Vortrag auf dem Symposium

"Evolution and Natural Selection: a Model for Understanding Societies"

Prague, March 10-11, 2000

The theory of evolution is, in a basic sense, a theory of information creation, transfer, and reproduction. The material with which these faculties are implemented was originally related to biological organisms, and the information was found to be expressed in an organism's phenotype. Information in such systems comes in packages which we call genes, and its materialistic elements, its letters, and some of its grammar are already quite well understood. The study of this biological system – as we have already heard – has unraveled some of the basic design principles of this system for information transfer, reproduction and generation: gene multiplication and transfer from one generation to the next must be highly reliable, thus reproduction of expression is a process of high fidelity, and creation results from the interplay between small variations (mutations, recombination) and selection in all its different forms. The driving force of the process appears to be an intrinsic property of the information packages – the **genes** – themselves: their selfishness as Richard Dawkins calls it, namely, their drive to penetrate into all possible implementations, particularly in the next generation.

Information comes in many forms. One, on which we want to focus in the context of our symposium, is mental information, the form of information which originates from the work done by nervous systems and brains, and which is transmitted between individuals (both animal and man) by **imitation**. Imitation is a process which depends on transfer and reproduction. It leads to traditions across generations and to similar habits and attitudes of groups of communicating organisms. In a general sense it produces a common knowledge base as a reference system in communities at various levels of complexity and size. When we ask whether evolution theory might be applicable to human societies we basically ask whether the design principles of the evolutionary process as discovered in biological systems are also applicable to the transfer, reproduction and creation of mental processes. A key issue in this context is the question whether mental information comes in packages, how they replicate and whether such packages have properties comparable to their biological correlates, the genes.

This question refers to the concept of memes (or "culturgens", as they also have been called). Richard Dawkins, who introduced this concept in his book "The selfish gene", and Susan Blackmore, who elaborated on it in her book "The meme machine" , provide intriguing arguments in favor of their view, but unfortunately their ideas have not yet been connected with neural processes, their organization and implementation. There is a line of thinking, presented mainly by Gerald Edelman on Darwinistic processes in the brain machine, which might provide some hints in this respect. I want to present just a very short outline of the direction one might have to go if one wanted to enrich the meme concept with brain theory.

Several questions come to mind:

What is a meme in the neural sense?

Are memes stable enough, do they self-replicate and do they provide copying fidelity?

How are memes selected and changed?

Is there a meme-based (natural) selection process?

Can the analogy between gene and meme do useful work?

One can think of many more such questions, but I will be not able to address even these few questions in any detail.

The neural elements of cognitive functions are functional ensembles of nerve cells. An ensemble consists of a large or small number of highly interconnected neurons forming a functional and highly adaptive network. Its structural basis is the neuroanatomical correlate of interconnected neurons, as seen, for example

(transparency 1)

in the structural modules of neurons in the cortex of mammals (here the cortex of the Cerebellum). An ensemble can have several to many discrete states, each of which codes for a particular percept, motor unit or cognitive element. This important property, namely, the multitude of functional states of one anatomically fixed circuit, has been particularly well proven in small circuits in which all connections are known, e.g. in a network generating the feeding rhythms of a particular part of the stomach in an invertebrate animal, the crab.

(transparency 2)

Formalization of such biological networks by technical neuronet models, e.g. in the form of associative matrices,

(transparency 3)

have demonstrated important properties: multiple stable states and thus multiple stored memories, a property known as completion: only a part of the original input is needed to create a concise output; the property of generalization, meaning that not only the original input but also similar other ones produce an output; and the property known as graceful degradation or fault tolerance: the neuronet functions even if some of its elements are damaged. Another way of talking about the properties of such neuronets is to describe them as attractor networks and adopt the lingo of complexity theory. In any case, the interesting point here is that these functions are also found in biological neural nets and result from the particular properties of synaptic interaction between the neurons. The most important rule at work here is that synapses adjust their strength according to the activity of the two neurons connected by the synapse. Whenever a pre- and a postsynaptic cell are simultaneously active they strengthen their synapse, when their activity is out of phase they weaken their synapse. This is known as the Hebb Rule. Gerald Edelman´s theory of “neural Darwinism” formulates the conditions (his key term is the principle of reentry) under which neural ensembles stabilize their multiple functional states and accumulate information by experience with the environment. Let us assume for the moment that functional states of ensembles are the substrate for memes. Since no one really knows what a meme is, we might well pursue this notion and see whether it makes any sense.

To put it briefly, we may thus ask: "Do ensembles/memes correspond to genes in such a way that one might fruitfully apply evolution theory?" Genes replicate, and that is an important function in the transfer of information within biological systems. Ensemble states do not need to replicate within the brain, because the material which encodes the information is not transmitted, but the information is encoded and recoded by the sender and the receiver without any material transport. Thus memes replicate by the transfer process itself, and their substrate – the ensembles – do not need to replicate.

The discrete states of ensembles may correspond to alleles of genes, and thus multiple phenes #phenotypes??# may result from multiple states of the ensemble. As in the gene-phenotype relationship, additional levels of complexity, depending on their interaction of genes (their “grammar”), come into play when ensemble activities are expressed (or lead to percepts), and these levels of complexity are not understood, neither in the gene-phenotype nor the ensemble-percept/motor act relationship.

We might expect that the extremely large number of neurons in the brain, and accordingly, the large number of ensembles, may lead to a practically unlimited number of functional states of the ensemble and thus unlimited expressions. In such a case the analogy between gene and ensemble would break down, because it is a basic requirement of evolution theory that functional states are limited and that transfer and reproduction of information packages is highly reliable. Let me show you a few examples demonstrating the high fidelity in our perceptual system (similar examples may be drawn from the motor system or from cognitive properties).

Transparencies and/or slides on illusions and deceptions

It thus appears that ensembles have discrete states which exclude each other and between which the ensemble can fluctuate under competitive input conditions.

Imitation – the analog process of transfer and reproduction – requires that ensembles be changed by experience. This is the condition in which the Hebb rule becomes particularly important: new functional states are formed according to this rule. At this stage we have to include into our discussion the tenet that there is a strong interrelationship between information provided from the biological system via the expression of genes and its transformation into the wiring of neurons during the development of the nervous system, and information from learning via imitation. The genetic control of the basic setup of ensemble connectivity constrains potential states and facilitates the establishment of genetically “prepared” states. Language acquisition as an interesting case, but this fundamental property is at work at all levels of basic and cognitive functions. Gerry Edelman´s theory has formulated mechanisms at the ensemble level which enable an organism to categorize an unlabeled world and allow it to behave in an adaptive fashion, not from instructions but from processes of selection upon variation. The principle is selection of correlated responses of ensembles that yield adaptive behavior.

Variations of ensemble coding correspond to mutations, because they create new potential states, and thus new memes. If we would be able to trace down synaptic connectivity in an ensemble we would certainly find variability in the pattern of synaptic connectivity for the same functional state in different animals. This variability is kept low due to self-controlling and self-repairing mechanisms at the ensemble level, but it will from time to time lead to new and more adaptive ensemble states. Multiplication of ensembles in larger brains, combinations of ensembles in the novel environment of a different brain, and loss of ensembles when they are no longer used may provide the material for selection.

Selection

It is important to keep in mind that - as at the level of genomics – the “grammar” of the interacting ensembles, with their hierarchies and parallel functions, their binding to functional groups of ensembles, and their competitive actions are quite important for both their fidelity and variability. We do not yet understand the functional principles of this complex level of interaction, but neuroscientists are busy trying to unravel parts of it, nowadays, in fact, with better chances of success using the powerful imaging techniques for brain functions.

Let me address a final point in my speculations, the question whether there is some analog for the selfishness of genes. Susan Blackmore´s book is all about this question. In her view memes want to be transmitted, and to be incorporated by imitation. In our terminology we would say, that the informational content of neural ensembles drives brains to express themselves, and to incorporate that content into a different brain. Indeed, learning is a fundamental and very effective property of any nervous system, and developmental programs of nervous systems are designed to incorporate experience into the wiring of neurons. Moreover, participation in the knowledge base of other individuals is a highly effective way of coping with a quickly changing environment. It appears to me that the most important evolutionary process during the transition from unicellular to multicellular organisms is the implementation of a network of communicating excitable cells. That was the step which allowed adaptive synapses to develop; they are preconditions for flexible behavior, a preadaptation for ensemble coding. In a next step, social life provides the opportunity to learn by imitation, and that appears to be the step when memes develop. Social life, and as a potential consequence of it, self-experience and consciousness, will certainly gain from effective imitation. The meme, therefore, would achieve increasing importance with increasing complexity of social life and self-recognition. Whether one needs to argue in favor of some selfish properties of the meme is highly debatable, as it is in the case of the gene. However, it goes without saying that arguments in favor of some “selfish” properties of memes capture at least one important aspect of neural ensembles, namely, their auto-regulatory and auto- processing properties.

Are these speculations at all useful? I have two points to add here.

(1) Cultural selectionists are keen to distinguish what can be transmitted by genes from what cannot. This attitude held in behavioral biology, too, for a long time. Mainly by analyzing brain mechanisms, we have learned that such a distinction is less fruitful than it appears. Genes are intimately involved in shaping the brain by experience, and the environment modulates the brain's genetically controlled developmental programs.

(2) Sociologists would say that such an atomistic view, as I presented it, is unnecessary and does not obey Occam´s razor. This is certainly true, but quantum theory is also not necessary for the explanation of most macrophenomena. Still, nobody would deny that quantum theory is helpful in the micro- and meso-domain, and will become more and more important with our capacity to handle extremely complex systems. We would not question that the brain is such a complex system, and we would not question that communicative processes in societies rely on the properties of these brains.