Prof., Dr. Sc.techn., Rolf Pfeifer
Artificial Intelligence Laboratory, University of Zurich

EMBODIED ARTIFICIAL INTELLIGENCE – ON THE ROLE OF MORPHOLOGY AND MATERIALS IN THE EMERGENCE OF ADAPTIVE BEHAVIOUR

In the early days of artificial intelligence the focus was on abstract thinking and problem solving. These phenomena could be naturally mapped onto algorithms, which is why originally artificial intelligence was considered to be part of computer science. Over time, it turned out that this view was too limited to understand natural forms of intelligence and that embodiment must be taken into account. As a consequence the focus changed to systems that are able to autonomously interact with their environment. The major implications of embodiment, dynamical and information theoretic, are illustrated in a number of case studies. Two grand challenges, evolving grounded intelligence and exploring ecological balance, i.e. the relation between task environment, morphology, materials, and control in an artificial organism, are discussed.

1. INTRODUCTION

Computer science has grown into an enormous discipline with many subfields and it is often hard to see how the different areas are still connected to form one discipline, except that they all, one way or other, deal with computers. What about Artificial Intelligence? For several decades, i.e. from the 50s until the mid-80s it was mostly concerned with algorithms, for example for playing chess, checkers (and other games), solving cryptarithmetic puzzles, or natural language processing of written text. Because of this perspective, it was considered a subdiscipline of computer science. As we will be arguing below, there have been severe limitations of this approach because of its focus on algorithms exclusively. Over time, it became clear that intelligence was not so much a question of algorithms but of the interaction of an agent with the real world and researchers started using robots as their workhorse. This change in orientation entails many new research issues that are well outside the field of computer science. Not all researchers in artificial intelligence have changed direction; many are continuing to pursue the algorithmic approach. Which direction one is interested in depends on the goals: If the goal is to find a solution to a problem, the algorithmic approach might be best (e.g. Wolfgang Wahlster, this volume). However, if it is to understand the principles underlying (naturally) intelligent behavior, the alternative approach, i.e. the one of embodiment is better suited, as will be argued later.

We begin with a short history of artificial intelligence. Then we introduce the concept of embodiment and provide a set of case studies to illustrate the different kinds of implications. Next we attempt to characterize the state of the art in the field of embodied artificial intelligence. This is followed by an outline of some of the grand challenges.

2. A BRIEF HISTORY OF ARTIFICIAL INTELLIGENCE

The field of artificial intelligence has dramatically changed during the past 15 years-or-so. Initially, starting in the fifties, intelligence was essentially considered to be synonymous with thinking, i.e. with problem solving, reasoning, and logical deduction. Within this framework thinking could naturally be conceptualized as a sequence of steps, as algorithms. The main idea of the classical or traditional approach in artificial intelligence can be captured in the so-called cognitivistic paradigm which states that cognition can be viewed as computation, cognition being a very general term for mental processes. This implies that intelligence can be studied at the level of algorithms and there is no need to investigate the underlying physical processes. Thus, there is a deliberate abstraction from the physical level. This paradigm has spawned a host of research and artificial intelligence grew into a large discipline consisting of many different subfields, including knowledge representation, logic, planning, natural language processing, problem solving and reasoning, expert systems, qualitative reasoning about physical processes, theorem proving, and machine learning.

During the 1980s artificial intelligence was booming, in particular the field of so-called expert systems. There had been high hopes that we would soon have computer programs capable of solving real-world problems like medical diagnosis, configuration and repair of complex devices, scheduling, commercial loan assessment, etc. By the end of the 1980s it had become clear that expert systems had not been successful. The idea underlying expert systems, that human expertise – or intelligence if you like – could be captured in a possibly large set of logical rules that could then be run on a computer, proved to be an inappropriate model of the true nature of human expertise (for a review of the arguments see, e.g. Clancey, 1997; Pfeifer and Scheier, 1999; Vinkhuyzen, 1998; Winograd and Flores, 1986). One of the most fundamental problems with such systems was the lack of grounding. Grounding means that an expert's skills are built on top of a long history of interaction with a physical and social world during which sensory-motor and perceptual skills have evolved. An implication of grounded intelligence is that abstract concepts and symbols can be meaningfully interpreted vis-à-vis the real world. It became apparent that intelligence could not be sensibly conceived of in purely computational terms.

In addition to these developments, evidence for the problems with the cognitivistic approach to artificial intelligence came from another area. Around the same time, i.e. also during the 1980s, many people started building robots. The basic idea of the traditional approach to robotics has been and still is that the essence of intelligence is to be seen in the internal, symbolic processing. All that would be required, so the rationale, is to attach a camera and some actuators in order to have a system that can interact with the real world. One could then map the camera image onto an internal representation, a model of the real world, generate a plan of action that could then be executed by the robot. In the meantime, it is well-known that this approach which constitutes the standard approach to computer vision did not pan out in general. It only worked in well-defined settings like factory environments. The limitations of viewing intelligence as a computational phenomenon exclusively became obvious. Given these insurmountable problems a radically new approach was required. Rodney Brooks of the MIT Artificial Intelligence Laboratory suggested that we forget about logic and problem solving, that we do away with thinking and with what people call high-level cognition and focus on the interaction with the real world (Brooks, 1991a, b). This interaction is, of course, always mediated by a body, i.e. the proposal was that intelligence be "embodied". What originally seemed nothing more than yet another buzzword turned out to have profound ramifications and rapidly changed the research disciplines of artificial intelligence and cognitive science. It is currently beginning to exert its influence on psychology, neurobiology, and ethology, as well as engineering.

Research in artificial intelligence employs a synthetic methodology, i.e. an approach that can be succinctly characterized as "understanding by building": by developing artifacts that mimic certain aspects of the behavior of natural systems, a deeper understanding of that behavior can be acquired. There are three aspects to the synthetic methodology: (1) building a model of some aspect of a natural system, (2) abstracting general principles of intelligence, and (3) applying these abstract principles to the design of intelligent systems. The artifacts of interest are either computer programs, as in classical artificial intelligence, or robots as in embodied artificial intelligence. In the embodied approach simulations are used as well, but they are of a particular type and include models of an independent environment that have their own dynamics, as well as the agent's sensory and motor interactions with these surroundings. The synthetic methodology contrasts with the analytic one where a given system is analyzed in a top-down manner, as is the standard way of proceeding in science.

3. EMBODIMENT

The goal of this section is to introduce the novel ideas that have been developed within the framework of embodied artificial intelligence. In particular we will show that embodiment means much more than simply "using a robot" – it requires an entirely new way of thinking, and it necessitates reflecting on the interaction with the real world; the latter is messy and not as neat as the world of the virtual machine. We start with a few comments on embodiment and then present a series of case studies.

3.1. Implications of embodiment

Embodiment has two main types of implications, physical and information theoretic. The former are concerned with physical forces, inertia, friction, vibrations, and energy dissipation, i.e. anything concerned with the (physical) dynamics of the system, the latter with the relation between sensory signals, motor control, and neural substrate. Rather than focusing on the neural substrate only, the focus is now on the complete organism which includes morphology (shape, distribution and physical characteristics of sensors and actuators, limbs, etc.) and materials. One of the surprising consequences is that often, problems that seem very hard if viewed from a purely computational perspective, turn out to be easy if the embodiment and the interaction with the environment are appropriately taken into account. For example, given a particular task environment, if the morphology is right, the amount of neural processing required may be significantly reduced (e.g. case study 1). Because of this perspective on embodiment, entirely new issues are raised and need to be taken into account. An important one concerns the so-called "ecological balance", i.e. the interplay between the sensory system, the motor system, the neural substrate, and the materials used (Hara and Pfeifer, 2000; Pfeifer, 1996; Pfeifer, 1999, 2000; Pfeifer and Scheier, 1999). Ten years of research in this new field have generated a large number of fascinating results and unexpected insights.

3.2. Case studies

We begin with a simple robotics experiment, the "Swiss Robots" and an example from artificial evolution which illustrate mostly the relation between behavior, sensor morphology, and internal mechanism. Then we discuss motor systems, in particular biped walking, and muscles where the exploitation of (physical) dynamics is demonstrated. Finally we will show how it all fits together.

3.2.1. Case study 1: The "Swiss Robots"

The "Swiss Robots" (figure 2a) can clean an arena cluttered with Styrofoam cubes (figure 2b) (which is why they are called the "Swiss Robots"). They can do this, even though they are only equipped with a simple obstacle avoidance reflex based in infrared (IR) sensors. The reflex can be described as "stimulation of right IR sensor, turn left", "stimulation of left IR sensor, turn right". If a robot happens to encounter a cube head-on, there will be no sensory stimulation because of the physical arrangement of the sensors and the robot will move forward and at the same time push the cube until it encounters another one on the side (figure 2c) at which point it will turn away. If the position of the sensors is changed (figure 2d), the robots no longer clean the arena, although the control program is exactly the same (for more detail, the reader is referred to Maris and te Boekhorst, 1996; Pfeifer and Scheier, 1998; or Pfeifer, 1999). Another powerful idea which is illustrated by this example is that if the morphology is right, control can become much simpler (in this case a simple obstacle avoidance reflex leads to clustering behavior). This point will be further illustrated when we discuss the trade-off between morphology and control in the following case study on the evolution of the morphology of an "insect eye" on a robot.

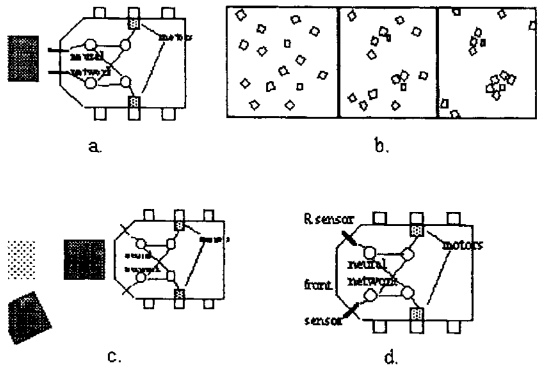


Figure 1: The "Swiss Robots". (a) Robot with IR sensors and neural network implementing a simple avoidance reflex. (b) Clustering process. (c) Explanation of cluster formation. (d) Changed morphology: modified sensor positioning (details: see text).

3.2.2. Case study 2: Evolving the morphology of an "insect eye" on a robot

When sitting in a train, looking out the window in the direction of the train, a light point, say a tree, will travel slowly across the visual field as long as the tree is well in front and far away. The closer we are getting, the more the tree will move to the side, and the faster it will move across the visual field. This is called the phenomenon of motion parallax; it is solely a result of the geometry of the system-environment interaction and does not depend on the characteristics of the visual system. If the agent is moving at a fixed lateral distance to an object with a constant speed we may want its motion detectors to deliver a constant value to reflect the constant speed. Assume now that we have an insect eye consisting of many facets or ommatidia. If they are evenly spaced, i.e. if the angles between them are constant (figure 2a), different motion detector circuits have to be used for each pair of facets. If they are more densely spaced toward the front (figured 2b), the same circuits can be used for motion detection in the entire eye. Indeed, this has been found to be the case in certain species of flies (Franceschini et al., 1992) where the same kind of motion detectors are used throughout the eye, the so-called EMDs, the Elementary Motion Detectors. Thus, if the cells are appropriately positioned much less computation has to be done. This is an illustration of how morphology can be traded for computation. Where this trade-off is chosen depends on the particular task environment, or in natural systems, on the ecological niche: natural evolution has come up with a particular solution because morphology and neural substrate have co-evolved.

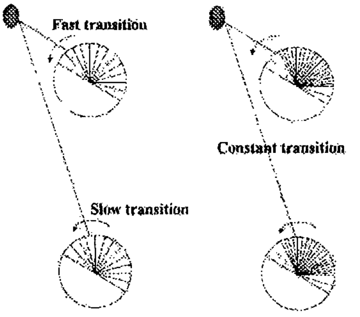


Figure 2: Trading morphology for computation. (a) Evenly spaced facets imply different motion detection circuits for different pairs of facets. (b) Inhomogeneous distribution of facets implying that the same neural circuits can be used for motion detection throughout the entire eye.

In order to explore these ideas, Lichtensteiger and Eggenberger (1999) evolved the morphology of an "insect eye" on a real robot: They fixed the neural substrate, i.e. the elementary motion detectors which were taken to be the same for all pairs of facets were not changed during the experiment, and they used a flexible morphology where they could adjust at what angles the facets were positioned (figure 3c). They used an evolution strategy (Rechenberg, 1973) to evolve the angles for the task of maintaining a minimal lateral distance to an object. The results confirm the theoretical predictions: the facets end up with an inhomogeneous distribution with a higher density towards the front (figure 3b). The idea of space-variant sensing (e.g. Ferrari et al., 1995; Toepfer et al., 1998) capitalizes on this trade-off and is gaining rapid acceptance in the field of robot vision.

Although these examples are very simple and obvious, they demonstrate the interdependence of morphology and control, a point that should always be explicitly taken into account but has todate not bee systematically studied.

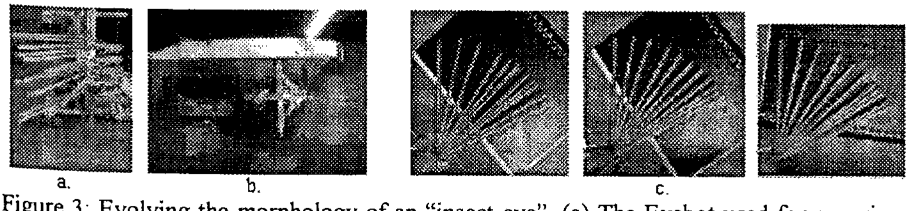
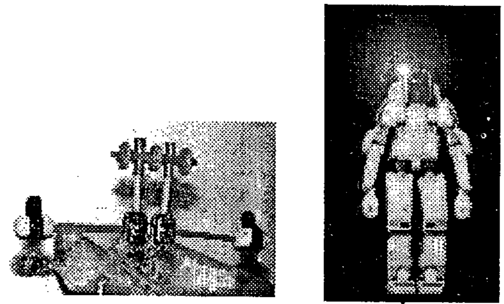


Figure 3: Evolving the morphology of an "insect eye". (a) The Eyebot used for experiments on motion parallax. (b) The experiment seen from the top. The robot has to maintain a miminal lateral distance to an obstacle (indicated by the vertical light tube) by modifying its morphology, i.e. the positioning of the facet tubes. This is under the control of an evolution strategy. The same EMDs are used for all pairs of facets. (c) Final distribution of facets from

three different runs. The front of the robot is towards the right. In all of runs, the distribution is more dense towards the front than on the side. In all of them, there are no facets directly in the front of the robot. This is because of the low resolution (the aperture) of the tubes.

3.2.3. Case study 3: The passive dynamic walker

Let us start with an example illustrating the relation between morphology, materials, and control. The passive dynamic walker (McGeer, 1990a, b), illustrated in figure 4, is a robot (or, if you like, a mechanical device) capable of walking down an incline without any actuation whatsoever. In other words, there are no motors and there is no control on the robot; it is brainless, so to speak. In order to achieve this task the passive dynamics of the robot, its body and its limbs, must be exploited. This kind of walking is very energy efficient but its "ecological niche" (i.e. the environment in which the robot is capable of operating) is extremely narrow: it only consists of inclines of certain angles. The strategy is to build a passive dynamic walker, and then to extend its ecological niche and have the robot walk on a flat surface (and later more complex environments) by only adding little actuation and control. Energy-efficiency is achieved because in this approach the robot is operated near one of its Eigenfrequencies.



a. b.

Figure 4: Two approaches to robot building. (a) The passive dynamic walker, (b) the Honda robot.

A different approach has been taken by the Honda design team. There the goal was to have a robot that could perform a large number of movements. The methodology was to record human movements and then to reproduce them on the robot which leads to a relatively natural behavior of the robot. On the other hand control is extremely complex and there is no exploitation of the intrinsic dynamics as in the case of the passive dynamic walker. The implication is also that the movement is not energy efficient. It should be noted that even if the agent is of high complexity as the Honda robot, there is nothing that prevents the exploitation of its passive dynamics.

There are two main conclusions that can be drawn from these examples. First, it is important to exploit the dynamics in order to achieve energy-efficient and natural kinds of movements. The term "natural" not only applies to biological systems, but artificial systems also have their intrinsic natural dynamics. Second, there is a kind of trade-off or balance: the better the exploitation of the dynamics, the simpler the control, the less neural processing will be required and vice versa.

3.2.4. Case study 4: Muscles – control from materials

Let us pursue this idea of exploiting the dynamics a little further and show how it can be taken into account to design actual robots. Most robot arms available today work with rigid materials and electrical motors. Natural arms, by contrast, are built of muscles, tendons, ligaments, and bones, materials that are non-rigid to varying degrees. All these materials have their own intrinsic properties like mass, stiffness, elasticity, viscosity, temporal characteristics, damping, and contraction ratio to mention but a few. These properties are all exploited in interesting ways in natural systems. For example, there is a natural position for a human arm which is determined by its anatomy and by these properties. Grasping an object like a cup with the right hand is normally done with the palm facing left, but could also be done – with considerable additional effort – the other way around. Assume now that the palm of your right hand is facing right and you let go. Your arm will immediately turn back into its natural position. This is not achieved by neural control but by the properties of the muscle-tendon system: On the one hand the system acts like a spring – the more you stretch it, the more force you have to apply and if you let go the spring moves back into its resting position. On the other there is intrinsic damping. Normally reaching equilibrium position and damping is conceived of in terms of electronic (or neural) control, whereas in this case, this is achieved (mostly) through the material properties.

These ideas can be transferred to robots. Many researchers have started building artificial muscles (for reviews of the various technologies see, e.g., Kornbluh et al., 1998 and Shahinpoor, 2000) and used them on robots, as illustrated in figure 5. ISAC, a "feeding robot", and the artificial hand by Lee and Shimoyama use pneumatic actuators, Cog the series elastic actuators, and the Face Robot shape memory alloys. Facial expressions also provide an interesting illustration for the point to be made here. If the facial tissue has the right sorts of material properties in terms of elasticity, deformability, stiffness, etc., the neural control for the facial expressions becomes much simpler. For example, for smiling, although it involves the entire face, the actuation is very simple: the "complexity" is added by the tissue properties.

a b c d

Figure 5: Robots with artificial muscles. (a) The service robot ISAC by Peters (Vanderbilt University) driven by McKibben pneumatic actuators. (b) The artificial hand by Lee and Shimoyama (University of Tokyo), driven by pneumatic actuators. (c) The humanoid robot Cog by Rodney Brooks (MIT AI Laboratory), driven by series-elastic actuators. (d) The "Face Robot" by Kobayashi, Hara, and Iida (Science University of Tokyo), driven by shape-memory alloys.

Let us briefly summarize the ideas concerning the interplay between morphology, materials, and control. First, given a particular task environment, the (physical) dynamics of the agent can be exploited which leads not only to a natural behavior of the agent, but also to higher energy-efficiency. Second, by exploiting the dynamics of the agent, often control can be significantly simplified. And third, materials have intrinsic control properties.

We have now talked about ants, simple robots, insect eyes, simple biped walkers and artificial muscles. How does this all fit together and how does it relate to intelligence? These are not questions that can be answered now but they constitute in fact, the major challenges in the field for the next 10 years.

4. STATE OF THE ART

Typically when discussing the state-of-the-art in artificial intelligence questions of the following sort are addressed: In the classical approach we could do high-level problem solving like medical diagnosis, theorem proving, and natural language processing. With the embodied approach we are doing tasks suited for robots like obstacle avoidance, navigation, homing, perhaps sorting objects into categories or manipulating physical objects. Where are we now? Have we given up on the original goal of trying to understand what people call high-level cognition? The problem with these questions is that we have never really been able to do medical diagnosis, at least not in the same way that human physicians would do it, simply because we have not yet understood the nature of human expertise. What we have been able to do is define sets of rules that capture the formal aspects of diagnostic knowledge. As shown earlier, there is now widespread agreement, that this is not a realistic way of modeling human expertise. A similar point could be made about natural language. This implies that if we are interested in the foundations of high-level cognition, this "detour" is necessary because there is increasing evidence that high-level thinking must be grounded in the sensory-motor history of an individual's interaction with its environment. We put "detour" between quotation marks to indicate that this is not actually a detour but a necessary research activity. Getting there is one of the great challenges (see section 5).

During the mid-1980s Rodney Brooks argued on the basis of natural evolution that we first need to understand simpler forms of intelligence before we can tackle higher levels as we find them in humans and that we should begin by working on insect-like robots. He developed a series of highly interesting robots such as Genghis and Hannibal that imitated at some level insect walking. These robots could learn to walk and climb over obstacles, for example. By the early 1990s he claimed that these robots had achieved insect-level intelligence and that it was time to move to something more challenging like human-level intelligence and he engaged in the Cog project (e.g. Brooks et al., 1999; see below).

While Genghis and Hannibal are fascinating and are indeed capable of imitating certain aspects of insect behavior, they are far from "insect intelligence". Just imagine what other things insects are capable of doing: the are excellent navigators, i.e. they can find their way with great precision in very taxing environments; they reproduce, they care for their offspring; they have sophisticated sensory-motor abilities; they can distinguish food from non-food; they can find food efficiently in the environment; they build amazing structures; and they often form complex societies. In this perspective Ghenghis and Hannibal are not very insect-like, and it is, in our view, an exaggeration to talk about insect-level intelligence. Again, this by no means implies that they are not interesting; they simply have not yet achieved insect-level intelligence in general.

The Cog project has the ambitious goal to eventually achieve human-level intelligence. A developmental approach is taken to the problem (for more detail, see section 5.1). The idea of this approach is to equip the robot with "human-like" sophisticated sensory and motor systems: There is a torso with arms a head with a neck, an active vision system, an acoustic system, touch sensors, and proprioceptive sensors (for measuring joint angles and forces of the robot's limbs). As it interacts with the real physical and social world it learns to make distinctions (i.e. it forms categories) and it acquires communication skills. In this way – and the is the grand goal of this project – what we call high-level cognition can be bootstrapped from this embodied interaction with the real world. Anything the robot learns is thus "grounded", to use the jargon of the field. The conviction underlying this project and that we

fully share, is that intelligence must be grounded in sensory-motor interactions. Perception is not mapping a sensory stimulation (e.g. a pixel array) onto an internal representation but a sensory-motor coordination (Dewey, 1896). Again, Cog is a fascinating robot and there is a lot of potential for research. However, talking about human-level would be an enormous exaggeration. While in the case of insects we might be inclined to believe that today's robots have achieved their level of intelligence, it is entirely obvious that this is not the case for humans – infants or adults.
The discussion about the state-of-the art in the field of robotics and artificial intelligence has always been difficult because of science fiction and horror scenarios. It is encumbered by numerous predictions that do not contribute to assessing what robots can and cannot do and how this will be in the future. Of course, nobody can predict the future, especially where technology is concerned. However, some scenarios are science fiction and do not belong into a scientific discourse.

5. SOME GRAND CHALLENGES

It will be a long way until we reach the romantic vision of understanding intelligence, whatever that would exactly mean. And there are many grand challenges that need to be resolved along the way. We discuss two that we believe cover most of the issues in the synthetic study of intelligence that we are currently aware of. They are both tightly intertwined but can be separated for the purpose of dividing up the research into manageable chunks. The first one implies understanding how we can evolve an artificial real-world agent, i.e. a robot, for high-level cognition, the second comprehension of "ecological balance".

5.1. The first challenge: Evolving grounded intelligence

One of the most fundamental abilities of agents—animals, humans, and robots—in the real world, is the capacity to make distinctions: food has to be distinguished from non-food, predators from con-specifics, the nest from the rest of the environment, and so forth. This ability is also called categorization and forms the basis of concept formation and ultimately high-level cognition. In order to elucidate the distinction between traditional computer models and embodied models, we briefly discuss a prominent traditional categorization model, ALCOVE (Kruschke, 1992). Indeed, ALCOVE is an excellent model: It can predict a large part of the experimental data published in the psychological categorization literature. In essence, ALCOVE is a connectionist model in which certain nodes, the category nodes, are activated whenever an instance of a particular category is encountered. In other words, these category nodes are representations of the categories. The task of ALCOVE can then be seen as one of mapping the input feature vector onto an internal representation of the category.
The main problem with ALCOVE, as is the problem with most models in classical cognitive psychology and classical artificial intelligence, is that it is not connected to the outside world: its inputs are feature vectors, its output activation levels of nodes in a neural network. In the real world, agents are exposed to a stream of continuously changing sensory stimulation, not to feature vectors, and they require a continuous stream of motor control. Moreover, there is the problem of object constancy, i.e. the phenomenon that the sensory stimulation from one and the same object varies enormously depending, for example, on distance, orientation, and lighting conditions. It turns out—and it has been discussed extensively in the literature—that categorization in the real world requires a completely different approach, as the history of computer vision teaches.
The insight that categorization in the real world is not an exclusively computational problem and requires that embodiment be taken into account is gaining increasing acceptance: It has

been demonstrated that categorization is best viewed as a process of sensory-motor coordination (Edelman, 1987; Metta et al., 1998; Pfeifer and Scheier, 1997; Scheier, Pfeifer, and Kuniyoshi, 1997). The sensory stimulation that the neural system has to process depends on the physical characteristics and on the positioning of the sensors on the agent. But not only that, it also crucially depends on the agent's behavior. For example, touching a bottle with a stiff hand yields entirely different sensory stimulation than fully grasping the bottle with the entire hand bent around the bottle. Note that this is a change in the morphology of the hand which leads to a change in the sensory stimulation. So, there are two closely related factors influencing the sensory stimulation, morphology, and sensory-motor coordination.

The question we have to ask now is how this all connects to the study of high-level intelligence or cognition. how does cognition come about? What we have shown is the basic ways in which neural processing, morphology, and environment are interconnected. An increasing number of people are becoming convinced that if we are to explain cognition, we must understand how it evolves during ontogenetic development (e.g. Clark, 1997; Edelman, 1987; Elman et al., 1997; Thelen and Smith, 1994). Thelen and Smith argue that while in human infants behavior is initially highly sensory-motor and is directly coupled to the system-environment interaction, during development some processes become "decoupled" from the direct sensory-motor interaction, but the underlying mechanisms, the neural substrate, is exactly the same. The advent of the discovery of mirror neurons (see, e.g. Rizzolatti et al., 2000, for an overview), i.e. neurons that are equally activated when performing or just observing an action, adds validity to this view. The question of what the mechanisms are through which, over time, this "decoupling" from the environment takes place is, to our knowledge, an unresolved research issue.

The challenge for artificial intelligence is to build robots that can mimic the processes of human infant development. This will on the one hand help us uncover the mechanisms underlying development, and on the other we will be able to build highly complex and intelligent systems. Of course, given the current state-of-the art, it is an illusion to build a robot that actually (physically) grows. Perhaps with progress in nanotechnology this may eventually be possible. But for now we have to work with non-growing robots. Given this limitation, one approach is to build a humanoid robot, i.e. a robot that has some similarity with humans in terms of morphology (shape), sensory and motor systems. The human sensory and motor systems are extremely sophisticated. For example, the entire body is covered with many sensors, e.g. touch and temperature, there are different types of sensory channels (vision, hearing, touch, smell, taste), and there are many internal (proprioceptive) sensors. We have already discussed the complex properties of muscles. Building a complex robot implies, in addition to the conceptual challenges, technological ones in terms of actuators, "tissue", and sensors. A grand challenge indeed, requiring the cooperation of many scientific disciplines from computer science, developmental psychology, neuroscience, engineering and materials science. Mimicking human infants (or toddlers) is one of the goals of the Cog project that was mentioned above.

The drawback is that we are stuck with one particular design, a complex and sophisticated one perhaps, but still a given one. Artificial intelligence has additional possibilities in that we can explore designs that do not exist in nature. But how should we design our systems, then? In order to answer this question we need to understand "ecological balance". A good method to explore a problem space is artificial evolution. We will show how it can be used to understand and explore "ecological balance" in systematic ways.

5.2 The second challenge: Understanding and exploring "ecological balance"

Using artificial evolution for design has a tradition in the field of evolutionary robotics. The standard approach there is to take a particular robot and use a genetic algorithm to evolve a control architecture for a particular task. However, if we want to explore ecological balance we must include morphology and materials into our evolutionary algorithms. The example of the Eyebot where the morphology of an "insect eye" was evolved, demonstrates another way in which evolution can be used: We fix the neural substrate and let evolution work on the morphology to solve the problem. Both of these approaches are not biologically plausible and can only be done in artificial systems.

The problem with including morphology and materials is that the search space which is already very large for control architectures, literally explodes. Moreover, if sophisticated shapes and sensors are to be evolved, the length of the genome which is required for encoding these shapes will grow very large and there is no hope that anything will ever converge. This issue can be approached in various ways, we just mention two. The first which we will not further discuss is to parameterize the shapes, thus bringing in biases from the designer on the types of shapes that are possible. In the eyebot the rods with the light-sensitive cells were given and only the angle could be adjusted, which makes the problem very simple, but then there is only little variation possible in the morphology. An example that has stirred a lot of commotion in the media recently is provided by Hod Lipson and Jordan Pollack's robots that were automatically produced (Lipson and Pollack, 2000). They decided that the morphology would consist of rods to which different types of joints could be attached. Rods can, for example, be parameterized as length, diameter, and material constants etc., thus limiting the space of possible shapes dramatically, but then the search space, even though it is still large, becomes manageable. While this example is impressive, it still implies a strong designer bias.

A more general and the more natural approach, is to not encode directly the structure of the organism in the genome but instead to encode the developmental processes. For example, it is not possible to encode the structure of the human brain in the genome because in the latter there is not enough information content. Once again, nature can be taken as a source of inspiration.

An illustration of how biological development might be modeled is given in Eggenberger (1997, 1999) who succeeded in growing 3-D shapes based on the Artificial Evolutionary System (AES). The AES implements the biological mechanisms of gene-based cell-to-cell communication. The final organism corresponds to an attractor of a highly complex dynamical system. Although these sorts of models are only in their initial stages, they will become increasingly important if we are to understand the principles of "ecological balance" and of agent design. The attempt behind the AES is to evolve entire organisms from one cell. The search space is, again, extremely vast and there is little hope that anything will converge within reasonable time. Natural systems have evolved mechanisms to impose constraints so that, for example, groups of genes couple together for certain periods of time during development (e.g. the hox genes) which enables, for example, the coordinated growth of organs or limbs.

If we have the mechanisms for co-evolving entire organisms' morphology, materials, and control, we have a powerful tool at hand by which we can explore the space of possible designs and thus "ecological balance". At the moment this is only possible in simulation; the experiments with artificial systems that can grow physically are only in their very initial stages. One way to get around this problem, at least to some extent, is on the one hand to have a good simulator that models the physics of an evolved individual and its interactions with the real world, on the other to have rapid robot building kits that enable the researchers to quickly build a robot to test some individuals in the real world. But even if done in simulation, evolving an organism from scratch is a grand challenge as well.

One of the problems with the examples and ideas presented in this paper is that they are mostly qualitative. Clearly, more quantitative statements will be required to make the story more compelling. But we hope that researchers will take up the challenges posed by embodiment.

6. CONCLUSIONS

We have tried to outline the history and the future of artificial intelligence, from its initial form as an algorithmic – cognitivistic – discipline all the way to its current embodied form. The big and frequently asked question is whether this embodied approach will indeed succeed to achieve in a bottom-up manner, higher levels of intelligence that go beyond direct sensory-motor tasks. We feel that this is indeed the case: It has been suggested, for example, that even abstract relationships like transitivity can be explained as emergent from embodied interactions with the environment (Linda Smith, pers. comm.). A similar argument has been made for mathematical concepts (e.g. Núñez and Lakoff, 1998). The jury is still out on whether this is a sound intuition or will turn out to be flawed; all we can do at the moment is outline a research program. But because embodiment provides a new perspective and many ideas for empirical studies on natural and artificial systems, as well as for new kinds of agents, we are optimistic that we can achieve a better understanding of intelligence in the future.

Acknowledgment

I would like to thank Reinhard Wilhelm for suggesting the topic of this paper and to the members of the Artificial Intelligence Laboratory for many discussions, in particular Lukas Lichtensteiger and Gabriel Gómez.

LITERATURE

[1] Brooks, R. A. (1991a). Intelligence without representation. *Artificial Intelligence*, 47, 139-160.

[2] Brooks, R. A. (1991b). Intelligence without reason. *Proceedings International Joint Conference on Artificial Intelligence-91*, 569-595.

[3] Brooks, R. A., Breazeal, C., Marjanovic, M., Scassellati, B., and Williamson, M. (1999). The Cog project: building a humanoid robot. http://www.ai.mit.edu/projects/cog/.

[4] Clancey, W.J. (1997). *Situated cognition. On human knowledge and computer representations*. Cambridge, UK: Cambridge University Press.

[5] Clark, A. (1998). *Being there: Putting brain, body, and world together again*. Cambridge, Mass.: MIT Press.

[6] Dewey, J. (1896). The reflex arc in psychology. *Psychol. Review*, **3**, 1981, 357-370. Reprinted in J.J. McDermott (eds.): *The Philosophy of John Dewey*. Chicago, IL: The University of Chicago Press, 136-148.

[7] Edelman, G.E. (1987). *Neural Darwinism. The theory of neuronal group selection.* New York: Basic Books.

[8] Eggenberger, P. (1997). Evolving morphologies of simulated 3d organisms based on differential gene expression. In: P. Husbands, and I. Harvey (eds.). *Proc. of the 4th European Conference on Artificial Life.* Cambridge, Mass.: MIT Press.

[9] Eggenberger, P. (1999). *Evolution of three-dimensional, artificial organisms: simulations of developmental processes.* Unpublished PhD Dissertation, Medical Faculty, University of Zurich, Switzerland.

[10] Elman, J.L., Bates, E.A., Johnson, M.H., Karmiloff-Smith, A., Parisi, D., and Plunkett, K. (1996). *Rethinking innateness. A connectionist perspective on development.* Cambridge, Mass.: MIT Press.

[11] Ferrari, F., Nielsen, P.Q.J., and Sandini, G. (1995). Space variant imaging. *Sensor Review*, **15**, 17-20.

[12] Franceschini, N., Pichon, J.M., and Blanes, C. (1992). From insect vision to robot vision. *Phil. Trans. R. Soc. Lond. B*, **337**, 283-294.

[13] Hara, F., and Pfeifer, R. (2000). On the relation among morphology, material and control in morpho-functional machines. *In Proc. SAB 2000*, 33-40.

[14] Kornbluh, R. D., Pelrine, R., Eckerle, J., and Joseph, J. (1998). Electrostrictive polymer artificial muscle actuators. In *Proc. 1998 IEEE Int. Conf. on Robotics and Automation.* New York, N.Y.: IEEE, 2147-2154.

[15] Kruschke, J.K. (1992). ALCOVE: An exemplar-based connectionist model of category learning. *Psychological Review*, **99**, 22-44.

[16] Lipson, H., and Pollack J. B. (2000), Automatic design and manufacture of artificial life forms. *Nature*, **406**, 974-978.

[17] Lichtensteiger, L., and Eggenberger, P. (1999). Evolving the morphology of a compound eye on a robot. In: Proc. of Eurobot'99.

[18] Maris, M., and te Boekhorst, R. (1996). Exploiting physical constraints: heap formation through behavioral error in a group of robots. In Proc. IROS'96, IEEE/RSJ International Conference on Intelligent Robots and Systems, 1655—1660.

[19] McGeer, T. (1990a). Passive dynamic walking. *Int. Journal of Robotics Research*, **9**, 62-82.

[20] McGeer, T. (1990b). Passive walking with knees. Proc. of the IEEE Conference on Robotics and Automation, 2, 1640-1645.

[21] Metta, G., Sandini, G., and Konczak, J. (1998). A developmental approach to sensori-motor coordination in artificial systems. In *Proc. of IEEE Conference on System, Man and Cybernetics*, San Diego (USA), 11-14.

[22] Núñez, R.E., and Lakoff, G. (1998). What did Weierstrass really define? The cognitive structure of natural and ε-δ continuity. *Mathematical Cognition*, **4**, 85-101.

[23] Pfeifer, R. (1996). Building "Fungus Eaters": Design principles of autonomous agents. In P. Maes, M. Mataric, J.-A. Meyer, J. Pollack, and S.W. Wilson (eds.): *From Animals to Animats. Proc. of the 4th Int. Conf. on Simulation of Adaptive Behavior*. Cambridge, Mass.: A Bradford Book, MIT Press, 3-12.

[24] Pfeifer, R. (1999). Dynamics, morphology, and materials in the emergence of cognition. *Proc. KI-99, Lecture Notes in Computer Science*. Berlin: Springer, 27-44.

[25] Pfeifer, R. (2000). On the role of morphology and materials in adaptive behavior. *Proc. SAB 2000*, 23-32.

[26] Pfeifer, R., and Scheier, C. (1997). Sensory-motor coordination: the metaphor and beyond. *Robotics and Autonomous Systems*, **20**, 157-178.

[27] Pfeifer, R., and Scheier, C. (1998). Representation in natural and artificial agents: an embodied cognitive science perspective. *Zeitschrift für Naturforschung*, **53c**, 480-503.

[28] Pfeifer, R., and Scheier, C. (1999). *Understanding intelligence. Cambridge*, Mass.: MIT Press.

[29] Rechenberg, I. (1973). *Evolutionsstrategie: Optimierung technischer Systeme nach Prinzipien der biologischen Evolution*. Stuttgart: Frommann-Holzboog.

[30] Rizzolatti, G., Fogassi, L., and Gallese, V. (2000). Cortical mechanisms subserving object grasping and action recognition: A new view of the cortical motor functions. In: M.S. Gazzaniga (ed.). *The new cognitive neurosciences.* Cambridge, Mass.: MIT Press.

[31] Scheier, C., Pfeifer, R., and Kuniyoshi, Y. (1998). Embedded neural networks: exploiting constraints. *Neural Networks*, **11**, 1551-1569.

[32] Shahinpoor, M., Bar-Cohen, Y., Simpson, J.O., and Smith, J. (2000). Ionic Polymer-Metal Composites (IPMC) as biomimetic sensors, actuators & artificial muscles- A review. http://www.unm.edu/~amri/paper.html

[33] Simon, H. A. (1969). *The sciences of the artificial (2nd ed.)*. Cambridge, MA: MIT Press.

[34] Thelen, E. and Smith, L. (1994). *A dynamic systems approach to the development of cognition and action.* Cambridge, Mass.: MIT Press, Bradford Books.

[35] Toepfer, C., Wende, M., Baratoff, G., and Neumann, H. (1998). Robot navigation by combining central and peripheral optical flow detection on a space-variant map. *Proc. Fourteenth Int. Conf. on Pattern Recognition.* Los Alamitos, CA: IEEE Computer Society, 1804-1807.

[36] Vinkhuyzen, E. (1998). Expert systems in practice. Unpublished PhD Dissertation, University of Zurich.

[37] Wingorad, T., and Flores, F. (1986). Understanding computers and cognition. Reading, Mass.: Addison-Wesley.