Optimal exemplar learning in cognitive systems

Uri Hershberg* and Anat Ninio +

*Interdisciplinary Center for Neuronal Computation, Hebrew University, Jerusalem Israel. Email:- uriher@cc.huji.ac.il

+Department of Psychology, Hebrew University, Jerusalem Israel.

Abstract

The following paper is the fruit of an interdisciplinary discourse between the study of the learning of syntax in infants and the study of the development of immunity. Showing specific examples from these fields, we propose a form of optimal exemplar learning. We suggest that this form of learning is how all cognitive systems reach the "rules" or general properties of their environment.

In such learning, the specific interaction of the system with any example is identical. However, not all examples of the environment are equally encountered in natural interactions with the environment. There is a class of examples that are ubiquitously encountered and generic of the general properties of the environment. This ensures that they will be the first noticed and learned by the system. The very ubiquity of the useful examples also ensures their rapid reinforcement and facilitates the learning of other examples of the same property.

In talking to infants, parents use with high frequency a set of verbs that are almost empty semantically, being generic to the subcategory to which they belong. The first verbs used by the infants in learning syntax belong to the same group of verbs. Once these first examples of a given syntactic combination are learned, the ability to learn further examples is greatly facilitated.

In living cells those proteins that the cells need most in trauma that are most highly conserved throughout evolution. These proteins are ubiquitous in both humans and microbes. The immune system creates the major antibody clones precisely to these proteins when combating diseases.

In summary, we have discovered an underlying similarity between the learning processes of two a priory unrelated systems. This leads us to believe that optimal exemplar learning may be operative in all cognitive systems. This is a strong hypothesis that should now be investigated in other cognitive systems.

1. Introduction.

There are many different cognitive systems, and there exist many descriptions of their characteristics and functioning. In terms of general system theory one trait that is quite clearly common to all such system regardless of their definition, is their self organization through interaction. In all cognitive systems the final capabilities of the system are not fully determined by the plan of the system. To reach these capabilities the system must interact with its environment (Hershberg & Efroni, 2001). In the following pages we present a theory of the common traits and learning strategies of cognitive systems. Specifically we would like to suggest a form of exemplar learning, called Optimal Exemplar Learning, that enables a cognitive system to learn to be a cognitive system. These common strategies do not reflect a similarity between the specific building blocks of each system. They are a result of a common relationship between cognitive systems and their environments. To demonstrate this point we will use two very disparate systems, language and immunity. It is very clear that both systems are very different both in their biological substrate, neurons and lymphocytes, and in their environments, the abstract information space of language and the molecular jungle of our bodies.

However, as we will see in the following pages, both share a common strategy in the manner by which they interact with their diverse environments and acquire their capabilities.

2. Optimal Exemplar Learning: How does a cognitive system learn to be a cognitive system?

The mode of learning by which a cognitive system acquires its capabilities and understanding of the general properties of the environment is by interacting with concrete exemplars of the environment. Learning is the result of unsupervised interactions with exemplars; all interactions with the environment are of the same type.

We think that this mode of learning is extremely efficient in cognitive systems because the environment itself is ordered, and because examples of the general properties of the environment are ubiquitous. We call these optimal exemplars of the environment "useful examples", because they are useful for learning the environment and its relevant general properties by the relevant cognitive systems.

Learning of examples of the general properties of the system does not change the types of interactions with new examples. However, due to the centrality of Useful Examples in the environment, learning them facilitates the acquiring of, or the correct reaction to, new examples. Useful Examples make for optimal exemplar learning.

The environment is ordered in a specific way which is especially pertinent to a naive unformed system. This order is not a chance occurrence; it is a reflection of the fact that cognitive systems are fitted to certain niches. Based on genetic inheritance and previous development, a cognitive system has certain tendencies, which give it the framework of its environment. This reflects the previous environments in which this system has developed / evolved and not the final capabilities of the system, which are dependent on future interactions with the environment.

It is tempting to speculate that if there were no statistical disparities among the exemplars of the environment, making the central exemplars easiest to assimilate by the learner, learning would be much more difficult, drawn-out and error-ridden than it is in fact. This is why we call exemplar leaning based on Useful Examples of the environment, Optimal Exemplar Learning.

Having given a general description of Optimal Exemplar learning we would like to show how two cognitive systems, language and the immune system fit in with these principles. We will start our examples with language, a more accepted cognitive system, and show how it works within the framework of optimal exemplar learning.

3. Language: The role of central examples in the acquisition of syntax.

We will deal with only one aspect of the acquisition of language – the learning of correct syntactic combinations, namely, how to construct a meaningful sentence from more than one word. However, our points should hold for other stages of language acquisition as well. Learning syntax is made possible by various innate and learned biases whose exact amount of influence is the scene of many an argument in the cognitive sciences which we will not go into. In any case, it is obvious that at the stage of syntactic development the language system already has a "tendency" which sets the stage for the learning of syntactic combination. This can be seen in the fact that the learning of sentence formation is usually preceded in children by a stage of rapid growth in vocabulary or "vocabulary explosion", that brings about the creation of the vocabulary necessary for syntactic combinations (Bates, Bretherton, Snyder & Beeghly, 1988).

It is a fact of long standing that, in general, the statistical shape of language is such that a relatively small subset of words are highly frequent while the rest are used at a lower frequency (Zipf, 1965). According to hypothesis, the highly frequent words have a special role in syntactic development (Ninio 1999a, 1999b).

The studies on the basis of which this hypothesis was formed, compared the order in which children learned to form syntactic combinations with intransitive and transitive verbs, with parents' use of verbs in syntactic combinations in conversations with their children. It was observed that parents use a small subset of verbs with very high frequency when talking to their children. Words like want, come, go and make account for a high fraction of the verbs used in parental conversation. All these high frequency verbs are very general, have uses that are almost empty semantically and can be said to be generic of the verb sub-categories to which they belong, defined by their following a certain set of syntactic rules. In other words, children are presented with high frequency exemplars that have very few irrelevant features as far as their syntactic behavior is concerned. In return, all the first verbs used by children in combination are usually drawn from this group of verbs.

Once the first verbs are learned in a certain syntactic construction, the speed of learning other verbs in the same syntactic construction, but not necessarily in other constructions, is greatly enhanced. This could be indicative of a scenario where the child learns the first examples with a great deal of investment of effort, after which the ability to learn new verbs in the same syntactic combination is greatly facilitated. Figure 1 presents as an example two graphs documenting the development of verb-object and subject-verb-object word-combinations in a child, as a function of age. The dependent measure is the cumulative number of different verbs participating in each type of construction as observed in recorded interaction sessions of child and caregiver. Each verb is counted at the age when it is first produced in the relevant syntactic construction.

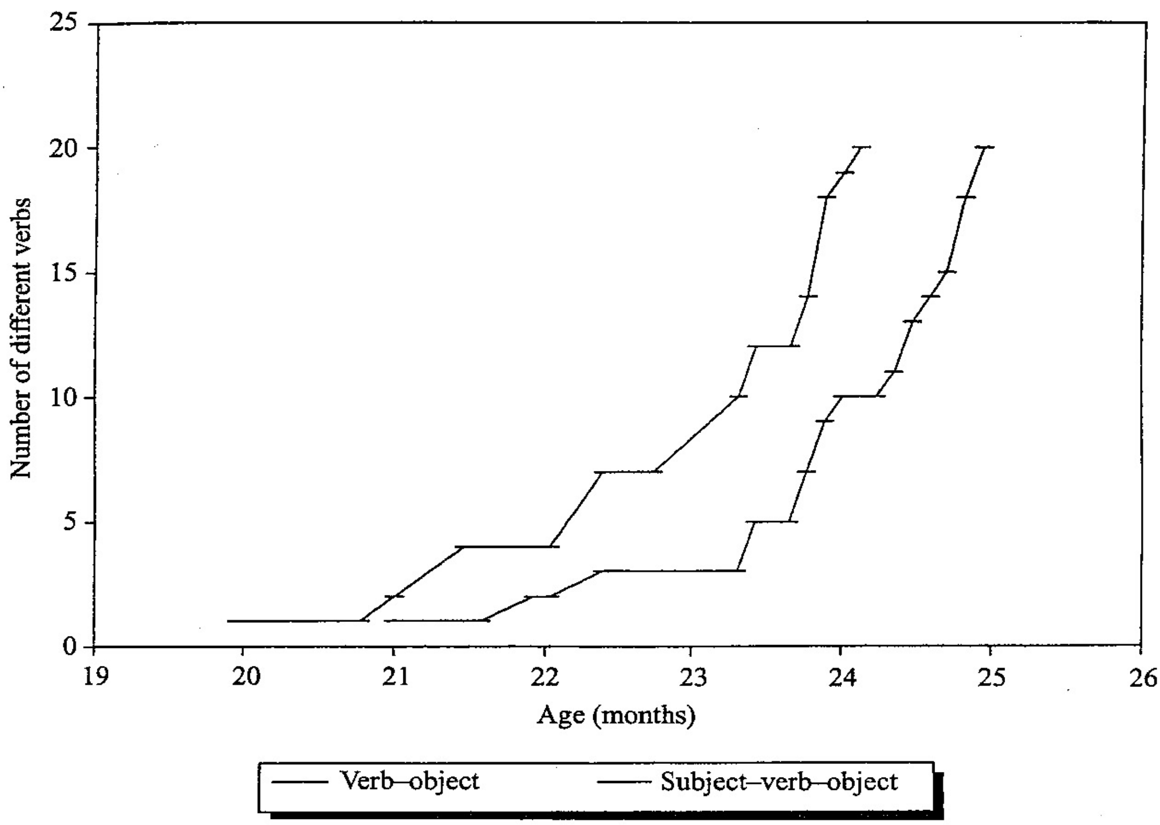

Fig. 1. Cumulative number of different verbs in VO and SVO word-combinations produced by a child acquiring English, as a function of age. Source: Ninio (1999a).

It is evident that these graphs have the characteristic shape of typical gradually accelerating learning curves: The time it takes to apply the new rule to yet another verb is much longer at the beginning of acquisition of that rule, and it gets shorter the more verbs the child has already learned to produce in the relevant pattern. Accelerated learning has also been observed with regard to other aspects of language acquisition such as the growth of vocabulary (Dromi 1987, Van Geert & van Dijk, 2002).

The conclusion is that languages are statistically shaped so that when we talk simply to those who are just learning to talk, we are accentuating the general properties of language. The statistical structure is a means of ascertaining that language can be learned through meeting with the concrete exemplars of the linguistic environment. The statistical structure of language, in other words, allows Optimal Exemplar Learning.

4. The Immune System: Learning from similarities.

The definition of the behavior of the immune system as a cognitive system is not new; it has been put forward by Cohen (2000), Varel (1994) and others (Hershberg & Efroni, 2001).

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 antigen examples from the body, through a process of negative and positive selection in which all receptors of too high or too low an affinity to these antigen examples are killed (along with the cells that produce them). These examples reflect the expression of proteins in the various cells of our bodies. As in other perceptual systems the selection of receptors is competitive. If a receptor spends a long time inactive, it will be pushed aside by more active receptors (Goldrath & Bevan, 1999). The immune system is left with a repertoire of cells that all share this level of affinity to the antigen examples of the body.

Let us look how this fits in to the principles of optimal exemplar learning. It seems bizarre to emphasize that all examples are concrete in the immune system. Compared to other, more obvious, cognitive systems, where it is apparent that abstraction and general concepts may exist as objects, in the immune system the receptors are all dealing with specific concrete molecular patterns. However, as in the other systems, the environment is ordered, and to reflect this order the receptor repertoire is better served by learning form certain examples and not others. All accumulation of knowledge from previously presented examples is by transference. Cells of the immune system react differently to an antigen or other chemical signals depending on previous experience and reaction to examples. This is true in the thymus where a previous reaction may cause cell death but also in the periphery where activation can cause differentiation of naive immune cells to active and to memory cell subsets.

To show how the immune system fits to the form of Optimal Exemplar Learning, let us start with the statistical distribution of examples. As in language the environment is ordered, the antigens, which select the cells, have a distribution which is similar in shape to that of words in language. In an Antigen Presenting Cell 50% of all antigens presented will belong to 200 of a possible 10^14 antigens (Barton & Rudensky, 1999).

In an attempt not to use too many names and other forms of biological jargon we have left out what exactly are the Useful Examples in the immune system and what are the general properties they embody. The general properties the immune system notices are based on the fact that as cells, both our selves and the pathogens that invade us are not far distant from each other. We all stem from the same evolutionary chain and share many biological mechanisms. Possible candidates as Useful Examples that reflect this fact could be housekeeping proteins (Cohen, 2000). This group of proteins have, as their name implies, essential cellular housekeeping functions that are expressed in times of cellular stress to help the cell maintain its viability in hard conditions. These are good candidates for Useful Examples of cellular status for three reasons first they are expressed in times of stress which is a good signal to raise the interest of the immune system, second they are necessary in all cells and so are expressed in all the cells of our body and outside it. Third and not least they are so important that they have changed very little over the evolution of life on earth and so are extremely similar in us and in the bacterial pathogens that invade us (11). Therefore a receptor repertoire built around such examples would have to change little to become efficient identifiers of foreign proteins and their derivatives. Beyond these theoretical considerations such housekeeping proteins and the immune receptors that are sensitive to them have been found to be important for good immune reactions to both foreign invaders, and to other immune functions, such as combating caner and the repair of damaged tissue (Cohen & Young, 1991; Moalem, Yoles, Leibowitz-Amit, Muller-Gilorr, Mor, Cohen, Schwartz, 2000 Schwartz & Cohen, 2000).

Much like in learning language, when the immune system interacts with its environment it will often encounter the Useful Examples causing the repertoire of receptors to be built around them. A repertoire built around other examples, not ubiquitous to all cells would have receptors that could be irrelevant to some pathogens, or even some cells of our bodies. As it is possible that some cells do not even express these proteins or any proteins similar to them. Even should some of the repertoire be built around other antigens these receptors will be less versatile in dealing with immune events and so over time should die out relative to the repertoire built around the Useful Examples which better reflect the common ground of cellular life and the factors essential to its viability.

5. Conclusions: Differences and unifying principles

Although we have dealt here only with two system, we believe that what we have said here may hold true for other cognitive systems as well. Cognitive systems and their environments differ in their interactions and in their building blocks. General properties and the examples that represent them have different reason for being signaled out by the perceptual tendencies of different cognitive systems. In the immune system it is the co-evolution of pathogens and hosts that defines the general properties. In language it is the fact that language is a cultural artifact, which needs to be passed on between users. 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 that the cognitive system undergoes with its environment. This reflects the fact that cognitive systems are fitted by evolution to specific niches. Therefore this difference basically represents the differences of environment and strengthens the feeling that in many disparate environments different cognitive systems are using the same rules to learn to be cognitive system.

Optimal Exemplar Learning allocates great importance to the statistical structure of the natural environment and in particular, to the identity of its most frequent exemplars. In attempts to study and simulate cognitive systems we must retain the statistical shape of the natural environments in which they develop, as this appears to be one of the essential features in their development. Even before we consider other aspects of cognitive learning implied by the principles of Optimal Exemplar Learning, for example the embodiment of examples in their natural environment; the importance of the behavioral and perceptual tendencies of the system; and so forth, we must consider the statistical distribution of the environments. This distribution by itself may direct us to the essential pieces of information that the systems use to learn their capabilities.

When studying a specific cognitive system and its environment or, for that matter, when trying to create an artificial system with the traits of a cognitive system, we should try and understand what are the Useful Examples they use to learn to be cognitive systems, as knowing what the Useful Examples are, gives a great amount of information about the cognitive system, its environment and the nature of the interactions between them.

References:

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

Bates, E., Bretherton, I., Snyder, L. & Beeghly, M. (1988). *From first words to grammar: Individual differences and dissociable mechanisms.* Cambridge University Press.

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

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

Dromi, E. (1987). *Early lexical development*. New York: Cambridge University Press.

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

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

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

Moalem, G., Yoles, E., Leibowitz-Amit, R., Muller-Gilor, S., Mor, F., Cohen,

I.R. & Schwartz, M. (2000). Autoimmune T cells retard the loss of function in injured rat optic nerves. *Journal of Neuroimmunology, 106*, 189-197.

Ninio, A. (1999a). Pathbreaking verbs in syntactic development and the question of prototypical transitivity. *Journal of Child Language, 26*, 619-653.

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

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

Van Geert, P.L.C. & van Dijk, M. (2002), Focus on variability: New tools to study intra-individual variability in developmental data. *Infant Behavior and Development,25*,340-374

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

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