Proposing a new focus for the study of natural and artificial cognitive systems

Uri Hershberg *
* Interdisciplinary Center for Neuronal Computation
Hebrew University
Jerusalem, Israel
uriher@cc.huji.ac.il

Abstract

In the study of systems the function of the system is often a good hint to how it works. In the following paper I would like to suggest that in studying or modeling a cognitive system our pre - knowledge of their functions should be treated carefully. We should focus on the statistical distribution of the system's environment and the ways this distribution affects the behavior and development of the cognitive system. I will show an example of how such a focus changes the view of the immune system. I would also like to show how this new outlook on the study of cognitive systems could affect attempts at creating artificial cognitive system.

1. Introduction

Epigenetic robotics is an attempt to model cognitive development with artificial systems. At its base is the belief that such modeling, whether to better understand cognitive systems or to build (if we can) artificial cognitive systems, requires an account of the embodied cognitive system interacting within its environment. Cognitive systems are by their nature adaptive in their formation and in their behavior in general. Together with Efroni I have suggested that this is, in fact, a criterion differentiating cognitive and non-cognitive perceptual systems(Hershberg and Efroni, 2001). In a cognitive system the capabilities of the system are not preordained merely by the plan of the system but need interaction with their environment to define them exactly. Cognitive systems are further differentiated from other adaptive systems by a common strategy of interaction with their environment which is a major factor in the form of their final capabilities (fig. 1). We have shown, citing examples from vision and language, that this criterion holds true for agreed-upon cognitive systems. We have further shown that the immune system adheres to this criterion (Hershberg and Efroni, 2001), supporting the claim that the immune system should be treated as a cognitive system (Cohen, 2000). We differ from Edelman's claims (Edelman, 1992) about the immune system, and about neuronaly based cognitive systems because we suggest that not only do they both depend on similar forces of selection to reach their capabilities but also that their overall relationship with their environment is the same. This is especially important as it means that our criterion applies to systems regardless of their being based on a neuronal substrate.

2. Optimal Exemplar Learning

The criterion cited above implies that every cognitive system must learn its eventual capabilities and sensitivities. I have, together with Ninio, suggested that this learning is a form of exemplar based learning which we called optimal exemplar learning (Hershberg and Ninio, 2002). Utilizing criteria used by Ninio's empirical study of the acquisition of syntax in young children (Ninio, 2001, Ninio, 1999) we used the example of language and the immune system to propose how this form of learning enables cognitive systems to reach the 'rules' or general properties of their environment.

These common learning strategies do not reflect a similarity between the specific building blocks of each system. They reflect a common relationship between each cognitive system and its environment. It is very clear that both systems are very different both in their biological substrate, neurons and immune cells, and in their environments. However both share a common strategy in the overall manner by which they interact with their environment, acquiring their capabilities.

In optimal exemplar learning, as in other forms of exemplar based learning like lazy learning and analogical modeling (Daelemans et al., 1997), all interactions with the environment are of the same type (fig. 1); and the system's learning is a result of unsupervised interactions with concrete examples of the environment. Only optimal exemplar learning em-

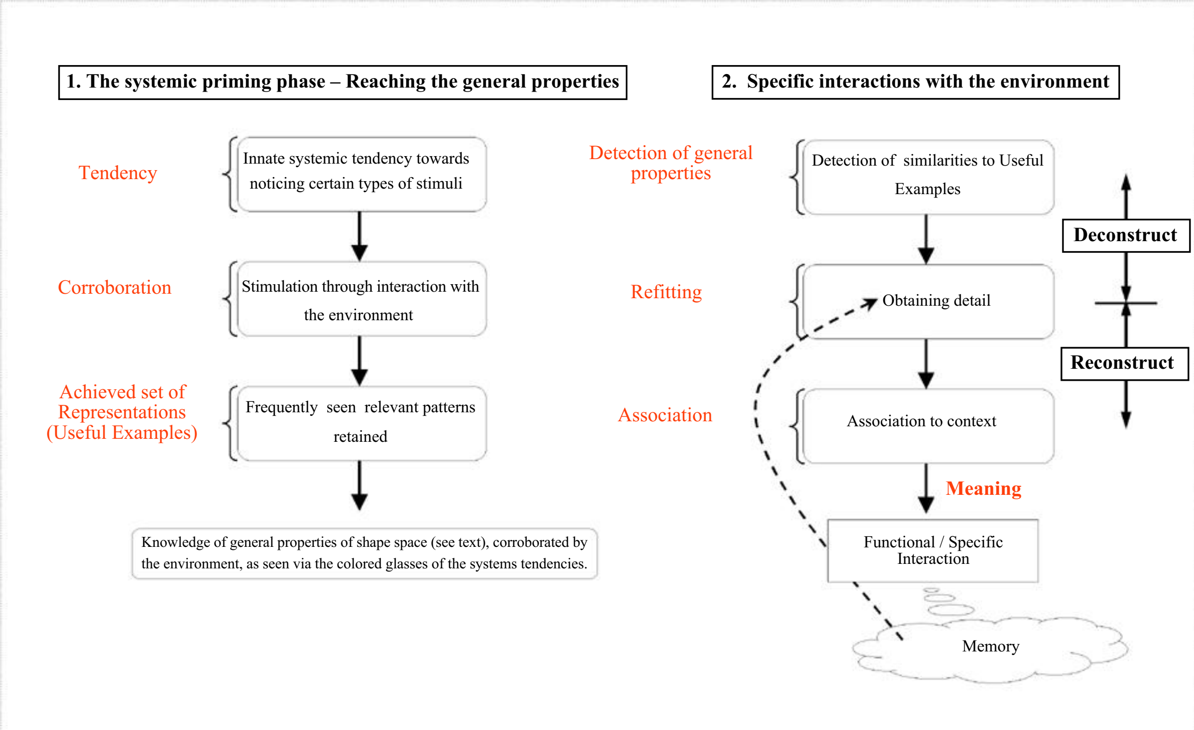

Figure 1: **The two phases encompassing cognitive action/perception:** Through specific interactions, and based on the previous **tendencies** of the system, certain general properties are **corroborated** by Useful Examples of these properties appearing in the environment. This **corroboration** leads 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 (Hershberg and Efroni, 2001)).

phasizes the importance to learning of the statistical distribution of examples encountered by the system, as it interacts with its environment. Despite its simplicity optimal exemplar learning is sufficient to allow the cognitive system an understanding of the environment because the environment is ordered. There are a few examples which are encountered at high frequency, while most examples are encountered less frequently. The high frequency examples are not only ubiquitous; they also reflect the core of some general property of the interaction of system and environment. We call these optimal exemplars of the environment Useful Examples because these are the examples most useful to the system when learning of the environment and its relevant general properties. This is reminiscent of Gibson's affordances (Gibson, 1966) with an added emphasis that these examples are especially pertinent to the nave, uneducated, cognitive system.

The distribution of examples in the environment is not arrived at by chance. For each cognitive system the reason for the singling out of specific Useful Examples is different. It depends on the specific environment and the specific types of interactions which the cognitive system undergoes with its environment. This reflects the fact that cognitive systems, as autopoetic entities (Maturana and Varela, 1980), are fitted by evolution to specific niches. Based on genetic inheritance and previous development, a cognitive system has certain tendencies, which give it the framework of its environment. Saying a cognitive system has certain tendencies is not the same as saying it is made for a certain function, it only means it has developed in a certain kind of environment. Its function depends to a great extent on the environment it is yet to encounter.

This view shows that if we wish to study the causal factors in the development of a cognitive system we should be careful in factoring-in their eventual function. We should possibly attribute a greater importance to the statistical structure of the natural environment, and in particular, to the identity of its most frequent exemplars - the Useful Examples. Without taking these interactions into consideration we may have misconceptions as to the full range of functions of the system.

3. The Immune System

The immune system's ability to identify and interact with a wide array of specific events and changes in the body, appears to imply that the immune system reacts to patterns and contexts. This has lent support to the treatment of the immune system as a cognitive perceptual system.(For further clarification of the immune system as a cognitive system see (Hershberg and Efroni, 2001),(Cohen, 2000)).

The immune system is a primitive cognitive system at best and is therefore not a very good example of the deeper cognitive traits. It is however a good example of the possible dividends to the study of cognitive systems that can come from focusing on the statistical distribution of interactions with examples of the environment. The function of the immune system is clear to most people, practitioners and laymen alike. Its function is to fight disease; i.e. to combat the invasion of foreign microbes, viruses and parasites into our body. I will try and show how this preconception of the immune system's function causes misconceptions in the common view of the immune system.

The immune system is a dispersed system that has many types of cells as well as effector and signaling substances. The two most important groups of cells are called B cells and T cells. Both of these cell families have a unique ability to create a repertoire of receptors of varied shapes. The shape of the receptor reflects the shape and type of molecule that will activate the receptor. This gives the immune system the potential ability to have receptors that can identify an astronomical number of molecular shapes, called antigens.

The receptors of B cells identify extra-cellular substances. The receptors of T cells identify intracellular substances by interacting with specialized antigen-presenting proteins known as Major Histocompatability Complex (MHC) receptors, which are expressed on the surface of everyone of the body's cells. MHCs present fragments of intracellular proteins, in effect mirroring the internal state of the cell. The immune system's identification of and reaction to a pathogen or other immune events is dependent on mutual reaction by both T cells and B cells to that event (Cohen, 2000).

The common view is that this plethora of receptor sensitivities is used only for the identification of the varied foreign antigens. If this is the function, why do we need self- reaction? Why do we have auto-immune diseases in which the immune system identifies and attacks tissues of our body? This problem gives rise to the need to give an account of how the system curbs self-recognition either by the whole system and/or by its receptors. At its most severe, this viewpoint, as put forth by the clonal selection paradigm, states that all receptors react only to foreign pathogens. This selectivity is reached by the destruction of any immune receptor that reacts to self-antigens (Burrnet, 1957). Less severe paradigms postulate various regulatory mechanisms that suppress or negate self-reaction. All such theories of the distinction between self and non-self in the immune system (Langman and Cohn, 2000) share a few problems. First, empirically benign self-reactivity exists in the body (Nemazee, 2000, Lacroix-Desmazes et al., 1998). Second, our invaders and our bodies are related. A common evolutionary source and eons of co- evolution mean that there are very few antigens which, without a wider context of interaction, can signal to the immune system: "I am foreign" (Cohen, 2000).

Instead of looking at the immune system's function let us look at the examples presented by the environment. The creation of the immune system's receptor repertoire is very complicated and I will confine my remarks here to cellular immunity and the creation of the T cell repertoire.

At an embryonic stage of development a population of T cells is created with a random repertoire of receptors. These cells then migrate to the thymus where they are culled to create our adult repertoire. The exact makeup of the protein exemplars used in the thymus is not yet known. However the fact that the form of the exemplars affects the final repertoire (Goldrath and Bevan, 1999) implies that interaction with the antigen examples forms the eventual sensitivities of the system. In the thymus the T cells interact with certain specific antigen examples presented in MHCs. The T cells must pass two tests to survive. First, those cells that do not react at all to the MHCs or to the protein examples they hold, die of neglect. After this stage those cells that react too strongly to the same examples are also killed (Goldrath and Bevan, 1999). This creates a repertoire of receptors with variable receptiveness to many types of self antigens and a common level of affinity to these examples. In essence in the thymus we have a precisely delineated form of the priming phase of a cognitive system (fig. 1). The **tendency** of the immune system, as expressed in the shape of MHCs, is **corroborated** by the protein examples presented by them, creating a receptor repertoire which reflects the **achieved set of representations** of our molecular self.

Let us look at the statistical distribution of the antigen examples the body uses to create its receptor repertoire and see if they fit in with the theory of Optimal Exemplar learning. The statistical distribution of examples is similar to the way words are distributed in language. Much as Zipf described the distribution of words (Zipf, 1935) the peptide MHC complexes found on the cells of our body have a distribution in which a few peptides are very frequent (200 out of 10^4 occupy 50 percent of the MHCs of a given cell) and there is a large group of low frequency peptides (Barton and Rudensky, 1999).

Exactly which peptides belong to the high frequency group is hard to check experimentally; however there are other hints as to who are good candidates for this role. Candidates for Useful Examples would have to have the following properties: they would have to be expressed by every cell; they would have to have remained relatively unchanged all through the history of the development of the immune system, for them to be such a major part of its function; and they should be relevant in times of stress. Indeed, it is possible to find such a set. The set corresponds to a group of antigens Cohen has named "homuncular antigens" (Cohen, 2000) which belong to a group called housekeeping or maintenance proteins. Housekeeping proteins are essential in all cells, since they are responsible for ongoing energy metabolism, protein construction and basic genetic manipulations. In situations of emergency and stress, which are also the ones where you may usually find immune response, these proteins, or some subset of theirs, tend to be highly expressed.

That alone suffices to mark them as Useful Examples to the immune system. But more than that: Since they include proteins whose function is essential to cellular viability, they are expressed in all cells. In fact, these proteins are ubiquitous in cells in times of stress and are highly preserved throughout evolution, from prokaryotes to multi- cellular organisms (Gupta, 1998).

Beyond these theoretical considerations, receptors reacting to such proteins do in fact have a special importance in immune reaction: Receptors to different types of endogenous and exogenous antigens of housekeeping proteins have been shown to be important to immune reaction to pathogens and also in benign self immune reactions for bodily maintenance such as fighting cancer and repairing trauma (Cohen and Young, 1991).

This raises a general objection to considering all immune functions to be relevant to foreign objects only. For functions of bodily maintenance an ability to be receptive to signals of changes in self seems quite beneficial. Furthermore, as I mentioned above, essential proteins are highly conserved between foreign and local cellular life, and so a certain reactivity to self also seems called for to allow greater reactivity to pathogens and recognition of their essential building blocks.

In the entire description above, the known functions of the immune system and especially those of anti-pathogen behaviors and inflammation play only a minor role. The only types of interactions mentioned here are basically related to the levels of different proteins in the body, proteins which behave as examples of the future inputs of the immune environment. These eventually build the immune system's receptor sensitivity enabling it to oversee the state of our bodies. This leads us to the point where we can no longer consider the immune system merely as a mechanism for fighting foreign invaders. This new outlook calls for a more definite appraisal of the actual functions of the immune system as the housekeeping agent of the body. By going beyond known functions and asking what are the examples from which the system learns, we can arrive at to more of the potential functions, which previously we ignored.

The conception of the immune system as housekeeper rather than gatekeeper leads to the obvious conclusion that in order to understand the immune system we must study it not only in the context of combating pathogens but also in health, much like we do other biological systems. It is also a good example of how in the study of cognitive systems, our preconceptions of the functions of the system should be considered carefully. The function of a system is important to our understanding of the system but to properly understand cognitive systems and their functions, we should study them through the statistical shape of the environment, the interactions between the system and the environment, and the Useful Examples that embody them.

Artificial Cognitive Systems

We come now to the question most relevant to the subject of this conference, namely, how the concept of Useful Examples, which I have shown to be important in the study of cognitive systems, pertains to the design of artificial systems. I would like to suggest a new formulation of the basic design question. Rather than asking how to create an artificial system that functions as a cognitive system, we should ask how to define Useful Examples in the environment so that through interaction with them the system creates itself.

By asking this we return to the basic tenants of Epigenetic Robotics. An artificial system attempting to emulate cognitive systems, as an aid in scientific study or as a technological tool, can not be an AI like system built to optimize some function. A computerized face identification algorithm is not cognitive, no matter how innovatively built or how well it scores on its tasks. It must capture the essence of the relationship between an autopoetic machine and its reciprocal environment. By this I do not mean that it must necessarily be 'organismic' (Ziemke, 2001), born and made of organic material. Autopoesis is defined as a form of organization(Maturana and Varela, 1980). The common factors of autopoesis are in the relationship between the machine, its development and its environment, not the materials it is made of.

This new design outlook is especially pertinent in complex environments where no natural cognitive system exists to be emulated, for instance the information space of the internet. Following the theory of Useful Examples, we would no longer be aiming to create agents which function intelligently in the internet. Rather, we should ask how we can divide the internet into niches in which different agents, by their interaction with these niches, will find Useful Examples. This rephrasing of the question is by no means a solution, which is unfortunately still beyond the scope of this paper. However, it does capture more of the essential aspects of cognitive systems than the usual question of how to create an artificial system functioning like a cognitive system. The new question is an improvement since it acknowledges the interconnectedness of the system, its development and its environment.

References

Barton, G. and Rudensky, A. (1999). Evaluating peptide repertoires within the context of thymocyte development. *Seminars in Immunology*, 11:417–422.

Burrnet, F. (1957). A modification of jerne's theory of antibody production using the concept of clonal selection. *Aust J. Sci.*, (20):67–69.

Cohen, I. (2000). *Tending Adam's Garden.* San Diego, Calif., Academic Press.

Cohen, I. and Young, D. (1991). Autoimmunity, microbial immunity and the immunological homunculus. *Immunol Today*, 12(4):105–110.

Daelemans, W., Gillis, S., and Duieux, G. (1997). Skousen's analogical modeling algorithem: A comparison with lazy learning. *New Methods in Language Processing*, pages 3–15. UCP.

Edelman, G. (1992). *Bright air, brilliant fire.* Basic Books.

Gibson, J. (1966). *The senses considered as perceptual systems.* Boston: Houghton.

Goldrath, A. and Bevan, M. (1999). Selecting and maintaining a diverse T cell repertoire. *Nature*, 402:255–262.

Gupta, R. S. (1998). Protein phylogenies and signature sequences: A reappraisal of evolutionary relationships among Archaebacteria, Eubacteria, and Eukaryotes. *Microbiology and Molecular Biology Reviews*, 62(4):1435–1491.

Hershberg, U. and Efroni, S. (2001). The immune system and other cognitive systems. *Complexity*, 6(5):14–21.

Hershberg, U. and Ninio, A. (2002). Optimal exemplar learning in cognitive systems. *Cognitive Systems.* (To be published in a special issue on multidisciplinary approaches to leaning and cognitive systems).

Lacroix-Desmazes, S., Kaveri, S.V.and Mouthon, L., Ayouba, A., Malanchere, E., Coutinho, A., and Kazatchkine, M. (1998). Self-reactive antibodies (natural autoantibodies) in healthy individuals receptor selection in B and T lymphocyte. *Journal Of Immunological Methods*, 216(1-2):117–137.

Langman, R. and Cohn, M., (Eds.) (2000). *Seminars in Immunology.* 12(3) - (Special issue on theories of self / non-self discrimination).

Maturana, H. and Varela, F. (1980). *Autopoesis and cognition.* D. Reidel.

Nemazee, D. (2000). Receptor selection in B and T lymphocyte. *Annu.Rev. Immunol.*, (18):19–51.

Ninio, A. (1999). Model learning in syntactic development: Intransitive verbs. *Inter. J. of Bilingualism*, (3):111–131. (Invited article for a special issue on Cross-linguistic Studies of Early Grammar, edited by M.M.Vihman).

Ninio, A. (2001). Linguistic rules and categories are learned from "good models": A proposal for a new learning process. (Poster presented biennial meeting of the Society for Research in Child Development.).

Ziemke, T. (2001). Are robots embodied? Proceedings of the 1st Intl. workshop on Epigenetic Robotics:75-85.

Zipf, G. (1935). *Psycho-Biology of Languages.* Houghton-Mifflin.