactions in the peripheral immune system, PhD thesis, The University of New Mexico, 2004.

[54] G. Weisbuch, A shape space approach to the dynamics of the immune system, J. Theor. Biol., **143** (1990), 507-522.

[55] S. Wolfram, Cellular Automata and Complexity, Addison-Wesley, Reading, MA, 1994.

Received: August, 2011