Self-Assertion versus Self-Recognition: A Tribute to Francisco Varela

Hugues Bersini
IRIDIA-CP 194/6
Université Libre de Bruxelles
50, av. Franklin Roosevelt
1050 Bruxelles – Belgique
bersini@ulb.ac.be

Abstract

Ten years ago, a group of researchers, led by Francisco Varela, were proposing an alternative vision of the immune system main behavior and function. I was part of this group. This new vision saw the immune system not as behaving distinctively with self and non-self or according to any dichotomy imposed a priori and from outside (the self-recognition vision), but rather as behaving in a unique way. From this indifferent behavior, any external impact would progressively been treated in two different ways, reactive and tolerant, but now, consequently and from inside the system (the self-assertion view). This paper will recall, through a very artificial simulation, the difference existing between these two visions. Also at that time, we believed that, from an engineering perspective, this new vision, emphasizing more the adaptability and the need for endogenous constraints than the recognition and the defensive ability, although less obvious to accept than the classical defensive one, should be more beneficial. These last ten years proved that we haven't been convincing enough, and in this paper I resume the crusade.

1 INTRODUCTION

Ten years ago, Varela, Coutinho and Stewart (Varela et al., 1988; Varela and Coutinho, 1991; Stewart, 1994) were proposing and defending a new vision of the immune system, largely in the continuation of Jerne's intuition and model (Jerne, 1974), in which the "self" and "foreignness" dichotomy collapses, for the system is complete unto itself. Based on simulations of the immune idiotypic network and some experimental data, they published several papers, although not in the mainstream journals of immunology. In an idiotypic network, there is no intrinsic difference between an antigen and an antibody, and any node of the network can bind and be bound by any others. My role is this group was two-fold. At that time not a biologist and still not today, I was responsible with Vera Calenbuhr and Vincent Detours for the development of a series of computer simulations that have been described in (Detours et al., 1994; Bersini and Calenbuhr, 1996; Calenbuhr et al., 1996). I was also responsible for trying to initiate the influence of this new "reading" of this biological system for the conception of engineering artifacts. A mission I tried to fulfill in (Bersini and Varela, 1993; Bersini, 1999).

Although we pushed hard for this alternative vision, we need to admit today that the classical view of the immune system as a defensive system, first able to distinguish between dangerous and innocent external impact, and thus to defense against the dangerous ones, has been the most influential one from an engineering perspective. It was clearly the most appealing to adopt, but it's a pity. First, this is a vision that is facing more and more opposition among the biologists themselves. But beyond that, I am convinced that we don't need to know how the immune system distinguishes, if it does so, between good and bad stimuli, in order to build performing two-classes classification system or any pattern recognition mechanism. Also, we don't need to know how the system creates good markers of self, if it does so, to build performing clustering and self-organizing systems. And finally, we don't need to know how the system protects the body from external damages, if it does so, to build good protective system for computers. The Panda antiviral system that by computer has adopted for one year now is one good illustration, among many, of that. I don't believe the developers of such effective software needed to know anything about how the immune system fights natural virus to develop their system for artificial ones.

In the first section, largely relying on a recent excellent survey of the Stanford Encyclopedia of Philosophy (Tauber, 2002), I will try to summarize what main lines of criticisms attack the vision of immune system as able to distinguish self from non self and able to protect from non-self. In the second section, I will present a very simple software simulation that will make easier to understand the difference between the self-recognition and the self-assertion views. This simulation is very reminiscent of a lot of simulations that we published years ago, although I'll try to simplify it to the basics in order to really shed the light on the key differences.

Finally, the last section will try to defend why the self-assertion view should inspire in a more creative way the conception of engineering artifacts. This vision leads to strongly adaptive systems, both parametrically and

structurally, but whose adaptability mainly aims at satisfying endogenous constraints instead of responding to exogenous impacts. This constraint satisfaction might provide the system with several adaptive advantages as, for instance, the capacity to respond to a large diversity of external stimuli and to memorize in an economical way a repertoire of adapted responses when facing a non-stationary environment.

2 THE PROBLEMS WITH SELF AND NON-SELF

Although it's important not to confuse the alternative view proposed by Polly Matzinger (Matzinger, 2002), today best known critics of classical immunology, with the one proposed by Varela's group, part of the criticisms addressed by Matzinger to classical immunology has to be taken as important flaws of this classical view. Why the immune defenses do not protect us from the air we breathe, the food we eat, the fetuses we carry, the tumors that kill us (even then it should)? Why are a lot of our lymphocytes autoreactive without any sign of autoimmune diseases? Indeed, a lot of evidences in recent years have shown that autoimmunity is a normal finding in healthy individuals. Clearly the problem with self and non-self lies in the determination, namely the nature and the location, of the frontier. The designation of "self" and the "other" ignores that such neat divisions were adopted with a certainty that remain problematic.

One first relaxation to the self/non-self dual view of the immune system is to maintain the duality, i.e. the immune system keeps two ways of being in response to external impact: defensive and tolerant, but not depending on an evasive frontier to cross. It is the position adopted by Matzinger who insists in getting rid of the self/non-self discrimination as the central tenet of immunology. What she proposes instead is an immune system that just fights what is dangerous for it. So the dichotomy is maintained but self/non-self is simply replaced by dangerous/inoffensive. The fact that this move finally consists in this simple semantic substitution makes a lot of immunologists very skeptic against Matzinger position. However it appears that fighting danger rather than foreignness entails doctors to adopt therapeutic strategies that show great successes for certain serious diseases.

Now exploring more logically Matzinger's position, and although the full model is still somewhat confused, it is important to understand better what does the immune system see as dangerous and why it does so. One view, the less radical one, would see the danger as resulting from some specific characteristic of the external perturbation. It might be an additional feature of the invading antigen. In such a case, from the outside, the external impact will be, prior to any interaction, dangerous or not, and the immune system would still need to somewhat behave in a dichotomous way, first recognizing the danger then fighting it. The external environment of the immune system will still be separated in two zones: a dangerous and an inoffensive one. This interpretation of what is dangerous or not is not such an exciting one, because it still demands from the system the ability to discriminate and to defend. The self/non-self frontier is just re-defined but still exists outside the system.

The most radical view, and for reasons to be discussed later, makes Matzinger and Varela closer than they appear to be (in her "science" article Matzinger said that after many years of finding Varela's model intriguing she finally agrees). Varela's view would see the danger as a consequence of the interaction between the external impact and the current state of the immune system. In such a case, a stimulus is no more dangerous per se, but is dangerous in the current context of the immune system. An outside separation in two classes, making the immune system behaves in two ways, simply collapses. We remain with an immune system behaving in one only way but, depending on its current state and the nature of the impact, proposing different responses to it. For instance, a same external impact could drive the system to react differently at different times.

The reason why this second, more innovative interpretation, is akin to Varela and his group vision can be easily understood by reading the following excerpt from the Stanford Encyclopedia (Tauber, 2002) about the later vision:

"*When the immune system is regarded as essentially self-reactive and interconnected, the meaning of immunogenicity, that is reactivity, must be sought in some larger framework. Antigenicity then is only a question of degree, where "self" evokes one kind of response, and the "foreign" another based not on it intrinsic foreignness, but rather because the immune system sees that foreign antigen in the context of invasion or degeneracy. …. In the Jernian network, "foreign" is defined as perturbation of the system above a certain threshold. Only as observers do we designate "self" and "non-self". From the immune system perspective it only knows itself…. While host defense is a critical function, it is hardly the only one of interest. Indeed the immune system might be regarded as primarily fulfilling an altogether different role if its phylogeny is carefully examined…. Immune reactivity is determined by context where agent and object played upon each other….."*

3 A VERY ARTIFICIAL MODEL TO DISTINGUISH THE TWO VISIONS

In this section, I will describe a very simple model built in a two dimensional space and very reminiscent of several models that I did build years ago with my colleagues John Stewart, Vera Calenbuhr and Vincent Detours (Detours et al. 1994, Calenbuhr et al., 1996, Bersini and Calenbuhr, 1996). It will provide an easy to understand illustration of the difference between the self-recognition and the self-assertion visions.

We will suppose that any immune cell (they could be antibodies) be identified by its position in a two

dimensional space. In agreement with the key-lock binding of immune cells with antigens, we will also suppose, like indicated in figure 1, that any immune cell exerts an affinity in a zone symmetrically situated with respect to its position. What we want to model by this artificial construction is the possibility for a cell to bind an antigen when it presents a shape symmetrical with respect to the one of the antibody. The affinity is not restricted to the symmetrical position but extends to a square domain of size L, the strength of the binding decreasing with the distance to the center of the square.

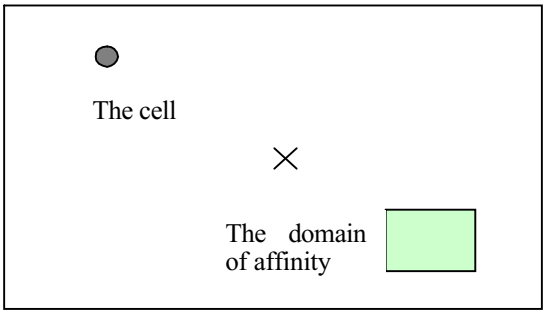

Figure 1: One cell and its symmetrical domain of affinity

At every simulation step, a new cell is randomly recruited anywhere in the system. It is added with an initial concentration Ci(0), C for concentration and i indexing the cell. Suppose the center of symmetry Xo, Yo. Suppose a cell i situated in position cx and cy and having concentration Ci(t) at time t . It will exerts an affinity on any cell situated in position x, y with value given by:

$$affinity = Ci(t)*(L - (|2Xo - cx - x| + |2Yo - cy - y|)/2)$$

So all cells will exert on any cell j a field of affinity *Affj* obtained by summing this affinity for all the cells currently present in the system:

$$Affj = \Sigma_i \, affinityOfCell_i$$

3.1 THE SELF-RECOGNITION VIEW

In the classical view, we will assume that only the antigens are subject to this field of affinity and, reciprocally, only the antigens exerts affinity on the cells.

At every time step, the concentration of the cell j evolves in the following way:

if (low < Affj < high) Cj(t) = Cj(t) +1

else

Cj(t) = Cj(t) – 1

if Cj(t) = 0 the cell j disappears from the system

In this case *Affj* is computed just by summing the field exerted by the antigens.

$$Affj = \Sigma_i \, affinityOfAntigen_i$$

Consitently with immunological facts, the cells will grow in concentration, i.e. simulating an immune response, if they receive a stimulating field in between two thresholds: *low* and *high*, whose precise values must be known for the simulation to run. The field must be sufficient enough but not too high due to the bell shape curve of the maturation and the proliferation of B lymphocites and antibodies.

From their side, the antigens will just decrease in concentration if they are bound enough by the immune cells. Take an antigen j, if finding enough cells to bind it i.e. *if Affj > low*, it will decrease in concentration according to:

if (Affj > low) Cj(t) = Cj(t) – k(Affj/low)*

k is a time rate

if Cj(t) = 0 the antigen j disappears from the system

The simulation proceeds as follows. Initially, cells are recruited randomly in the system, but in the absence of antigens, so with no stimulating field, they can't survive and disappear as soon as they get in. When an antigen enters the space, the simulation behaves as illustrated in figure 2.

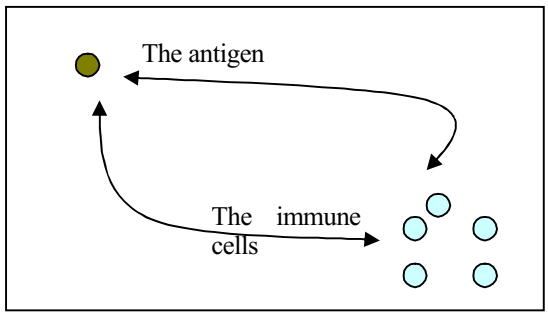

Figure 2: The reciprocal stimulation and elimination of antigen and immune cells

The antigen, now by the field of affinity it exerts symmetrically to its position, allows some cells to survive and to grow in concentration. These cells in turn exert a suppressing field on the antigen. The antigen will decrease in concentration until it disappears completely. Once it is cancelled from the system, the cells responsible for its disappearance are no longer stimulated and slowly

die, driving back the whole system to the initial situation: random recruitment of not surviving cells.

Playing with the concentration increasing and decreasing rates (for instance the constant *k*), the immune cells can take some time to disappear, akin to a sort of inertial memory of the antigen encounter. The next time a same antigen gets in, its cancellation will be faster like for any secondary immune response.

What needs to be understood, in contrast with the section to come, is that, in the classical case, cells show affinity only with antigen and not at all among themselves, although they occupy the same two-dimensional description space. Although nothing really differentiates an antigen from any cell, there must be a magical demon to tell the cells that the dot in the space is an antigen and not a cell.

3.2 THE SELF-ASSERTION VIEW

In this less classical view, all cells bind to all cells. To quote again the Encyclopedia: "*there is no essential difference between the "recognized" and the "recognizer", since any given antibody might serve either, or both, functions. Immune regulation is based on the reactivity of antibody with its own repertoire forming a set of self-reactive, self-reflective, self-defining immune activities*".

In the simulation now, the way we will compute the *Affj* received by any cell is as follows:

$$Affj = \alpha \Sigma_i \, affinityOfCell_i + \beta \, \Sigma_i \, affinityOfAntigen_i$$

This time, the affinity received by any cell is a weighted sum of the exogenous stimulation of the antigens and the endogenous stimulation of the cells themselves. Give a value 0 to α and you are back to the previous case. There is no way for any cell to discriminate between the exogenous and the endogenous impact. All impacts mix together to stimulate the change in concentration of any immune cell.

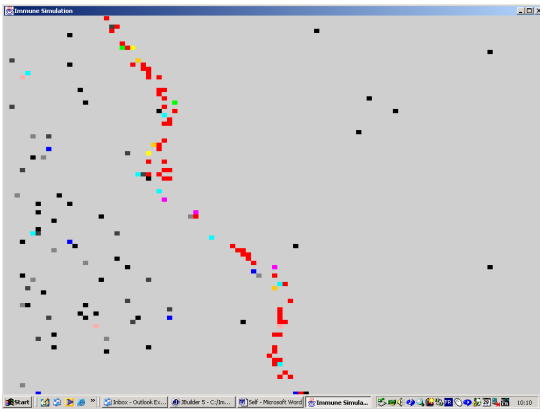

Figure 3: Snapshot of the self-assertion simulation

In the absence of any antigen, the simulation goes as shown in figure 3 (a snapshot of the simulation). The simulation slowly produces a sort of line or a band of self-sustained cells. Due to the way the affinity is computed (symmetrically with respect to the center of the space), cells in the line mutually stimulate themselves. A part of the line sustains another part of the same line. We speak of self-assertion since, indeed, this line can be roughly viewed as a signature of the immune self.

As shown in figure 4, first the system needs to be triggered off, and during the first time steps a lot of cells are recruited and very few are killed. During a second period, when the line of self-sustained cells begins to form, a lot of cells (not integrated in the line) are killed. This elimination phase can be roughly assimilated to the so called clonal selection phase taking place during the prenatal development and exercising a purging function of self-reactive cells. It is the period during which the tolerant zones are learned by the system itself. Finally the system tends to stabilize its rate of destruction and, while working at normal regime, integrates and kills new cells at a constant rate.

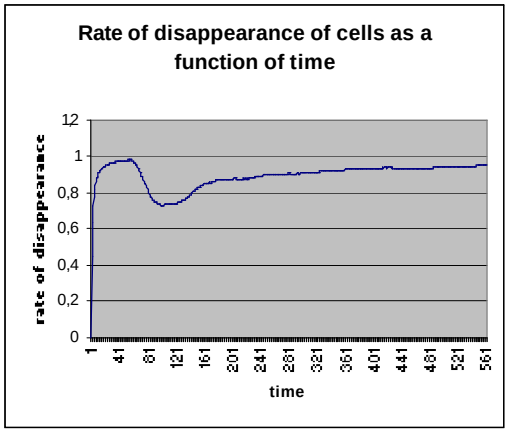

Figure 4: rate of disappearance of cells as a function of time

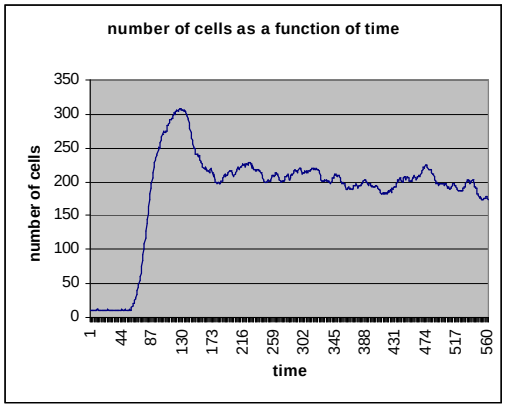

Figure 5. Number of cells as a function of time

In figure 5, plotting the number of cells as a function of time, again you can see the three successive phases of the simulation: first very few cells, then a short triggering period when a lot of new cells are recruited, and finally a stable regime.

One key observation is that the presence of the line divides the space in two zones: a reactive zone on the right and a tolerant zone on the left. If you add an antigen on the left, it will be tolerated since there is no cell on the right able to bind it. In contrast, an antigen on the right will be rapidly destroyed since a lot of cells on the left are still able to bind it. Basically the shape of the line is responsible for this division of the space in these two zones.

It must be clear that these two resulting zones are not shaped from the outside. There is no a priori division of the space into reactive and tolerant zones. This division is self-asserted by the system. The system creates, by its own evolution, its own zone of tolerance and own zone of reactivity. You might ask why a completely symmetrical simulation lead to unsymmetrical outcome. It is a simple artefactual effect of the random generation of cells that is amplified in time.

However the final separation of the space in a tolerant and a reactive zone will always be in relation with the history of the system. If you initially favor the recruitment in a given zone, this zone will naturally tend to become the tolerant one, a finding that qualitatively agrees with the Burnett's clonal selection theory.

This qualitative phenomenon i.e. the emergence of some geometrical patterns of self-sustained cells dividing the space in tolerant and reactive zones is very robust and largely independent of the values given to the parameters. This explains why I don't need to indicate the precise the values taken by the parameters of the simulation: α, β, *low, high*. The same qualitative outcome will be observed for a large range of values.

However, what's of crucial importance here is that no recognition and discrimination is at work. The system does not need to discriminate between an immune cell and an antigen, between self and non-self or along any prior arbitrary division applied to its biological environment.

4 TAKING AN ENGINEERING PERSPECTIVE

We are not biologists but are trying to be influenced by biology to create new ways of designing useful artifacts. As I already wrote in a previous paper (Bersini, 1999), I believe that the self-recognition interpretation of the immune system is not the most fruitful one. The basic reason is that this interpretation does not need biology to be expressed and understood. That the immune system can discriminate between two classes of external impacts can be easily translated into a classical pattern recognition problem. So far I haven't read any better ways of classifying, clustering data or constructing defensive systems, beyond classical ones, which have been discovered thanks to the immune analogy.

Also I don't want to pretend that the self-assertion view has been much more productive. Obviously, there have been fewer trials. In (Bersini, 1999) I discussed several engineering applications I was involved with that gained some benefits from applying here and there hints coming from immunology. The principal one that was implemented in all these application is the endogenous double plasticity inspired from immune networks. This endogenous double plasticity complies with the following principles:

1. the structural adjustments (akin to the recruitment of new cells) intermittently occur following a longer time scale than the parametric adjustment.
2. the structural plasticity amounts to the addition of new elements and the suppression of redundant elements from the system
3. again like in the artificial world shown above, the structural adjustments are dependent on the temporal evolution of the internal parameters (in the simulation, the current concentration of the cells). When and how to perform a structural change should depend on data related to the dynamics of the parametric change. So the network endogenous behavior and now exogenous criteria will guide these structural changes. Remember the immune system which only sees and knows itself.
4. these structural endogenous alterations have to be done in a network spirit by applying simple heuristics like "compensate for the weakest elements", "maintain diversity", "suppress redundancy".

In the same paper, I presented three practical illustrations of systems capable of evolving in time their structure and parameters while executing their task: neural net classifiers, autonomous agents that adapt by reinforcement learning, and controllers of chaotic systems.

In none of them, the biological influence was so strong to claim that I could not have done the same in the absence of any immunological knowledge, but in all of them, the way I tackled the problem, reinforcing the adaptability and the respect of the endogenous constraints, came from this knowledge.

5 CONCLUSIONS

The paper basic motivation is to better understand the difference existing between the classical self-recognition and the more "exotic" self-assertion visions of the immune system. Although the later is gaining more and more attention in the biological community, it is not

receiving the same attention in an engineering perspective. I believe it should.

We all need to admit that the immune algorithms, whatever they really turn out to be, did not provoke the same wave of interests as genetic algorithms or neural nets did for engineering applications. One key reason could be that in their initial presentation, both GA and neural nets were proposed in a very coherent and convincing way as simple algorithms, easy to implement, and associated with a precise and well-defined operational context: optimization for GA and pattern-recognition for neural nets. As a matter of fact, a lot of researchers discovered the whole problematic of optimization or pattern recognition by applying GA or neural nets.

Immune algorithms were never sold in such a persuasive way. No precise and complete algorithm was proposed and no clear operational context was associated with them: pattern recognition, defensive system, optimization, or robotics? Now, when maturing, researchers slowly realize that just playing with the initial basic GA or the initial neural nets does not give good results. What they do instead is to preserve some good mechanisms originating from this biological inspiration: population/selection based search or crossover for GA, multiplayer for neural nets, but turn them into a more operational form.

This is really what we are all doing today, based on our respective understanding of how the immune system behaves: gleaning here and there some inspirations and turning them into a more operational form. However, we should keep open our mind to more marginal voices, since if they are telling the truth, the radical revision they will entail in their community could have repercussions up to our own.

References

N. Jerne (1974) Towards a network theory of the immune system – *Annals of Institute Pasteur/Immunology (Paris)* – 125C: 373-389.

F. Varela, A. Coutinho, B. Dupire and N. Vaz (1988). Cognitive Networks : Immune, Neural and Otherwise, in A. Perelson (Ed.) – *Theoretical Immunology*, Vol.2 SFI – Series on the Science of Complexity – Addison Wesley, New Jersey pp. 359 – 375.

F. Varela and A. Coutinho (1991): Second Generation Immune Network – in *Immunology Today,* Vol. 12 No5 – pp. 159-166.

H. Bersini and F. Varela (1993) : "The Immune Learning Mechanisms: Recruitment Reinforcement and their applications" in *Computing with Biological Metaphors* - Chapman and Hall - R. Patton (Ed.)

V. Detours, Bersini, H., Stewart, J. and F. Varela (1994): Development of an Idiotypic Network in Shape Space - in *Journal of Theoretical Biology* - 170.

J. Stewart (1994) Cognition without neurons: Adaptation, learning and memory in the immune system. *CC AI 11*: pp.7-30.

Calenbuhr, V., Varela, F. and H. Bersini (1996). Immune Idiotypic Network - in *International Journal of Bifurcation and Chaos* - Vol. 6 No 9 - pp. 1691-1702

H. Bersini, and V. Calenbuhr (1996). Frustrated Chaos in Biological Networks - in *Journal of Theoretical Biology*, Vol. 188, No 2, pp. 187-200.

H. Bersini (1999). The Endogenous Double Plasticity of the Immune Network and the Inspiration to be drawn for Engineering Artifacts. In "*Artificial Immune Systems and Their Applications"* - Springer Verlag, pp.22 - 44

P. Matzinger (2002) The Danger Model: A Renewed Sense of Self – In *Science* – Vol. 296. pp. 301 – 305

A. Tauber (2002) The Biological Notion of Self and Non-Self – *Stanford Encyclopedia of Philosophy* – http://plato.stanford.edu/entries/biology-self.