Essay 8

Constructivist Artificial Life, and Beyond

Alexander Riegler

Institute for Software Technology,
Technical University of Vienna,
Resselgasse 3/2/188,
A-1040 Wien, Austria.

Abstract

Within this paper I provide an epistemological context for Artificial Life projects. Later on, the insights which such projects will exhibit may be used as a general direction for further Artificial Life implementations. The purpose of such a model is to demonstrate by way of simulation how higher cognitive structures may emerge from building invariants by simple sensorimotor beings. By using the bottom-up methodology of Artificial Life, it is hoped to overcome problems that arise from dealing with complex systems, such as the phenomenon of cognition. The research will lead to both epistemological and technical implications.

The proposed ALife model is intended to point out the usefulness of an interdisciplinary approach including methodological approaches from disciplines such as Artificial Intelligence, Cognitive Science, Theoretical Biology, and Artificial Life. I try to put them in one single context. The epistemological background which is necessary for this purpose comes from the ideas developed in both epistemological and psychological Constructivism.

The model differs from other ALife approaches—and is somewhat radical in this sense—as it tries to start on the lowest possible level, i.e. avoids several *a priori* assumptions and anthropocentric ascriptions. Due to this characterization, the project may be alternatively viewed as testing the complementary relationship between epistemology and methodology.

Keywords

Artficial Life, Cognitive Science, Technische Kognitionswissenschaft, Theoretical Biology, Computational Neuroethology, Insect Intelligence, Autonomous Agents, Robotics, Theory of Science, Epistemology, Computational Neuroepistemology.

8.1 Statement of the Problem and Situation of Research

8.1.1 Description of the Scientific Problem

The paper is intended to provide an interdisciplinary approach to specific basic cognitive abilities including methodological approaches from disciplines such as Artificial Intelligence, Cognitive Science, Theoretical Biology and Artificial Life. The epistemological background comes from the ideas developed in Radical Constructivism (as described in section 8.1.2.3 below). For some reasons, the interdisciplinarity provides tools that allow us to circumvent problems of traditional ways of studying cognitive principles:

- Artificial Intelligence (AI) and Cognitive Science (CS) serve well in describing specific facets of the entire spectrum of intelligence. But most cognitive phenomena are nonlinear and therefore are too complex to be directly realized by a human scientist. This fact results in a situation wherein cognition is a black box and various hypotheses may serve as explanation for it. Unfortunately, Artificial Intelligence failed to recognize this problem of combinatorial explosion in its early days (Lighthill 1973). Any increase of the complexity of an 'intelligent' system increases exponentially the effort to describe analytically such a system.[1] Therefore, the expected expansibility of results following from microworld studies (where especially AI scientists tried to find out basic cognitive mechanisms by reducing the complexity of the real world to the simplicity of a toy world) failed. Since then representation of the environment has become a main key in the AI paradigm, whereas the integration of learning and representation has been neglected.

- Similar to AI and CS, which are both faced with complex natural phenomena, Artificial Life (ALife) is concerned with the "... study of man-made systems that exhibit behaviors characteristic of natural living systems" (Langton 1989b). By inverting the traditional analysis to synthesis, ALife offers the facility of programming non-linear systems since complex behavior need not have complex roots. Furthermore, it may contribute to Theoretical Biology by abstraction of possible life forms.

- Theoretical Biology provides a lot of insights about evolution, i.e. about the phylogenetic development of the living that are deduced from empirical investigations. Above all, Theoretical Biology allows to investigate what 'problem solving' means for (natural) systems in their environment. It is argued that the notion of 'problem solving' is to be replaced by some sort of evolutionary system conditions.

---

[1] Note that the problem of combinatorial explosion in complex systems will not necessarily prevent the construction of such complex system if one renounces the (impossible) complete *a priori* analysis.

- (Radical) Constructivism in the psychological meaning (e.g. Piaget 1954) stresses the importance of the intellectual development of human beings, i.e. the ontogenetic evolution. Cognition must not be seen as static ability but rather as dynamic process that has its origin in the sensorimotor stage of early childhood.

  In its epistemological meaning, Radical Constructivism (as formulated e.g. by Heinz von Foerster 1973) draws attention to problems of statements about the real world. In this Constructivist view, knowledge 'about the world' and cognition in general is not seen as a mapping of features of an external world but rather as the ability to act adequately in the environment. Ultimately, this leads to the renewed question of what theories are in the context of Constructivism.

8.1.2 Outline of the Disciplines

In the following, Artificial Intelligence, Cognitive Science, and Artificial Life are reviewed for both motivational aspects and to point out corresponding aspects with Radical Constructivism.

8.1.2.1 Artificial Intelligence & Cognitive Science

Artificial Intelligence can be characterized as the synthesis of behavioral patterns that are interpreted as being intelligent with the methods of computation. Therefore, the ultimate goal of AI is a device that exhibits some intelligent behavior, e.g. being able to generalize perceived objects and experienced events in order to generate plans for its further actions. This example describes the so-called human *information processing* paradigm (IPP) that is fundamental for most AI approaches: First, objects are perceived; then, the intelligent creature's internal mechanism uses that informational stream to generate some hypotheses about its environment; and lastly, that mechanism provides a sequence of actions that are carried out by the creature. The IPP fundamentally depends on the *Physical Symbol System Hypothesis* (PSSH). According to Newell and Simon the PSSH postulates that intelligence is grounded in the capability to manipulate symbols: "A physical-symbol system has the necessary and sufficient means for general intelligent action" (Newell & Simon 1976).

Although Cognitive Science is also concerned with the phenomenon of intellectual, or at least cognitive, abilities, it emphasizes the empirical aspect in that it tries to come up with an appropriate model of cognition rather than any devices that only exhibit some cognitive behavior. By definition (e.g. Gardner 1985), the subject of Cognitive Science is to investigate human thinking, especially cognitive, communicative, and perceptual processes. As such questions have a long tradition, it is considered to be an approach that consists of psychology, philosophy, linguistics, neuroscience, and—last, not least—computer science. Actually, it has turned out that just computer science takes a leading position in Cognitive Science in that the computational metaphor serves as foundation for most models in this area, i.e. the information processing paradigm.

Several approaches to AI and CS start with the (empirical) investigation of human intelligent behavioral patterns as provided by psychology. Hence, there are a lot of subdomains of AI which focus their attention on a small piece of the entire spectrum of the phenomenon of intelligence. Examples for the subdisciplines are automatic proof, vision, expert systems, problem solving, natural language understanding, and last, but not least, robotics.

The underlying assumptions of those traditional AI domains are the following:

- Priority of knowledge representation and conceptualization.
- Cognition can be represented in an abstract way, i.e. formalized. Therefore, in order to retrieve a solution to a given problem the intelligent device only has to search through the problem space. Furthermore, Cognition underlies a general structure.
- Cognition can be described in natural language (since every programming language is a mathematical formalization, and every formalization is a condensed form of natural language).
- There is a separation between cognition/knowledge representation and learning processes/ intellectual development.
- In an epistemological point of view, there is a projection of the self-description of the observer (i.e. the system designer and/or user) in the AI system. This will be called the *problem of projection.*

For some reasons, traditional AI systems exhibit the following shortcomings:

Lack of flexibility:

Historically, the problem space of AI systems decreased over the years: starting with the General Problem Solver (GPS) with a very broad intellectual claim in the sixties, the piece have become smaller with the era of micro-world programs like `Shrdlu`, which reduced the complexity of real world to a world of toys, and became very narrow with expert systems (XPS), which are designed to serve as intelligent decision and information tool in very specific domains. On reason for this development can be found in the so-called *Frame Problem* (FP), i.e. the problem of finding the (or at least an) appropriate knowledge representation. Dennett (1984) provides an illustration of the FP: The basic requirement for a robot device is (according to the IPP) to develop plans in order to foresee consequences of its action. Furthermore the robot should have the ability to deduce side-effects of its actions which may have an influence on its future actions. As the real world is very complex, a complete deduction of all side effects would take too long for taking any action in real-time. Hence, the robot must know to distinguish between relevant and non relevant deductions. But even this process of discrimination needs a lot of computation and therefore evaluation time as each of the deductions has to be assigned with some (quantitative) credit to evaluate their usefulness within a certain situation. This problem causes inflexibility in unforeseen situations.

Lack of robustness: The IP paradigm directly leads to a bottleneck architecture that decomposes tasks in a functional manner, i.e., received information from outside passes sequentially through various steps before any action is taken: The generation of an internal model follows perception (i.e. sensual input) and is followed by a planning process; the plan is interpreted by a plan executor module which lastly performs the (appropriate) action. The breakdown (or at least the weakness) of only one module causes the breakdown of the whole system—e.g., if the perception module is not able to provide complete information about the environment, the generation of the internal model will not cover all eventualities that have to be taken into consideration by the planning module.

Microworld trap: Traditional AI systems get their input problem in symbolic form and deliver the output solution in symbolic form, too. Both the symbolic input and output is provided by human programmers and users that interpret the symbolic output (the 0s and 1s so to say) as the solution of their specific problem.

Complex problem spaces: Due to the intellectual limitations of human programmers, i.e. the impossibility to grasp complex problem spaces, explicitly programed AI systems lack complexity. Historically, this can be documented by the development of AI systems such as *General* Problem Solver to `Shrdlu` to expert systems whose boundaries are rather small.

8.1.2.2 Artificial Life

Artificial Life (ALife) is a widespread discipline containing various approaches to the phenomenon of natural and artificial life. Christopher Langton characterizes ALife as "… a field of study devoted to understanding life by attempting to abstract the fundamental dynamical principles underlying biological phenomena, and recreating these dynamics in other physical media—such as computers—making them accessible to new kinds of experimental manipulation and testing" (Langton 1992).

Three major components of ALife systems may be outlined (and contrasted to traditional conceptions):

Bottom-up: Rejection of analytical top-down decompositions of complex systems into their components (cf. Cognitive Science). Instead, emphasis is put on a synthetic bottom-up approach (in contrast to AI's synthetic top-down aproach).

Emergence: Properties and behavior of creatures emerge from the interaction between components following local rules rather than global parameters.

Goallessness: ALife emphasizes the structural development of systems over time.

8.1.2.2.1 Bottom-Up

Representation of knowledge is done by means of local rules instead of global governing rules. As living systems may not be characterized as complex machines, which are steered from outside, but as some self-organizing systems entities, all regulative processes (as described by an observer) have to be determined by the internal structure of the system.

8.1.2.2.2 Emergence

Properties and behavior of creatures emerge only from the interaction of local rules. As we will see later, this idea corresponds closely to aspects of a constructivist view of ALife. The underlying methodological assumption is a non-reductionistic, mechanistic[2] way of explanation: the operation of certain phenomena must never be reduced to the operation of the components involved in the phenomena. In such devices the interesting global behavior (i.e. the behavior an observer attributes to the whole system) emerges from those interactions. They are referred as an Emergent Computation if the emergent behavior is also computation (cf. Forrest 1990). Therefore, Emergent Computation, which is considered to be the methodology of ALife, may be characterized in the following manner:

- The ALife system consists of a set of *agents* that are primitive computational units. They follow their own local (often individual) instructions called *low-level instructions.* Currently, there is no answer to the question of the size of the genotype for interesting non-trivial behavior to occur.
- The local units interact with each other forming a global pattern for an observer. The global pattern in turn perturbs the local interactions in that they provide the context within which the latter apply.
- Furthermore, the global pattern is naturally interpreted as computations. Being aware of the microworld trap, the concept of the observer (i.e. the system designer and/or user) is of crucial importance. Most ALife systems are interpreted by perceiving visualizations of the dynamic variables to reveal the phenomena of interest.

The methodology of Emergent Computation leads to the following features:

- In complex and dynamic environments any predefined explicit instructions will fail due to contingencies and opportunities that any intelligent device may happen to encounter. This matter of fact decreases the flexibility of the entire system. In emergent computation there are no global plans. Therefore, it is up to the emergent behavioral pattern to interact with the environment in an adaptive way.
- Nonlinear and complex systems are hard to formalize at the emergent level, if at all. Braitenberg (1984) criticized this analytic methodology and suggested a synthetic approach instead. Particularly, complex psychological behaviors may generated by simple low-level instructions. Braitenberg calls this Synthetic Psychology. His law of uphill analysis and downhill synthesis has become a paradigm in ALife. The synthetic methodology covers many problems of representation in traditional Artificial Intelligent systems (cf. the Frame Problem; i.e. the problem of appropriate representation mechanisms, as mentioned above) in that there is no need to specify any representational structures at the global level at all.
- As emergent computation does not handle symbols but rather low-level instructions that are directly connected to the domain of interest (e.g. via a sensory apparatus) the symbol grounding problem does not arise. On the contrary, it may help to find out what the use of symbols means for an agent and may therefore contribute to a theory of cognition.

8.1.2.2.3 Goallessness

In Artificial Life systems, emphasis is put on structural development over time, which is the ultimate scientific goal of both ALife and Cognitive Science—instead of providing solutions for some problems as AI does. Generally speaking, the aim of ALife research can be characterized as the generation of life-like[3] behavior (Langton 1989b) emerging from the interaction between local rules. One may tend to call this behavioristic, meaning that the output of any system is a function of the input. But this problem only emerges if ALife is put on a rationalistic background, where the projection of observer generated goals in the (learning) creature inevitably leads to some constraint with respect to correct interpretations of the observed behavior.

8.1.2.3 Radical Constructivism

In this section, the key features of Radical Constructivism are presented that may be applied to the interdisciplinary methodology of any Constructivist implementation. Due to several reasons, there

---

[2] The term 'mechanistic' does not refer to the idea of a clockwork universe wherein everything can be deduced from a certain starting point. Rather, 'mechanistic' only indicates that there are no additional meta-physical forces etc. necessary for explanations (cf. Mechanistic Explanations, section 8.1.2.3.1 below).

[3] Of course, in this context the term 'life-like' does not have any biological foundation. Rather, it has only some paradigmatic meaning in that it shows the original motivation of ALife.

will be drawn a distinction between epistemological and psychological Constructivism: Psychological Constructivism emphasizes the cognitive development of beings, especially human beings. Therefore, its starting point is a psychological one: the, to some extent, mentally 'naked' child. On the contrary, epistemological Constructivism primarily asks for what we know about the world. The main point is the concept of the observer, i.e. starting with the assertion that observing is the only access to the 'world'. That is due to the fact that an observer is a so-called operational closed system, wherein nervous signals are unspecified, i.e. visual stimuli affect the same kind of internal signals as tactile ones. Since observing—in the sense of having experiences—is a coherent coordination of actions in a community of observers, Constructivism is *not* a solipsistic philosophy.

In the following, both aspects of Constructivism are discussed in detail.

8.1.2.3.1 Epistemological Constructivism

Mechanistic Explanations. As a first step, the differentiation shall be applied that machines and living entities are mechanistic in such a way that one can explain them as deterministic systems related to their construction, but the former need not be, and the latter are not, deterministic in relation to their behavior, i.e., instead of believing in a mechanistic worldview which assumes that the explanation of a system's behavior depends on the knowledge of initial and marginal conditions, the difference between being able to explain an observed object depending on the observer, and the underlying deterministic character 'behind' the observed, is to be emphasized.

Subsequently, it is useful to distinguish between two domains:

phenomenal: built up by the components that constitute the entity.

descriptive: constructed by an entity that is determined by means of the operation of definition.

To put it differently, what occurs within a composed entity (e.g. living systems) is completely different from what happens to this entity. This is particularly evident in the case of living systems that exist in the domain of physiology and in the domain of behavior. These two phenomenal domains do not intersect since the description of a composite unity takes place in a meta-domain with respect to the description of the components that constitute that unity. An observer may simultaneously look at both.

Viability. Within the theory of Radical Constructivism, von Glasersfeld (1988) introduces the concept of *viability* in order to explain the utility of theories, mental models, and ideas. His conception is consistent with the tradition of instrumentalism. The instrumentalist point of view emphasizes the notion of a knowledge that fits observations, or, as von Glasersfeld puts it, "It is knowledge that human reason derives from experience. It does not represent a picture of the real world but provides structure and organization to experience".

At first glance one tends to believe that the notion of 'to fit' is quite different from the notion of 'to match'—the function being postulated by the realist epistemology—but on closer look we find that it depends on the level of observation. The concept of fitness (e.g., viability) also refers to a relation similar to that of to match but on the pragmatic level. In spite of all efforts to establish a function to rate theories all approaches refer to an iconic match on a certain level of observation. This is also true for the concept of von Glasersfeld.

There are important consequences for Artificial Life that come close to the Frame Problem: The artificial organism has to evaluate a model or theory in order to calculate its usefulness for the organism in a certain situation. This process is very difficult to model because of its wide reference to the organism's knowledge itself.

Moreover the process depends on the time and situation of the evaluation. This means that the ALife organism can produce a completely different evaluation of a theory in another context. So the ALife designer must analyze whether the theory of viability should be considered for the conception of an artificial organism.

Ascription. The von Glasersfeld (1988) concept of ascription (which was originated by Kant) is strongly connected to the notion of viability. If a theory is viable in a certain context the theory can also be attributed to a perception (observation) in order to explain the behavior of the perception. So one can distinguish two dimensions of viability: the first refers to the utility of the theory for oneself and the second refers to the possibility of ascription (e.g., the predictability) of a certain phenomenon.

Therefore, the concept of ascription is intended rather to explain something than to describe or to formalize something. This has to be kept in mind if we want to understand the reason for the use of words in a realist fashion. Ascriptions are ontologized in order to isolate them from their context and to generalize them. In other words, the idea of objectivity arises through the successful ascription of viable experiences. But as the components of complex constructions are anchored in sensorimotor experiences any constitution of objectivity of meaning through reference is impossible.

From a psychological point of view, acription has the following meaning: By way of constructing permanent objects the organism externalizes some invariants, which it abstracts from its experience and which it uses as independent external things from now on. The modular way of learning, which is also described by Piaget & Inhelder (1969), is basic to the notion of Constructivism, since it precisely describes the process of creating individual realities by constructing conceptual clusters out of smaller pieces, that are acquired before. But it is up to a realist epistemology to view the process of constructing as a function of an input-output information transfer between a creature and its environment (for example, reading this application paper). Indeed, the distinction between a creature and its environment (e.g. between the reader and this paper) itself is a construction of an observer (namely the reader of this paper).

For ALife, ascription is a key concept due to the necessity for organisms to predict their environment.

No Teleonomy Teleonomy and significance are characteristics of an observer's description only. Hence, a mechanistic explanation is not just an attempt to clarify why some facts lead to others but rather to reproduce the observed phenomenon in a symbolic representation. What happens in the operation of living systems at any given time takes place in a way that is strictly determined by local interactions, as opposed to being the result of external control, i.e. the result of the purpose of an external designer. Furthermore, the concept of ascription also leads to the conclusion that any behavior of an organism may only be 'purposeful' in the eye of the beholder. That also applies to a human being that observes itself (i.e. self-reflection).

The Question of Material The question whether the actual material out of which models of cognitive entities are built has any influence on the dynamics which occur can be answered by the following distinction (Maturana & Varela 1980):

Organization: the relations which determine a system as a unit and its possible interactions and transformations, so that the unit is defined as a member of a specific class.

Structure: the actual components of an existing system and the relations between them which have to be fulfilled in order to constitute the system.

Hence, as long as a system does not change its organization, even in the case of structural variations, it is a member of a specific class and therefore it makes no difference whether systems (also living systems) are made of some materials (which are defined as actual relations in a given system) or other materials. In other words: There may be various possible structures that all have the same organization. Systems undergoing changes in their structures without leaving a distinct class display a *structural plasticity*, and, as a matter of fact, the components of all living systems are being perpetually disassembled and rebuilt, such as during metabolic processes. Since the structure only determines the space of the component's interactions among themselves, it is *insufficient* to describe a system simply by reproducing its structure.

8.1.2.3.2 Psychological Constructivism

To implement cognitive creatures, the work of Jean Piaget (1954; 1969), especially his theory of sensorimotor development, is of crucial importance.

According to Piaget, the cognitive development of human beings can be divided into four main levels:

Sensorimotor: Transition of responses of simple reflexes to the use of symbols.

Preoperative: Thinking like a film that allows no flexibility as in actual symbolic thinking.

Concrete-operative: Arbitrary use of symbols and capability to group allows flexibility and coherence.

Formal-operative: Amplification of cognitive operations to abstract entities and hypothetical-deductive reasoning.

In contrast to the reductionism of AI which separates cognitive components into different domains, Piaget emphasizes both a holism with regard to the domains and an object holism. The latter claims a holistic organization and functioning of cognition. He proceeds on the assumption that the initial mental equipment of neonates is rather poor and restricted to a small set of isolated reflexes such as crying, sucking, kicking, and so on. Furthermore, the neonate is completely unable to differentiate between its 'ego', its body, and the environment. Piaget calls this *adualism.*

These presuppositions imply that there is a great challenge to discover mechanisms which lead from the most primitive innate reflexes to complex forms of cognition. In general, two mechanisms accomplish this development:

- *Assimilation* is responsible for integrating stimuli into already existing internal structures. It is very important to note that a response is never a mere reaction to external stimuli. Rather, a response is always based on the internal structure of the individual.
- The notion of *accommodation* may be contrasted to assimilation in that the former connotes the tendency of the internal structure to fit some environmental events.

The first period, the sensorimotor stage, within which cognition is linked to the content of specific sensory inputs or motoric actions, will be subject to an implementation, i.e. the question how it happens that cognitive creatures get ‘symbolic ideas’ about their world. According to Piaget, this symbol grounding process of getting knowledge about the world happens as follows (Furth 1969):

1. Behavioral patterns consists mainly of innate reflexes.
2. Development of new behavior by coordination of reflexes, but no relationship between means and goals, i.e. the infant is not able to recognize some causal relationships.
3. The infant ‘discovers’ its environment by better coordination of sensors, especially the movements of hand and eye. Circular responses are the reason to repeat a fascinating action over and over again.
4. When the infant successfully controls basic sensorimotor coordinations it starts to imitate events in an explorative manner. It learns to distinguish between means and goals. Furthermore, the infant develops abstract anticipatory capabilities, i.e. it is able to anticipate events that do not depend on its own activity. This is the first step to recognizing that there are independent objects.
5. The infant tries to find out new plans (*schème*) by trial and error always monitoring the utility of the taken action.
6. Here, new plans are also developed by internal coordination, i.e. the process of *interiorization* takes place which serves as precondition for the use of symbols.

The slow coordination of cognitive schemata is fundamental for even the logic-mathematical intellect. For a Constructivist implementation, the fact of increasing *multimodal* coordination is very important—e.g. to grasp means the concurrent use of visual and tactile sensors. Working with multimodal experience, the infant constructs the idea of invariant permanent objects.

8.2 Outline of a Constructivist Artificial Life Implementation

As mentioned above, this paper seeks to outline the importance of emergent computation through a radical constructivist implementation of a cognitive model. The purpose differs from traditional ALife approaches—and is somewhat radical in this sense —as it tries to start on the lowest possible level, i.e. avoids several *a priori* assumptions and anthropocentric ascriptions. Due to this characterization, an appropriate implementation may be alternatively viewed as testing the complementary relationship between epistemology and methodology: each step towards fulfilling the Constructivist claim makes the actual implementation more difficult as it entails more unusual techniques. On the other hand, using traditional concepts (cf. the symbolic approach in AI) runs the risk of neglecting those essential epistemological aspects which are addressed by Constructivism.

For the purpose of modelling cognitive entities we might separate the implementation into four modules (cf. Peschl 1991) detailed in following sections.

8.2.1 Design of the cognitive systems

Here, the low-level architecture of each individual is specified. According to the Braitenberg (1984) “bricks”, this design may consist of several distinct components, as follows.

8.2.1.1 A *sensory* apparatus

This component is capable to receive various environmental perturbations of various kinds, e.g. visual, acoustic, tactile, olfactory, proprioceptive[4] perturbations etc. as well.

---

[4] Proprioception connotes the direct ‘perception’ of the orientation of limbs or eyes via muscle tension etc. Again, the organism has to grasp this relationship from scratch.

Much emphasis will be put on the visual apparatus which will be equipped with a kind of primitive retina, i.e. a mapping system that reflects some physical features of objects in an individual's neighborhood but which is heavily influenced by internal states. Following von Foerster (1973), the ratio of internal to external sensors in humans is about $10^6 : 1$. This relation expresses the enormous psychological Constructivist aspect of perception.

Another interesting issue might be the *reafference principle,* i.e. how animals distinguishes between reafferent (internal) and exafferent (external) perturbations and how they compensate only for the reafferent stimuli. The *animate vision* approach (Cliff 1991) primarily deals with the reafference principle in that it emphasizes the dynamic aspect of vision. By moving around in a systematic manner, the visual capability is improved, since the movement helps to abstract invariants. Therefore, vision and perception in general (as the animate vision approach may be extended to other modalities, too) is rather behavior than passive mapping.

Furthermore, much emphasis will be put on proprioceptive sensors that are capable of sensing internal states. They are important to avoid pure reactive systems[5] in which the sensor input is directly mapped on the effector input. It is not expected that such architectures will exhibit complex behaviors due to their directness. Instead, 'short-circuits' in the effector-environment-sensor loop that leaves out the environment are the prerequisite for learning and increase extremely the spectrum of behavior.

There are a lot of further kinds of sensors one may think of—e.g. it is known that pigeons use three kinds of 'compasses': they are capable of perceiving the magnetic field of the earth, they may get navigational aids by perceiving polarized light as well as through recognizing stellar constellations. The orientation of bees is based on 'knowing' the angle between the location of food and the sun. Hence, it seems quite reasonable to provide an external reference point, a 'sun' over the two dimensional area so to say. As it is a main motivation of a Constructivist ALife system to investigate how it comes that something serves as external reference point, this matter of fact is of course *not* known *a priori* by the creatures.

8.2.1.2 *Effectors*

The effectors can be thought of as locomotion, acoustic and visual utterance (in order to developtment some kinds of communication through building consensual domains between several creatures), grasping, etc. Actually, the Radical Constructivist methodology also apply to the design of the effectors. For instance, although some components of the creature are used for locomotion as they were designed by the programmer, the knowledge about the functionality of their body parts is not *a priori* 'known' (i.e. the program that embodies a creature cannot explicitly make use of this matter of fact) by the creatures themselves. They don't even 'know' that they can move around. It is only by experience that the algorithm may learn that the activation of a locomotion element causes reafferent stimuli.

8.2.1.3 The *'black box'* in between

This component serves as connection between sensors and effectors. As shown by Braitenberg, those connections need not to be complicated in order to exhibit complex emergent behavioral patterns. Simple Connections between two sensors and two locomotion elements exhibit forms of attractive behavior, such as photo- and chemo-taxis.

Several computational methodologies may apply for performing the connections: Neural Networks, Finite State Machines (FSA), etc. Although computationally Turing-equivalent, recent research (e.g. Jefferson *et al.* 1992) has shown that FSA are slightly better that neural networks for this problem due to better internal representation. It is up to the intended implementation to verify this result. FSA are also more appropriate to serve as vehicle for some kind of Piagetian schema mechanism as their data structures are quite similar to schemes.

8.2.1.3.1 Learning mechanisms

Winograd & Flores (1986) reviewed various approaches to learning in Artificial Intelligence. They distinguish three different kinds of learning:

- *Parameter adjustment.*

  "... a fixed structure is in place, and the learning consists of adjusting some kind of weights to achieve a higher measure of performance..."

  This approach partially corresponds with the paradigm of learning in neural networks due to similarities of adjusting weights which represent the knowledge of the system. Within a mechanistic world view, all learning approaches

---

[5] Several approaches in current ALife approach follow the reactive systems approach that is hoped to circumvent problems of complex internal representations as they appear in traditional AI robotic devices (Maes 1990b).

can be (theoretically) reduced to this approach. This matter of fact depends on both the level of observation and the intended consequences of the manipulation of the chosen representational structure. Although the authors emphasize the limits of this approach we cannot resist mentioning that these limits have general scope and will also define the space of possible results of other kinds of learning.

- *Combinatorial concept formation.*

  "… the programmer begins by creating a representation, and the learning consists of finding (and storing for later use) combinations of its elements that satisfy some criterion…"

  This approach is widely applied in the field of traditional AI and related disciplines. At this point it should be pointed out that the representational structure, which does not incorporate the background knowledge, determines the limits of the considered formal system and can lead to a breakdown.

- *Evolution of structure.*

  "… the initial system does not have a structure directly related to the task as seen by its designer…"

  This approach is compatible with the ALife view of learning. Here, the problem of background does not arise since the approach is not intended to determine the representational structure of the actual problem by means of symbolic presuppositions.

Although the first two kinds of learning are applied at a microscopic level the third approach differs with regard to the epistemological foundation: There are no fixed representational implications determining the behavior of the overall system.

If we want to appreciate the problem of learning in Artificial Life correctly we have to consider the various approaches on learning. The distinctions above are a very general form of classifying different paradigms of learning. According to the author's opinion the discussion, especially the Evolution of Structure approach, does not adequately deal with the epistemological problems of knowledge. For one thing, the authors argue that, however, evolution is massively parallel so that today's computational performance is far from being capable of simulating evolution; and then, one cannot speed up the origin of species during evolution due to the process of structural coupling, as described by Maturana. According to Winograd and Flores, former evolutionary approaches in AI failed because they ignored the structural complexity (in terms of Maturana) assuming that organisms mainly consist of uniform components and that most behavioral patterns are learned instead of being innate. Above all, these arguments may not apply to an ALife project since we do not want to simulate the evolution of human beings or mammals as it has happened from the first unicellular organism until today. Rather, ALife emphasizes alternative forms of life as well. Even the process of structural coupling doesn't say anything about the redundancy in evolution. To ascribe any purpose to evolution (such as that the purpose was the emergence of human beings) is exactly an ascription in terms of von Glasersfeld and therefore in contradiction to Constructivism which rejects teleonomy. Hence, from a Constructivist point of view evolution is not a process of optimization but rather a structural drift, as Maturana puts it. Even so, one must not underrate the overall complexity. Especially, the question of *a priori*, i.e. innate, mechanisms, is of great importance and one may define ALife as discipline to find out these foundations.

From a methodological point of view, quite popular algorithms are *Classifier Systems*, which are a kind of genetic machine learning in that they are motivated by biological aspects (e.g. Booker *et al.* 1989). Technically, they are rule-based systems typically used for inexactly and implicitly defined tasks in environment which provide large amounts of noisy and irrelevant data.

Several problems that arise from the usage of classifier systems for ALife models may be addressed by the implementation, e.g.:

- *Coding of genotypes/schemata.*

  Originally, genotypes are binary coded such that each genome can either be set or not. For quite simple problems, this coding is sufficient.[6] Research on more complex ALife problems—e.g. (Collins & Jefferson 1992) where genomes are used that are 25590 bits long in order to encode whole neural networks—suggests that alternative coding strategies may apply better.

- *Evaluation of the 'fitness' function.*

  The use of genetic algorithms depends heavily on a function that serves as criterion for which agents will reproduce. In several ALife systems

---

[6] Experiments have been done upon the question whether gray coding is better than ordinary binary coding (Caruana & Schaffer 1988). Due to the so-called 'Hamming cliff' in binary coding which represents a counterproductive bias for certain problems. However, there is no general rule of thumb.

this fitness function corresponds with some internal 'level of energy' in each agent, which is the overall scoring of 'successful' actions minus the sum of 'mistakes' scores (i.e. actions that were not successful). This is contrary to a Constructivist epistemology: Success and mistake are not ontological observer independent criteria but only real in the domain of reality that is brought forth by operationally coherent actions. To commit a mistake means only that the reality expected by the observer is different to the one within which the 'not successful' action takes place.

8.2.2 Design of the environment

For performance reasons, the environment, within which the creatures act, is assumed to be two dimensional and both spatially and temporally discrete. This does not necessarily mean to decrease the explanatory power since the essential criterion is to avoid *a priori* cognitive structures in the individuals that reflects ascribed anthropocentric notions, i.e. to escape the microworld trap. Hence, the rich populated environment serves rather as source of perturbation than as blocksworld in that there are static objects whose anthropocentric attributes are ascribed to the creatures' cognitive world. The individuals will gain their own ideas (*Begrifflichkeit*).

The environment itself consists of different objects (both static and animated) that can be distinguished by their visual appearance, such as their size, location, albedo, etc., their acoustic appearance, their energy content (in order to serve as sources of energy for the creatures).

The objects are *not* considered to differ with regard to their shape. Only on the graphic user interface they look different but not for the modeled creatures. This restriction avoids computational costs that may arise due to complex ray tracing algorithms otherwise.

8.2.3 Modelling the interaction between cognitive systems and their environment

This module consists of algorithms that manage the 'physics' of the implementation, i.e. the module computes the perturbations each creature experiences in its environment. To maintain the constructivist idea in the implementation, there must not be any *a priori* categories that take place in the cognitive apparatus of the creatures. That is to say, in most related implementations the creatures 'perceive' animated and inanimated objects by their name (e.g. creatures are able to distinguish *a priori* between predators and members of the same species) and act with regard to what they 'see'. These presumptions are in conflict with a radical constructivist methodology. Hence, the main question in the intended implementation may be formulated as: How does it *come* to recognize predators, food sources, and even communication signals?

8.2.4 Interface and Control Device

As stated above, most ALife systems relying on the methodology of Emergent Computation are interpreted by perceiving the visualizations of the dynamic variables to reveal the phenomena of interest. That becomes evidently true for a Constructivist implementation as there are a huge number of variables on the one hand and the interesting phenomena are biologically motivated. Therefore, visualization techniques will be an integral part of the model.

Methodologically important, the user interface should provide many degrees of freedom for the experimentalist to systematically vary the parameters of the implementation as well as different way of investigating the complex relationships between low-level rules and behavioral patterns.

Furthermore, there has to be a well defined interface between this module and the other three components as the latter may become too computationally intensive to be performed on the same computing device that serves as control and investigation module.

8.3 Purpose, Intended Methodology, and Implications of a Constructivist Artificial Life Model

8.3.1 Purpose

The purpose of the model is to show how an individual creates his or her own world. This is in a line with Jean Piaget's "L'intelligence … organise le monde en s'organisant elle-meme" (Piaget 1954).

The underlying motivation for an Constructivist ALife model is that—contrary to Artificial Intelligence in general and Parallel Distributed Processing (PDP) in particular—Artificial Life approaches provide a more general view on phenomena such as life and cognition. As in AI the solution is the ultimate goal and PDP amplifies the pool of methodologies by borrowing ideas from neurophysiology. ALife is a more widespread discipline looking in nonreductionistic fashion at the phenomena of life, cognition, behavior, learning, and so on. Therefore, the development of an alternative understanding of the cognitive sciences can be achieved within two steps (see sections 8.3.2 and 8.3.3 for a more comprehensive description): The first step will be to enrich ALife with the epistemological basis of Radical Constructivism in order to form a *Radical Constructivist* Artificial Life (RACAL). In the second step RACAL and traditional cognitive sciences may be fused with respect to their methodology and scientific goals. This leads to a Technical Cognitive Science, or in short: TKW.[7] This conception of future developments fits traditional disciplines as shown in Figure 8.1.

8.3.2 Step One: Radical Constructivist Artificial Life—RACAL

Applying the idea of a distinction between structure and organization to define a Radical Constructivist Artificial Life (RACAL) means to separate the (physical, chemical and so on) properties of components from the properties observers attribute to the observed creatures. As far as the physical aspects are concerned, the components of living creatures are not unusual. Hence, arbitrary mixtures of these components make no sense, and any reproduction of the structure of a system is insufficient with regard to explaining the system. On the other hand, putting components together in such a way that the constituted entity belongs to the intended class of systems (as defined by its organization), the reproduction is successful.

8.3.2.1 Autopoiesis

The notion of autopoiesis has been adapted in many disciplines to characterize self-organizing entities in general. Indeed, autopoiesis is not just a synonym for such systems but has a very precise meaning.

On the one hand the components of autopoietic systems take part in the recursive production of the network of production of components that produced those components. In other words, contrary to common ideas of purposeful machines with input and output, a living system is primarily concerned with maintaining itself by keeping its organization (not necessarily its structure) constant.[8] Activities which appear as the maintenance of organization occurs are considered purposeful and meaningful by observers and *only by observers*.[9] As discussed above, such descriptions can lead to confusion: We are accustomed to stating that there are representations of the world inside the brain with the help of which living creatures cope with the environment. We usually take these metaphorical conceptions at face value. Hence, it is not surprising that Artificial Intelligence dealing with machine learning tries to find the ‘correct’ forms of representation, e.g. semantic representations of the environment. Their usual lack of success speaks for itself.

Furthermore, an entity exists in the space within which the components exist by determining the topology of the network of processes.

The difference between allopoietic and autopoietic systems can be characterized by the products of their operation: An autopoietic system, defined as a unity through relations of production of its components, not through the components themselves, is always and only producing itself, while an allopoietic system’s products are something different from the system itself. Systems constructed by Artificial Intelligence scientists to solve a problem or to exhibit some intentionally useful behavior are always allopoietic and never ‘alive’. Therefore, discussing the lifelikeness of artificial systems argued on the basis of behavior is misleading.

8.3.2.2 The importance of background modelling for ALife

Winograd & Flores (1986) put the emphasis on the background knowledge being essential for our understanding, but which is neglected by traditional methods of representation. Consequently, this aspect may form a basic distinction between Artificial Intelligence and ALife. The tradition which does not support this point of view but emphasizes the formal descriptions regarding language as a system of systems referring to certain entities in the world is called the rationalistic tradition. It it scarcely necessary to point out that the advantages of this scientific approach consist in the computational power

---

[7] Abbreviation of the German translation: *Technische Kognitionswissenschaft.*

[8] Similarly, living creatures may also be characterized by the maintenance of their organization against the environment.

[9] Technically, for Maturana an observer is “. . . a system that through recursive interactions with its own linguistic states may always linguistically interact with its own states as if with representations of its interactions” (Maturana & Varela 1980).

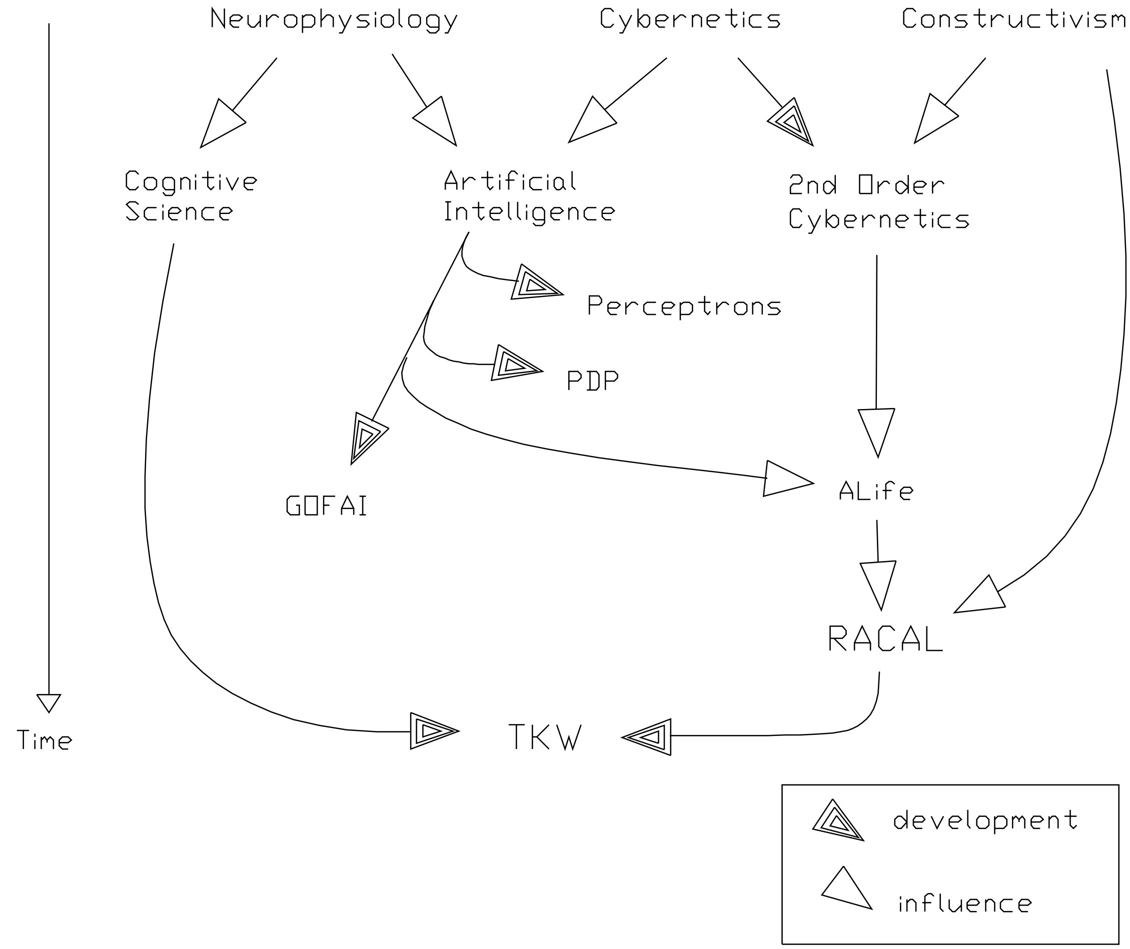

Figure 8.1: Relationships between disciplines.

of this reductionistic world model due to its simplicity. By way of contrast, in Artificial Life powerful computational performance is needed in order to adequately simulate the complexity of the world. Due to structural coupling, which may be seen as a requirement for intelligence, the environment of the simulated creatures has to exhibit a similar structural plasticity in order to allow the development of a domain wherein interactions can take place. This aspect may be considered as the most remarkable drawback of Artificial Life at present.

8.3.2.3 Consequences

- *Organism and environment are not independent of each other.*

  From this it follows that the structural change of the environment as a result of the structural coupling with the organism is as important as the structural change in the organism. So ALife should not only focus on the simulation of perception in the organism but also consider the possible changes in the environment. As it is common to speak of an ontogenetic development of an organism it should also become common to do so with regard to the environment.

- *Organisms use their mechanism to react, rarely by explicitly formulated goals.*

  This thesis holds that for the explanation of the behavior of an organism using different level of rules, as it is common in Artificial Intelligence, is not adequate. From this it follows that the methods applied in AI cannot perfectly simulate the behavior essential for the development of intelligence. Furthermore, the role of mental models has to be questioned. Meaning denotes a social construct, established in the structure of language and forms of social interactions. Interactions between different organism lead to a mutual ontogenetic structural coupling.

Furthermore it has to be pointed out that this matter of fact does not determine the structure of the mechanisms which naturally can be based on rules.

- Thinking is not a process of manipulating representations of an external world.

  AI is based on this misconception which has its roots in the rationalistic tradition as defined by Winograd & Flores (1986). This epistemological background of AI is determined by the suppositions that the external world consists of entities with well defined properties.

8.3.3 Step Two: Technical Cognitive Science—TKW

8.3.3.1 Cognitive domain

Since autopoietic systems are structurally plastic systems, they can interact with other systems without losing their identity. This is called their cognitive domain. In evolution, the existence of any nervous system has increased that domain of possible interactions. To put it differently, cognition is not considered a qualitative property, only emergent in living creatures at a high level (e.g. vertebrates or only mammals) but rather connotes a system's capability to cope with its environment. According to Maturana & Varela (1980) it may be said that "... Living systems ... are not made to handle a medium [their 'environment'], although it has been through the evolution of their handling of their medium that they have become what they are, so that we can say what we can say about them." This rather general assertion needs to be made more precise—for instance, by putting emphasis on viewing the systematic correlation between invariants and internal 'processing' mechanisms.

However, there is no sense to provide any simulation of cognitive systems with a kind of fitness function against which the systems are compared.

Put another way, on the one hand a Radical Constructivist Cognitive Science negates any teleonomic claim when explaining the phenomenon of living systems, on the other hand it views the operation of the living as subsequent structural coupling with the medium in which it takes place. Therefore, cognition is seen as a general feature of the interaction between living systems and their environment if it is based on adequate interpretation from the environment as stimulus for own action by the organism. These terms are only useful in the description of an external observer who distinguishes and observes between organism and environment simultaneously. This is exactly the matching point between ALife and Cognitive Science: in TKW to deal with the phenomenon cognition means no longer to deal with a new quality that can be separated into various *subcognitive* domains, such as representational aspects, vision and so on. Rather, it offers the possibility to integrate *evolutionary* (both ontogenetic and phylogenetic) aspects.

8.3.3.2 External and internal worlds

At this point, it is useful to introduce two concepts that are borrowed from the German language in which there is a distinction between two words for reality:

- *Realität* (from the Latin *res*, meaning "thing") connotes the ontological given environment every realist makes reference to. In this context, scientific investigations seek for 'true' knowledge, i.e. knowledge that could be said to correspond to that ontological *Realität.*
- *Wirklichkeit* (comes from the German verb *wirken* and means to have an effect on) designates the "constructed" world in our minds, as the constructivist position proposes. Thus, *Wirklichkeit* connotes a sequence of "effects" (perturbations) which appear at any time at any place. The reference elements for "knowledge acquisition", which we call "phenomena" or "facts", are therefore spatially and temporally constrained configurations of effects.

This distinction between two forms of reality fits into Maturana's conception[10] who emphasizes the subject-dependency of cognition and observation.[11] There is no objective world, no *Realität* in the sense used above, which is independent of an actual living system. The particular world (*Wirklichkeit*) is constructed through the ongoing interaction of the living system in its cognitive domain, i.e. in that environment that is determined by the possible interactions the system can enter without loosing its

---

[10] It is necessary to point out, that Maturana only refers to the English expression "reality" that connotes that domain which is specified by the operations of the observer (cf. Maturana & Varela 1980). Hence, "reality" is synonymous with *Wirklichkeit.*

[11] "Everything said is said by an observer to another observer that could be himself" (Maturana & Varela 1980).

characteristic identity. Thus there is no (pictorial) representation of entities in an outside world (*Realität*), but rather a network of dynamic correlations between the sensory and effectors. At this point we can state that cognition is no new quality that appears in living systems; instead, an observer should view the cognitive domain as the domain of possibilities a system can take to cope with its environment. Hence, there is no need for any criteria to attribute a system's behavior to be cognitive or intelligent. Cognition is a substantial component of the living. The existence of a nervous system is not a criterion either, since it only amplifies the possibilities of interaction and therefore the cognitive domain. Or, to put this concept into the context of TKW, we can use cognition and behavior in a methodologically synonymous way.

The inside and outside of a closed system exist only for an observer, not for the system *itself*. Thus, there are no sensory and effector surfaces that separate a cognitive organism from its environment. Instead, the environment in which the observer exists acts only as an intervening element through which the effector and sensory neurons interact.

To put it differently, according to Constructivism the organism (and its nervous system) is closed. Hence, it cannot distinguish between internal and external activities. Therefore interactions are not instructive, and perception is only the expression of an observer's description of the close interaction (*structural coupling*) between an organism and its medium. Therefore, as the cognitive domain of an organism is closely tied with changes of internal states, the evolution of that cognitive domain cannot be seen as an increase of the quality of representations of *Realität*. Rather, the constitution and evolution of cognitive domains, of an organism's *Wirklichkeit* depends upon an active sensorimotor capturing of that individual *Wirklichkeit*.

8.3.4 Reasons for the Evolution of Cognitive Sciences into TKW

In (Stary *et al.* 1992) we pointed out that the more biological concepts are simulated, the more fundamental become epistemological issues. Hence, an epistemological foundation—such as Radical Constructivism—is important to the subject of Artificial Life. Leaving out the above advantage of having a better approach to the phenomenon of life, Radical Constructivism and ALife addresses several problems, discussed in following sections.

8.3.4.1 Knowledge Representation

Cognitive systems gain their 'knowledge' (in terms of an observer) by way of acting in the world. Hence, we must not try to model some features of *Realität* in a direct manner, i.e. ascribing a formalization of our understanding of *Wirklichkeit* to creatures, but rather to let them develop their own *Wirklichkeit*.

8.3.4.2 A Methodological Issue for TKW

Many problems arising through the possibility of creating ALife as well as Cognitive Science are nonlinear and therefore are too complex to be directly realized by a human scientist. Therefore, in a pragmatic view it is useful to distinguish between linear systems and nonlinear systems.[12] The key features of *linear* systems are that the complete system can be understood by studying its parts in vacuo (i.e. isolated), and that their behavior can be predicted by modelling the observed system, i.e. to take apart the system, to analyze it and to put it together in the desired way. Hence, analysis—or top-down approach—is the appropriate method for linear allopoietic systems.

On the other hand, *the behavior of a nonlinear system* "… is *more* than the sum of its parts" (Langton 1989b). In this case, a different approach has to be chosen which goes bottom-up: synthesis that has become the underlying paradigm of Artificial Life: "Rather than start with the behavior of interest and attempting to analyze it into its constituent parts, we start with constituent parts and put them together in the attempt to synthesize the behavior of interest" (Langton 1989b). In epistemological Constructivism, the latter case demonstrates the difference between the domain of the behavior of a composite entity and the domain of behavior of its parts. Now to formulate a strong Cognitive Science, cognitive systems can be understood *only* by synthesis.[13] That is the reason for using the adjective *Technical* for TKW (Riegler 1991).

---

[12] These terms are borrowed from mathematics: The superposition principle connotes the fact that the sum of the solution of two homogenous differential equations is itself the solution of such an equation. If a system obeys the superposition principle it is said to be linear, otherwise nonlinear. In ALife, these notions are used in an analogous way (Langton 1989b).

[13] This argumentation goes back to the Italian philosopher G. Vico. His *verum ipsum factum* (we know only what we put together) expresses the fact that only constructed, i.e. technical, systems can be understood.

8.3.5 Technical Results to be expected

Traditionally, robotics is an important subdomain of Artificial Intelligence whose purpose is to construct robot devices that can survive and fullfil some tasks in an uncertain environment. Unfortunately, a pure symbolic approach in which the whole behavioral pattern (e.g. the stimulus-response patterns) are *a priori* programmed. Whenever the robot enters a new unforeseen situation there is no 'correct' answer to that challenge.

According to Luc Steels (1990), there are three reasons for the robot's failure:

- It is not clear how enough information may be retrieved from the environment in real-time.
- The internal *a priori* model must always only be an approximation to given situations
- As the robot is entirely constructed by humans, certain presumptions which take place in the implementation drive the robot to be inflexible and brittle.

However, it has been argued that robotics may overcome these problems by avoiding the functional decomposition in the design of robot architectures. Instead, as described by Brooks (e.g. Brooks 1991b), each subsystem of a robot should be concerned with a complete task, e.g. wandering around, avoiding obstacles, etc. None of the subsystems is governed or controlled by other subsystem. Rather, they are activated according to specific sensor signals and control the robot in certain situations. Unfortunately, no actual learning or evolutionary element has been brought in the conceptual design of this so-called subsumption architecture.

In contrast to the subsumption architecture, a (future) goal of a Constructivist ALife model may be a robot device that have the same working methodology, i.e. to select appropriate schemata in order to react to certain stimuli from its environment. The schemata itself are the product of a (quasi-Piagetian) learning mechanism that is similar to the development of humans at their sensorimotor period.

Acknowledgements

I would like to thank Stuart A. Umpleby who was the initiator of the an early version of this paper, and Peter Krall and Ludwig Huber for their comments.