# Useful Examples and a New Model of the Immune System

**Uri Hershberg**

Interdisciplinary Centre for Neuronal Computation
Hebrew University of Jerusalem,
Israel.
*uriher@cc.huji.ac.il*

*ABSTRACT: It is central to the development of cognitive systems that they have in their environment certain examples that are both high frequency and generic to some general property of their interaction with the environment. We have called such examples Useful Examples and suggested that in order to understand cognitive systems and their capabilities we must study these examples. In the following pages I will use a model of immune cell repertoire dynamics in the immune response to pathogens as an illustration of this point. This model is based on the application to immunology of the theory of Useful Examples and on existing knowledge of immune dynamics. Its conclusions place self-reactivity at the basis of immune specificity of reaction, causing a paradigm shift in existing conceptions about immune functions and sensitivities. In closing I will go over conclusions about the general importance of Useful Examples in the study of cognitive systems.*

KEY WORDS: *complex systems, immunology, cognitive systems, learning, development, self organization .*

Revue d'Intelligence Artificielle. Volume 17 – No. 5/6 2003, pages 763 to 773

## 1. Introduction

When talking of analogies and their use in the behaviour and development of cognitive systems we focus mostly on how much the two 'ends' of the analogy are the same. We ask how this sameness is used by cognitive systems to relate and learn about the world. This ignores a basic facet of analogies, which is that by finding things similar we are implicitly saying that they are not the same. What we learn form the analogy are the differences. The similarities are used as framework / scaffold to teach us about the new and the different.

We have previously suggested that cognitive systems share a common strategy of interaction with their environment to acquire their capabilities (Hershberg & Efroni 2001). Cognitive systems must learn to be cognitive systems. We have presented a view of the system's mode of learning in which we described it as a direct result of unsupervised interactions with concrete examples of the environment (see fig. 1 and Hershberg & Ninio 2003).

Such a simple form of learning is sufficient to allow the system an understanding of its environment because the environment is ordered. There are a few examples which are encountered with high frequency while most other examples are encountered less frequently. The high frequency examples are not ubiquitous by chance. Their prevalence reflects a multiplicity of usages for these examples within the interaction of system and environment. Viewed as separate instances, the variety of contexts in which such examples exist appears a deterrent to learning. However, when examined in the light of the complete set of interactions between system and environment, this variety becomes an asset. In effect, these examples embody important properties of the environment for the cognitive system.

These examples are constantly encountered; once acquired, they can greatly enhance the ability to learn new less generic examples by similarity-matching; and finally, both in their own right and due to their similarity to other examples, they will be consistently reinforced by the system's interaction with its environment. We call such ubiquitous and generic examples of the environment Useful Examples because they are useful to the system for learning the environment, and its relevant general properties. The distribution of examples in the environment is not arrived at by chance. For each cognitive system the reason for signalling out of specific Useful Examples is different. It depends on the specific types of interactions which the cognitive system undergoes with its environment. This reflects the fact that cognitive systems are fitted by evolution to specific niches. Based on genetic inheritance and previous cognitive development a cognitive system has certain tendencies, which give it the framework of its environment.

As an example of the application of this theory of Useful Examples to the study of cognitive systems I will in the rest of this paper, deal mostly with the immune system. I will show how studying the Useful Examples, used by a given system to acquire its capabilities, can influence our scientific understanding of that cognitive

system and its function. I will explain how by leading from the theory of Useful Examples we reach a new model of immune reaction and the importance of self reactivity to the proper maintenance of immune capability.

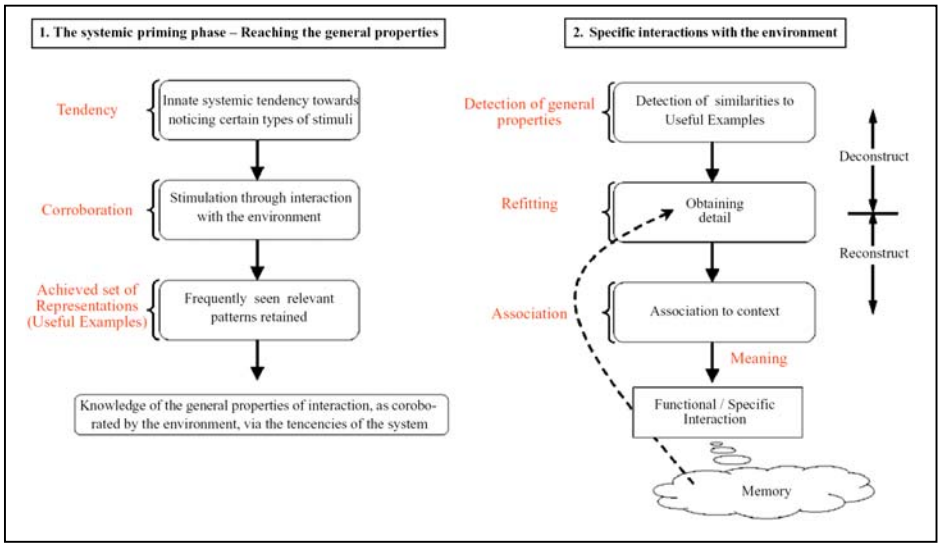


***Figure 1. The two phases encompassing cognitive action/perception:*** *Through specific interactions, and based on the previous (innate)* ***tendencies*** *of the system, certain general properties are* ***corroborated*** *by Useful Examples of these properties appearing in the environment. This* ***corroboration*** *leads to the formation of an* ***achieved set of representations*** *in the system. Specific interactions with the environment start with the* ***deconstruction*** *of the specific patterns encountered via the* ***detection*** *of similarities to the Useful Examples embodied in the* ***achieved set.*** *After* ***detection*** *there is a* ***refitting*** *and fine-tuning of the perception to the specific elements of the event. Having identified the specific elements of the event there is now a reconstruction of the raw event, accompanied by* ***association*** *to contextual factors into a* ***functional*** *or* ***meaningful*** *event, which can be appropriately reacted to and added to the* ***memory*** *of the system. The two phases are chronologically mingled. Every interaction is a specific interaction even while the cognitive system is still building its understanding of the general properties. In addition in many cognitive systems it is not clear if the process of defining new general properties ever comes to a complete stop. 'Young' cognitive systems are less fluent in working with the general properties and it is harder to teach an 'Old' cognitive system new tricks, but all interactions with the environment have elements of both phases. (Adapted from Hershebrg and Efroni 2001).*

## 2. Paradigms of immune specificity.

To show how the theory of Useful Examples changes our view of the immune systems sensitivities we must first go over the existing view. The immune system starts with a large random collection of receptors from which only a subset survives. The receptors of the immune system are selected according to a certain level of affinity to certain molecular shapes from the body, called antigens. The receptors are selected through a process of negative and positive selection in which all receptors of too high or too low affinities to these antigens are killed (along with the cells that produce them).

The resulting immune system is unique among the systems in our body in having an amazingly diverse variety of receptors. This great variance is used to identify a near infinite number of different antigens. The high sensitivity of different antigen receptors is however highly regimented allowing the immune system to react to one type of antigen in a given context while ignoring other antigens. It is as yet an open question how the immune system maintains the specificity of reaction. - The possible answers lie in two camps: Receptor specificity and systemic response specificity. By receptor specificity we mean that antigen receptors themselves are each highly focused and specific. Thus the specificities of an immune response is based on a repertoire built only of certain kinds of receptors. By systemic specificity we mean a mechanism that does not rely on the specificity of the receptors, but rather on some global behaviour of the system.

The mainstream paradigms of immunology answer this question by various mechanisms of receptor specificity (Langman et al 2000) based to some extent on Jerne's clonal selection theory: The immune systems function is to defeat pathogens. The immune system identifies foreign antigens and destroys them. The identification of the foreign is made possible by removing, in the immune systems prenatal development, all receptors that recognize self. Anything that an immune receptor identifies "must be the enemy" (Burrnet 1957).

The fact that we can find receptors sensitive to practically any substance and the importance of negative selection in the building of the immune system's naïve repertoire have supported this paradigm based on receptor specificity. However, a closer look at the way immune cells are selected in the bone marrow and the thymus indicates that this interpretation is problematic:

– Immune cells are not completely specific; they are degenerate and cross-reactive.

– A certain level of benign affinity to self-antigens exists in all receptors.

These points hold true for all cells of the adaptive immune system, and weaken the claims that the immune system's specificity is based on any type of receptor specificity. The adaptive immune system is usually divided into two main groups of

cells, humoral immunity (B cells) and cellular immunity (T cells). We present a model of T cells, focusing on the theoretical problems relevant to T cells.

All T cells are highly cross-reactive, each T cell can react to 3*10^7 different antigen types (Borras et al 2002). Furthermore, although T cells of a high affinity to self antigens are culled by negative selection, all T cells must have some level of affinity to self or they will not pass the positive selection phase and die of neglect (Goldrath and Bevan 1999). Positive selection appears to be especially important as it is now found to be an essential factor in the maintenance of the T-cell repertoire in the periphery. Low level affinity to self is not only essential as a maintaining signal of T cells but is also prolifirative in its effect (Ernst et al. 1999). We thus find a situation where the actual T cell repertoire is not, as is commonly suggested by immune theory, built of specific receptors reactive to foreign antigens. The repertoire is, in point of fact, built of highly cross-reactive receptors that all share some minimal common affinity to self antigens.

## 3. Looking at the order of the immune environment.

The answer rests in replacing the mechanistic view of the immune system with a view of the immune system that considers it to be comparable to other cognitive perceptual systems (Varela 1994, Cohen 2001, Hershberg and Efroni 2001). This view leads us, as we suggest from the study of other cognitive systems (Hershberg and Ninio 2003), to study the distribution of examples that are naturally encountered by the immune system. When we do this we see that not all antigens are equally encountered. Much as has been seen in language (Zipf 1935) the distribution is such that a few antigens are highly frequent while most potential antigen types are rare or non-existent. In fact 200 antigens, out of $10^{14}$ potential types, will be presented in 50% of T cell encounters (Barton and Rudensky 1999). This over expression in itself is not enough for us to apply the theory of Useful Examples as it does not suggest any generic trait of immune interaction. However if we add that most, if not all, high frequency antigens are self antigens the picture changes. We would like to suggest that the above inhomogeniety of presentation reflects the fact that certain families of proteins are ubiquitously expressed in all cells while others are only sometimes expressed. The proteins that are most expressed are those which are essential for cellular life, such as chaperones and other housekeeping proteins. Such housekeeping proteins have been suggested as important to immune function and they are expressed in all cells and at high levels (Cohen 2000). Because they are so important these proteins are also the most highly conserved (Gupta 1998).

We share a common heritage with the pathogens that invade us. The points of similarity between invader and host represent the essential factors of cellular life – the manipulation of genetic material and energy production. Those molecules that we share with our pathogens are essential to them and to us. It should therefore be no surprise that antigens derived from such proteins make effective immune signals.

They are older than the immune system, can be counted on to exist in all immune environments and cell types, and are over-expressed in times of stress. In fact, it has been found that antigens derived from certain foreign housekeeping proteins are very immunogenic, and induce better activation of the immune system the more they are similar to self proteins (Cohen & Young, 1991; Schwartz & Cohen, 2000). Considering the wealth of information regarding points of similarity between the self and its pathogen invaders, building an immune system that ignores the self is like constructing an eye without a fovea. In the immune system the general properties of interaction are based on the co-evolution of cellular life. The Useful Examples of such an environment are those proteins essential to cellular viability. Such proteins, due to their importance to all cells, are ubiquitous and highly conserved throughout evolution. Their derivatives are therefore ideal as base signals of cellular stress and change.

## 4. A model of primary immune reaction based on affinity to self.

We have presented the following model to show how a repertoire of cross reactive T cell receptors with a common affinity to self antigens could function effectively against a foreign pathogen (Hershberg et al 2003). In this model, we focus on the relationship between similar antigens (be they foreign or self) and the T cells that react to them. To do so, we ignore the physical space of immune interaction and look at the multi dimensional shape space of T cells and antigens (for illustration figure 2 shows a two dimensional representation of shape space). Every point in shape space represents the various similar antigen shapes that a single clone of T-cell is reactive to with high affinity (figure 2 (iii)); high affinity is that affinity which leads to negative selection in the thymus (A). First neighbours in shape space are those antigens and T cells that have enough of an affinity to be affected by the positive selection of protein antigens (A').[1]

We present a generic scenario that models the immune reaction in terms of the following observed mechanisms (Goldrath et al. 1999):

– A steady feeding of T cells from the thymus to the periphery regardless of immune activity (figure 2 (ii)). The T cells will only arrive into points in shape space that are close enough to the high frequency self antigens ((A)) so that they where positively selected in the thymus.

– T cells proliferate at a rate relative to the level of antigen they encounter and to their level of affinity to each other (figure 2 (iii)), regardless of the source of the antigen (self or non self).

[1] This supposes the existence of a relatively permanent group of self- antigens, similar in the thymus and the periphery.

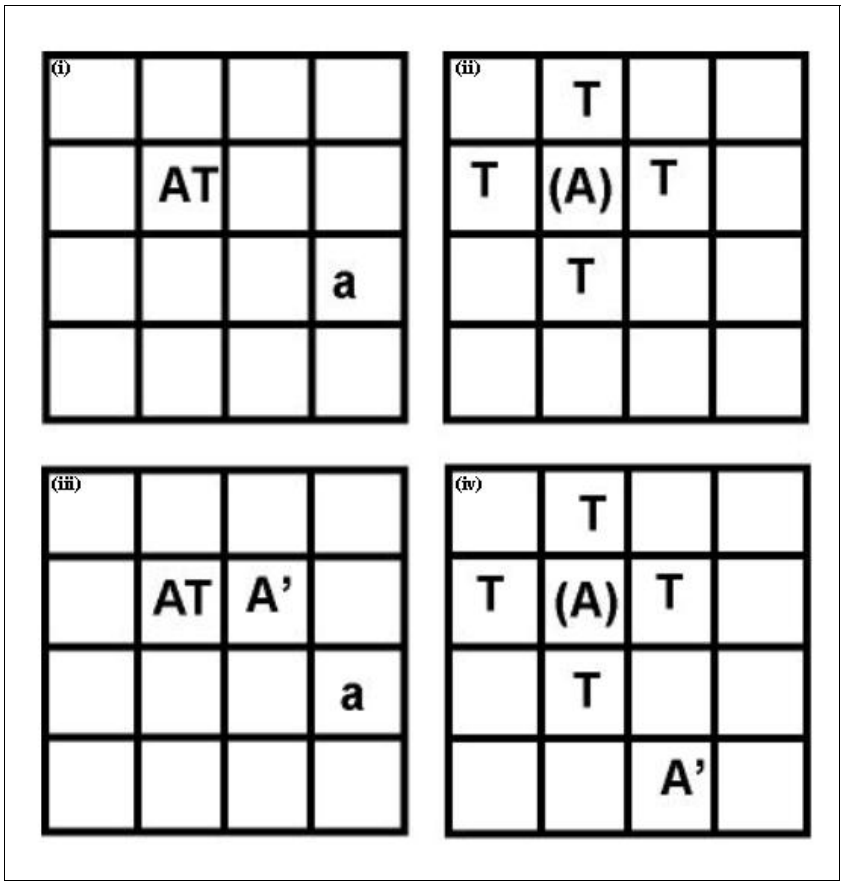


***Figure 2. Reactions in shape space*** *– shape space is an imaginary space in which every point represents the shapes of receptor/ligands that are reactive to each other As represent antigens and Ts represent T cells . (i) If a T cell and an antigen are at the same point (AT) they have high affinity to each other. If they are far away (T a) they have a low affinity to each other.*

– To represent the homeostatic competition between T cell clone types the T cells die at a rate proportional to the total population of T cells.

To these three reactions of T cell repertoire maintenance we add an effector mechanism to kill pathogens. Based on existing ideas suggesting that the effector mechanism may be activated not merely by antigen activation but rather by changes in T cell activity (Grossman et al 2001), we have built in the model an effector mechanism which is activated against antigens that change the level of T cell activity. In this model only positive changes in activity are considered.

## 5. Results and conclusions for immunity.

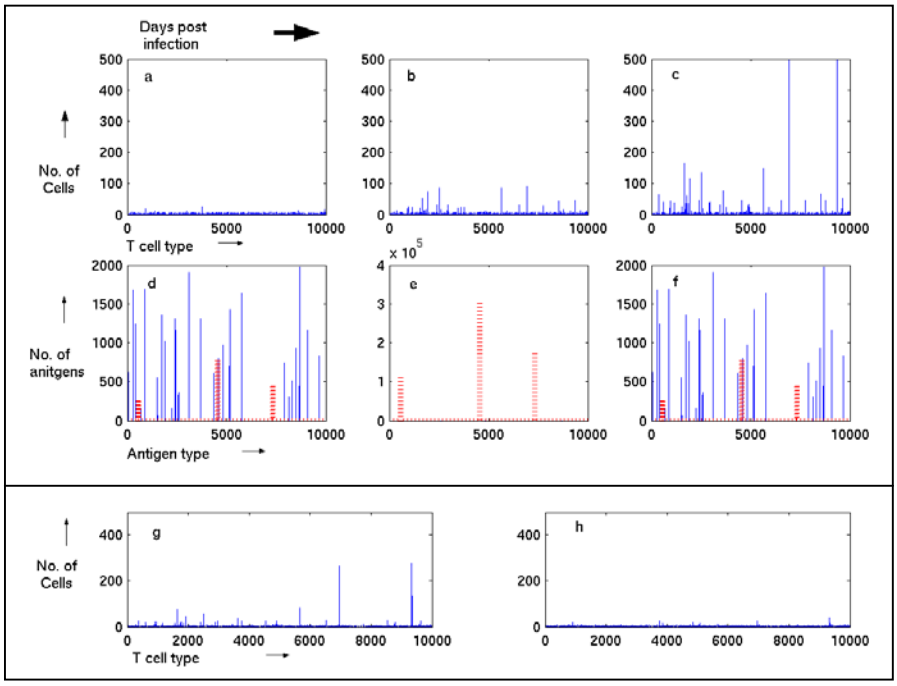


***Figure 3. Results of the simulation:*** *Upper panel: Snapshots at progressive days post infection until the height of T cell reaction at approximately eight days post infection. Changes in T cell repertoire activity [a-c] Changes in the antigen repertoire [d-f] of self antigens (blue) and non-self antigens (red). Lower panel: T cell repertoire three days [g] and three weeks [h] after the foreign antigens are cleared.*

Built around a permanent set of antigens, representing the high frequency (self) antigens, the above reactions result in a steady homeostatic population of T cells (fig. 3 [a]), without reactivity of the effector mechanism, as long as no pathogen is evident. We therefore add a primary infection in the form of new antigen populations in shape space (fig. 2 (iv) - A'). We place them so that no T cell will have a higher reactivity to them than to the self antigens that maintain them. Therefore there is no possibility of receptor specificity.

The foreign antigens differ in two traits. They are in a new place in shape space causing an imbalance in the T cell competition over antigens; and they have different dynamics from self antigens – they are proliferating rapidly.

Figure 3 shows the effect of a primary infection on the T cell repertoire. In these "snapshots"' we see the change of the T cell repertoire [a-c] until the clearance of foreign antigen, and the relative types of antigens [d-f] of self (blue) and non self (red) that push this change. At first when the level of the foreign antigens is low [d] the repertoire is at its stable resting state [a]. The rapid proliferation of pathogens [e] causes the immune system to turn its attention and resources towards the change in the pattern of antigens in shape space [b]. This change in T cell activity activates the effector mechanism, which rapidly decimates the pathogens and the antigens they exhibit [f]. Much as in reality the highest point of T cell activation is shortly after the height of the disease [c]. Those T cell types that are most highly expressed at this stage become the memory cells that allow more rapid reaction to secondary infections. The entire time course of the simulation, from infection to the height of immune reaction takes eight days[2], much as is seen in actual infections. Once the foreign pathogens are cleared the two final snapshots show the return of the repertoire to its resting state. Within a few days without the drive of foreign antigens the overall level of T cells is close to normal, however some T cell types are still over expressed [g]. Approximately three weeks later [h] the repertoire is once more at its resting state shaped around the self antigens of the body.

Thus without ignoring T cell cross-reactivity we have built a model that uses a small set of self antigens to educate the T cell repertoire. In doing so the simulation successfully recaptures the dynamics of a primary infection without the need for T cell specificity. We have shown a way that the immune system could react to foreign pathogens without the need for a special affinity to non self antigens. This success strengthens the plausibility of the paradigm of cognitive immunity and of systemic specificity in immune reactions. Most importantly it causes a complete revaluation of the place of self immunity in immune function. Self immunity is no longer viewed as the bane of an over reactive immune system, instead it is seen to be the basis of immune specificity. The ability to detect novel antigens and maintain a stable repertoire is based on a background signal built of the affinity to self.

The re-evaluation of the place of self immunity does not merely change our conceptual view of immune dynamics - it calls for an actual change in the way immunology is studied. Immunology is highly motivated by our wish to combat disease and by our perception of disease. This has lead to a unique scientific situation. Usually in science the system studied is compared at rest to its perturbed states. In immunology the view that when not activated by foreign antigens the immune system is inactive and in dynamical sense non-existent, has lead to a skewed scientific research situation. Most, if not all, immune research is done on the perturbed (disease state) immune system. However, with this new view of self affinity as a calibrating background signal, we see that the immune system is constantly active and that to understand immune disease dynamics we must first study the immune system in healthy states.

[2] A day in the simulation is calculated according to the steady rate of T cell arrival form the Thymus which is known to be ~2% a day of the resting population.

## 6. Useful Examples and study of other cognitive systems.

In the last sections I have shown how considering the natural distribution of examples encountered by the immune system has suggested a model of immune interaction that radically changed the basic paradigm of immune study. In closing I would like to stress that this is not unique to the immune system. The immune system is brought here as a primitive example of cognitive systems in general. The example is not at the level of the building blocks. There is no claim here that immune cells and neurons are the same. The similarities rest in the common strategies of interaction and in the need to create a fit between system and environment though their natural interaction. This actually strengthens the theories relevance for the creation of artificial systems. It tells us something about the general design principles of a cognitive system regardless of its specific building blocks or function.

If we wish to understand or to recreate cognitive systems we must look at the initial developmental stages that turn a naïve system into a fully functional cognitive one. The outward representation of this process is the skewed distribution of examples in the environment. The networks of analogies are not evenly distributed. Some nodes are far more central and, once learned, give the system a much easier access to the rest of the network. We have called these nodes Useful Examples; by studying them we have an observable handle to study cognitive systems. Such a handle is essential, as cognitive systems are so intuitively known to us as to make it very hard for us to study them objectively. This paper has shown, using the immunes system as an example, that our intuitive view of the function of a cognitive system is not a good place to start studying them. We should instead, when attempting to study or emulate a cognitive system, focus on the actual patterns of interaction with the environment and the Useful Examples that embody them.

This form of research has been shown to be relevant to the development of several cognitive systems - language (Ninio 1999, Hershberg &Ninio 2003), and as seen above immunity. Due to its generality we anticipate that it should hold true also for the study of other cognitive perceptual modalities.

Acknowledgements

I would like to thank Prof. Anat Ninio, Prof. Sorin Solomon and Prof. Irun Cohen for their guidance; Pierre Bohn, Yoram Louzoun, Eldad Bettelheim and Sol Efroni for arguing with me, and Shira Ninio for helping me make sure it was convincing.

Uri Hershberg is a recipient of The Levi Eshkol fellowship given by the Israeli MOS.

## 7. References


Barton, G. M. & Rudensky, A. Y. Evaluating peptide repertoires within the context of thymocyte development. *Seminars in Immunology*, 11, (1999) pp.417-422.

Borras E., Martin R., Judkowski V., Shukaliak J., Zhao Y., Rubio-Godoy V., Valmori D., Wilson D., Simon R., Houghten R. and Pinilla C. Findings on T cell specificity revealed by synthetic combinatorial libraries, *Journal of Immunological method*s 267 (2002) pp.79-97.

Burrnet FM, A Modification of Jerne's Theory of antibody Production Using the Concept of Clonal Selection. *Aust J. Sci. 20:67-69.* (1957).

Cohen, I. R. *Tending Adam's garden*. (2000) San Diego, CA: Academic Press.

Cohen, I. R. & Young, D. B Autoimmunity, microbial immunity and the immunological homunculus. *Immunology Today*, 12 . (1991) pp. 105-110.

Ernst B., Dong-Sup Lee, Chang, J.M., Sprent J., and Surh C.D. The Peptide Ligands Mediating Positive Selection in the Thymus Control T Cell Survival and Homeostatic Proliferation in the Periphery *Immunity*, Vol. 11 (1999) pp.173–181.

Goldrath, A. W. & Bevan, M. J. Selecting and maintaining a diverse T cell repertoire. *Nature*, 402, (1999) pp. 255-262.

Grossman Z. and Paul W.E. Autoreactivity, dynamic tuning and selectivity *Current Opinion in Immunology*, 13 (2001) pp. 687–698

Gupta, Redhay S. Protein Phylogenies and Signature Sequences: A Reappraisal of Evolutionary Relationships among Archaebacteria, Eubacteria, and Eukaryotes*. Microbiology and Molecular Biology Reviews*, 62, (1998) pp. 1435-1491.

Hershberg. U. & Efroni, S. The cognitive immune system: reaching the general properties. *Complexity,* 6 (2001) pp.14-21.

Hershberg U. & Ninio A. Optimal Exemplar Learning *Cognitive Systems* Special issue on multidisciplinary approaches to leaning in cognitive systems( to appear 2003)

Hershberg U., Solomon S. and Cohen, IR. What is the basis of the immune system's specificity? , special issue of *MIRIAM* devoted to papers from the 5th international conference of the European society for Mathematical and theoretical Biology (to appear 2003).

Langman RE and Cohn M, A minimal model for the self-nonself discrimination: a return to the basics *Seminars in Immunology* Vol. 12, No.3 (June 2000) pp. 189-195.

Ninio, A. Pathbreaking verbs in syntactic development and the question of prototypical transitivity. *Journal of Child Language*, 26, (1999) pp. 619-53.

Schwartz, M. & Cohen, I. R. Autoimmunity can benefit self- maintenance. *Immunology Today*, 21, (2000) pp. 265-268.

Varela, F. A cognitive view of the immune system*. World Futures, 42,*(1994) 31-40.

Zipf, G. K. *Psycho-biology of languages*. Cambridge, MA: MIT Press (1935/1965 first published by Houghton Mifflin).