Proceedings of the 2nd International Symposium
on Adaptive Motion of Animals and Machines,
Kyoto, March.4-8, 2003

Jumping, walking, dancing, reaching: moving into the future Design principles for adaptive motion*

Rolf Pfeifer
Artificial Intelligence Laboratory
Deptartment of Information Technology, University of Zurich
Andreasstrasse 15, CH-8050 Zurich, Switzerland
pfeifer@ifi.unizh.ch, phone: +41 1 63 5 4320/31, fax: +41 1 635 68 09,
http://www.ifi.unizh.ch/~pfeifer

Paper prepared for the AMAM Conference, Kyoto, 3-8 March 2003

Abstract

Designing for adaptive motion is still largely considered an art. In recent years, we have been developing a set of heuristics or design principles, that on the one hand capture theoretical insights about adaptive systems, and on the other provide guidance in actually designing and building adaptive systems. In this paper we discuss, in particular, the principle of "ecological balance" which is about the relation between morphology, materials, and control. As we will argue, artificial evolution together with morphogenesis is not only "nice to have" but turns out to be a necessary design tool for adaptive motion.

1. Introduction

The field of adaptive systems, as loosely characterized by conferences such as SAB (Simulation of Adaptive Behavior) or AMAM (Adaptive Motion in Animals and Machines), Artificial Life, etc., is very heterogeneous and there is a definite lack of consensus on the theoretical foundations. As a consequence, agent design is – typically – performed in an ad hoc and intuitive way. Although there have been some attempts at elaborating principles, general agreement is still lacking. In addition, much of the work on designing adaptive systems is focused on software, i.e. the programming of the robots. But what we are really interested in is not so much the programming aspects, but designing entire systems. The research conducted in our laboratory, but also by many others, has demonstrated that often, better, cheaper, more robust and adaptive agents can be developed if the entire agent is the design target rather than the program only. This implies taking embodiment into account and going beyond the programming level proper. Therefore we prefer to use the term "engineering agents for adaptive motion" rather than "programming agents".

If this idea of engineering agents is the goal, the question arises what form the theory should have, i.e. how the experience gained so far can be captured in a concise scientific way. The obvious candidate is the mathematical theory of dynamical systems, and there seem to be many indications that ultimately this may be the tool of choice for formulating a theory of adaptive behavior. For the time being, it seems that progress over the last few years in the field has been slow, and we may be well-advised to search for an intermediate solution. The form of design principles seems well-suited for a number of reasons. First, at least at the moment, there don't seem to be any real alternatives. The information processing paradigm, another potential candidate, has proven ill-suited to come to grips with natural, adaptive forms of intelligence. Second, because of the unfinished status of the theory, a set of principles is flexible and can be dynamically changed and extended. Third, design principles represent heuristics for actually building systems. In this sense, they instantiate the synthetic methodology (see below). And fourth, evolution can also be seen as a designer, a "blind one" perhaps, but an extremely powerful one. We hope to convince the reader that this is a good idea, and that some will take it up, modify the principles, add new ones, and try to make the entire set more comprehensive and coherent. The response so far has been highly encouraging and researchers as well as educated lay people seem to be able to relate to these principles very easily.

Although most of the literature is still about programming, some of the research explicitly deals with complete agent design and includes aspects of morphology (e.g. Bongard,

* This is a modified version of a paper published in the proceedings of the workshop "The legacy of Grey Walter", Bristol, UK, 14-16 Aug., 2002.

2002; Bongard and Pfeifer, 2001; Hara and Pfeifer, 2000a; Lipson and Pollack, 1999; Pfeifer, 1996; Pfeifer and Scheier, 1999; Pfeifer, 2003; Sims, 1994a, b). Our own approach over the last six years or so has been to try and systematize the insights gained in the fields of adaptive behavior and adaptive motion by incorporating ideas from biology, psychology, neuroscience, engineering, and artificial intelligence into a set of design principles, as argued above; they form the main topic of this paper.

A first version of the design principles was published at the 1996 conference on Simulation of Adaptive Behavior (SAB 1996, "From Animals to Animats") (Pfeifer, 1996). A more elaborate version has been published in the book "Understanding Intelligence" (Pfeifer and Scheier, 1999). More recently, some principles have been extended to incorporate ideas on the relation between morphology, materials, and control (Ishiguro et al., this volume; Hara and Pfeifer, 2000a; Pfeifer, 2003).

We start by giving a very short overview of the principles. We then pick out and discuss in detail "ecological balance" and provide a number of examples for illustration. We then show how artificial evolution together with morphogenesis can be employed to design ecologically balanced systems. It is clear that these considerations are only applicable to embodied systems.

This is not a technical paper but a conceptual one. The goal is to provide a framework within which technical research can be conducted that takes into account the most recent insights in the field. In our argumentation we will resort to research conducted in our own laboratory but also to research performed in the community at large. In this sense, the paper has somewhat of a tutorial and overview flavor and should be viewed as such.

2. Design principles: Overview

There are different types of design principles: Some are concerned with the general "philosophy" of the approach. We call them "design procedure principles", as they do not directly pertain to the design of the agents but more to the way of proceeding. Another set of principles is concerned more with the actual design of the agent. We use the qualifier "more" to express the fact that we are often not designing the agent directly but rather the initial conditions and the learning and developmental processes or the evolutionary mechanisms and the encoding in the genome as we will elaborate later. The current overview will, for reasons of space, be very brief; a more extended version is in preparation (Pfeifer and Glatzeder, in preparation).

Table 1: Overview of the design principles

Number	Name	Description
		Design procedure principles
P-Princ. 1	Synthetic methodology	Understanding by building
P-Princ 2	Emergence	Systems designed for emergence are more adaptive
P-Princ 3	Diversity-compliance	Tradeoff between exploiting the givens and generating diversity solved in interesting ways
P-Princ 4	Time perspectives	Three perspectives required: "Here and now", ontogenetic, phylogenetic
P-Princ 5	Frame-of-reference	Three aspects must be distinguished: perspective, behavior vs. mechanisms, complexity
		Agent design principles
A-Princ 1	Three constituents	Task environment (ecological niche, tasks), and agent must always be taken into account
A-Princ 2	Complete agent	Embodied, autonomous, self-sufficient, situated agents are of interest
A-Princ 3	Parallel, loosely coupled procs	Parallel, asynchronous, partly autonomous procs, largely coupled through interaction with environment
A-Princ 4	Sensory-motor coordination	Behavior sensory-motor coordinated with respect to target; self-generated sensory stimulation
A-Princ 5	Cheap design	Exploitation of niche and interaction; parsimony
A-Princ 6	Redundancy	Partial overlap of functionality based on different physical processes
A-Princ 7	Ecological balance	Balance in complexity of sensory, motor, and neural systems: task distribution between morphology, materials, and control
A-Princ 8	Value	Driving forces;

		developmental mechanisms; self-organization

P-Princ 1**:** ***The synthetic methodology principle***. The synthetic methodology, "understanding by building", implies on the one hand constructing a model – computer simulation or robot – of some phenomenon of interest (e.g. how an insect walks, how a monkey is grasping a banana, or how we recognize a face in a crowd). On the other we want to abstract general principles (some examples are given below). The term "synthetic methodology" was adopted from Braitenberg's seminal book "Vehicles: Experiments in synthetic psychology" (Braitenberg, 1984).

P-Princ 2: The principle of emergence. If we are interested in designing adaptive systems we should aim for emergence. The term emergence is controversial, but we use it in a very pragmatic way, in the sense of not being preprogrammed. When designing for emergence, the final structure of the agent is the result of the history of its interaction with the – simulated or real world – environment. Strictly speaking, behavior is always emergent, as it cannot be reduced to internal mechanism only; it is always the result of a system-environment interaction. In this sense, emergence is not all or none, but a matter of degree: the further removed from the actual behavior the designer commitments are made, the more we call the resulting behavior emergent. Systems designed for emergence tend to be more adaptive and robust.

P-Princ 3: The diversity-compliance principle. Intelligent agents are characterized by the fact that they are on the one hand exploiting the specifics of the ecological niche and on the other by behavioral diversity. In a conversation I have to comply with the rules of grammar of the particular language, and then I have to react to what the other individual says, and depending on that, I have to say something different. Always uttering one and the same sentence irrespective of what the other is saying would not demonstrate great behavioral diversity.

P-Princ 4: The time perspectives principle. A comprehensive explanation of behavior of any system must incorporate at least three perspectives: (a) state-oriented, the "here and now", (b) learning and development, the ontogenetic view, and (c) evolutionary, the phylogenetic perspective. The fact that these perspectives are adopted by no means implies that they are separate. On the contrary, they are tightly intertwined, but it is useful to tease them apart for the purpose of scientific investigation. Note the connection to the principle of emergence: If a time perspective can be explained as being emergent from another, we have a deeper kind of explanation. For example, if the "here and now" can be explained as being emergent from the ontogenetic one, this constitutes progress.

P-Princ 5: The frame-of-reference principle. There are three aspects to distinguish whenever designing an agent: (a) the perspective, i.e. are we talking about the world from the agent's perspective, the one of the observer, or the designer; (b) behavior is not reducible to internal mechanism; trying to do that would constitute a category error; and (c) apparently complex behavior of an agent does not imply complexity of the underlying mechanism. Although it seems obvious that the world "looks" very different to a robot than to a human because the robot has completely different sensory systems than a human, this fact is surprisingly often ignored. Second, behavior cannot be completely programmed, but is always the result of a system-environment interaction. Again, it is surprising how often this obvious fact is ignored even by roboticists. And third, the complexity of the environment plays an essential role in the behavior and thus in the ways in which this complexity is perceived by an observer. Thus, the behavioral complexity cannot be attributed to the agent alone, but to the agent-environment interaction.

A-Princ 1: The three-constituents principle. This very often neglected principle states that whenever designing an agent we have to consider three components. (a) the definition of the ecological niche (the environment), (b) the desired behaviors and tasks, and (c) the agent itself. The main point of this principle is that it would be a fundamental mistake to design the agent in isolation. This is particularly important because much can be gained by exploiting the physical and social environment.

A-Princ 2: The complete agent principle. The agents of interest are autonomous, self-sufficient, embodied and situated. This view, although extremely powerful and obvious, is not very often considered explicitly.

A-Princ 3: The principle of parallel, loosely coupled processes. Intelligence is emergent from an agent-environment interaction based on a large number of parallel, loosely coupled processes that run asynchronously and are connected to the agent's sensory-motor apparatus. The term "loosely coupled" is used in contrast to hierarchically coupled processes where there is a program calling a subroutine and the calling program has to wait for the subroutine to complete its task before it can continue. In that sense, this hierarchical control corresponds to very strong coupling. However, on a complete agent, there can be a very strong coupling of processes by the fact that the system is embodied: two joints coupled by a physical link (bones) are very strongly coupled as well. "Loosely coupled" also refers to the coupling through the interaction with the environment.

A-Princ 4: The principle of sensory-motor coordination. All intelligent behavior (e.g. perception, categorization, memory) is to be conceived as a sensory-motor coordination. This sensory-motor coordination, in addition to enabling the agent to interact efficiently with the environment, serves the purpose of structuring its sensory input. One of the powerful implications is that the problem of categorization in the real world is greatly simplified through the interaction with the real world because the latter supports the generation of "good" patterns of sensory stimulation, "good" meaning correlated, and stationary (at least for a short period of time).

A-Princ 5: The principle of cheap design. Designs must be parsimonious, and exploit the physics and the constraints of the ecological niche. This principle is related to the diversity compliance principle in that it implies, for example, compliance with the laws of physics. A trivial example are robots with wheels that exploit the fact that the ground is mostly flat.

A-Princ 6: The redundancy principle. Agents should be designed such that there is an overlap of functionality of the different subsystems. Examples are sensory systems where, for example, the visual and the haptic systems both deliver spatial information, but they are based on different physical processes (electromagnetic waves vs. mechanical touch). Merely duplicating components does not lead to very interesting redundancy; the partial overlap of functionality and the different physical processes are essential. Note that redundancy is required for diversity of behavior and to make a system adaptive. If there is a haptic system in addition to the visual one, the agent can also function in complete dark, whereas one with 10 cameras it ceases to function if the light goes out.

A-Princ 7: The principle of ecological balance. This principle consists of two parts, the first one concerns the relation between the sensory system, the motor system, and the neural control. The "complexity" of the agent has to match the complexity of the task environment, in particular: given a certain task environment, there has to be a match in the complexity of the sensory, motor, and neural system. The second is about the relation between morphology, materials, and control: Given a particular task environment, there is a certain balance or task distribution between morphology, materials, and control (for references to both ideas, see, e.g. Hara and Pfeifer, 2000a; Pfeifer, 1996; Pfeifer, 1999, 2000; Pfeifer and Scheier, 1999). Often, if the morphology and the materials are right, control will be much cheaper. Because we are dealing with embodied systems, there will be two dynamics, the physical one or body dynamics and the control or neural dynamics. There is the deep and important question of how the two can be coupled in optimal ways. The research initiated by Ishiguro and his colleagues (e.g. Ishiguro et al., 2003) promises deep and important pertinent insights. We will be giving examples of this principle later in the paper.

A-Princ 8: The value principle. This principle is, in essence, about motivation. It is about why the agent does anything in the first place. Moreover, a value system tells the agent whether an action was good or bad, and depending on the result, the probability of repetition of an action will be increased or decreased. Because of the unknowns in the real world, learning must be based on mechanisms of self-organization. There is a frame-of-reference issue in that values can be implicit or explicit. If an agent is equipped, say, with neural networks for Hebbian learning, we can, as outside observers, say that this constitutes value to the agent because in this way it can learn correlations, certainly a useful thing. If as a result of a particular action, a particular internal signal – neural or hormonal – is generated that modulates learning, we talk about an explicit value system (we adopted the term from Edelman, 1987). The issue of value systems is central to agent design and must be somehow resolved. However, it seems that to date no generally accepted solutions have been developed. Research on artificial motivation and emotion, is highly relevant in this context (e.g. Breazeal, 2002; Manzotti, 2000; Picard, 1997; Pfeifer, 2000b). As this is not the central topic of this paper, we will not further elaborate on this issue.

Although it does capture some of the essential characteristics of adaptive systems, this set is by no means complete. A set of principles for designing evolutionary systems and collective systems, are currently under development.

As mentioned earlier, all these principles only hold for embodied systems. In this paper, we focus on the principle of ecological balance which is at the heart of embodiment.

3. Information theoretic implications of embodiment

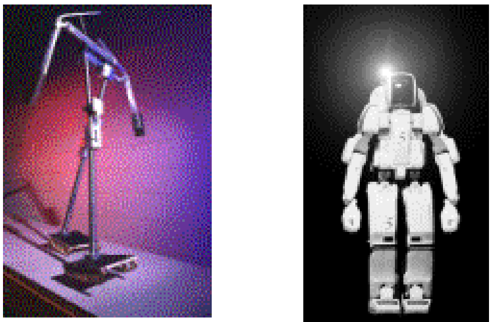
There is a trivial meaning of embodiment namely that "intelligence requires a body". In this sense, anyone using robots for his or her research is doing embodied artificial intelligence. It is also obvious that if we are dealing with a physical agent, we have to take into account gravity, friction, torques, inertia, energy dissipation, etc. However, there is a non-trivial meaning of embodiment, namely that there is a tight interplay between the physical and the information theoretic aspects of an agent. The design principles all directly or indirectly refer to this issue, but some focus specifically on this interplay, i.e. the principle of sensory-motor coordination where through the embodied interaction with the environment sensory-motor patterns are induced, the principle of cheap design where the proper embodiment leads to simpler and more robust control, the redundancy principle which states that proper choice and positioning of sensors leads to robust behavior, and the principle of ecological balance that explicitly capitalizes on the relation between morphology, materials, and neural control. For the purpose of illustration we will capitalize on the latter in this paper. We proceed by presenting a number of case studies illustrating the application of these principles to designing adaptive motion.

A short note on terminology is in place, here. We talked about information theoretic implications of embodiment. What we mean is the effect of morphology on neural processing, or better, the interplay between the two. The important point is that the implications are not only of a purely physical nature.

In previous papers we have investigated in detail the effect of changing sensor morphology on neural processing (e.g. Lichtensteiger and Eggenberger, 1999; Maris and te Boekhorst, 1996; Pfeifer, 2000a, b; Pfeifer and Scheier, 1999). In this paper we focus on the motor system.

The passive dynamic walker

The passive dynamic walker which goes back to McGeer (1990a, b), illustrated in figure 1a, is a robot (or, if you like, a mechanical device) capable of walking down an incline without any actuation and without control. In other words, there are no motors and there is no microprocessor 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 and there is an intrinsic naturalness to it. However, 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. Energy-efficiency is achieved because in this approach the robot is – loosely speaking – operated near one of its Eigenfrequencies. To make this work, a lot of attention was devoted to morphology and materials. For example, the robot is equipped with wide feet of a particular shape to guide lateral motion, soft heels to reduce instability at heel strike, counter-swinging arms to negate yaw induced by leg swinging, and lateral-swinging arms to stabilize side-to-side lean (Collins et al., 2001).



a. b.

Figure 1. Two approaches to robot building. (a) The passive dynamic walker by Steve Collins (Collins et al., 2001), (b) the Honda robot Asimo.

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 different types 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 – or the neural processing, if you like – 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 in principle that prevents the exploitation of its passive dynamics. In human walking for example – and humans are certainly highly complex systems – the forward swing of the leg is largely passive as well. Of course, the Honda robot can do many things like walking up and down the stairs, pushing a cart, opening a door, etc., whereas the ecological niche of the passive dynamic walker is confined to inclines of a particular angle. In other words, the ecological niche of Asimo is considerably larger.

In terms of the design principles, this case study illustrates the principles of cheap design and ecological balance. The passive dynamic walker fully exploits the fact that it is always put on inclines that provide its energy source and generates the proper dynamics for walking. Loosely speaking, we can also say that the control tasks, the neural processing, is taken over by having the proper morphology and the right materials. In fact, the neural processing reduces to zero. At the same time, energy efficiency is achieved. However, if anything is changed, e.g. the angle of the incline, the agent ceases to function. This is the trade-off of cheap design. In order to make it adaptive, we would have to add redundancy. There is no contradiction between cheap design and redundancy: even highly redundant systems such as humans exploit the givens.

Even though the passive dynamic walker is an artificial system (and a very simple one), it has a very natural feel to it. The term "natural" not only applies to biological systems, but artificial systems also have their intrinsic natural dynamics. Perhaps the natural feel comes from the exploitation of the dynamics, e.g. the passive swing of the leg.

In conclusion, as suggested by the principle of ecological balance, 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.

Muscles – control from materials: reaching and grasping

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. Reaching for and 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. Or put differently, the morphology (the anatomy), and the materials provide physical constraints that make the control problem much easier – at least for the standard kinds of movements. The main task of the brain, if you like, is to set the material properties of the muscles, the spring constants. Once these constraints are given, the control task is much simpler.

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 2. 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.

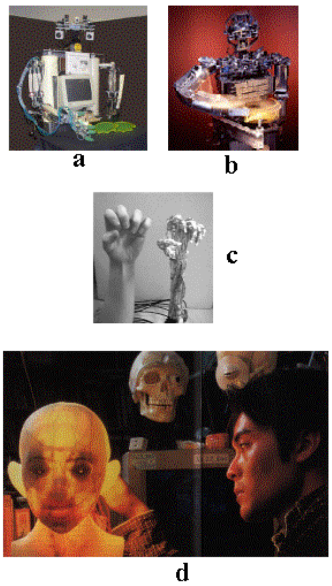


Figure 2. Robots with artificial muscles. (a) The service robot ISAC by Peters (Vanderbilt University) driven by McKibben pneumatic actuators. (b) The artificial hand by Lee and Shimoyama (University of Tokyo), driven by pneumatic actuators. (c) The humanoid robot Cog by Rodney Brooks (MIT AI Laboratory), driven by series-elastic actuators. (d) The "Face Robot" by Kobayashi, Hara, and Iida (Science University of Tokyo), driven by shape-memory alloys.

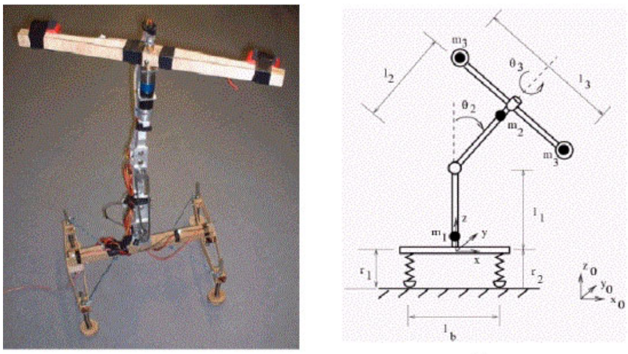
Facial expressions also provide an interesting illustration for the point to be made here. If the facial tissue has the right sorts of material properties in terms of elasticity, deformability, stiffness, etc., the neural control for the facial expressions becomes much simpler. For example, for smiling, although it involves the entire face, the actuation is very simple: the "complexity" is added by the tissue properties. Another highly desirable property that one gets for free if using the right kinds of artificial muscles is passive compliance: if an arm, for example, encounters resistance it will yield elastically rather than pushing harder. In the case of the pneumatic actuators this is due to the elastic properties of the rubber tubes.

The important point here is that we are not simply replacing one type of actuator, an electrical motor, by a different one. This would not be very interesting. The point is that the new type of actuator – e.g. a pneumatic one – has intrinsic physical properties such as elasticity and damping, that can be exploited by the neural control.

The dancing robot Stumpy – a synthesis

Recently, there has been an increased interest in applying and further investigating these ideas to the construction of robots. An illustrative example is the walking and hopping robot Stumpy (Paul et al, in press-a, b) (figure 3). Stumpy's lower body is made of an inverted "T" mounted on wide springy feet. The upper body is an upright "T" connected to the lower body by a rotary joint, the "waist" joint, providing one degree of freedom in the frontal plane. The horizontal beam on the top is weighted on the ends to increase its moment of inertia. It is connected to the vertical beam by a second rotary joint, providing one rotational degree of freedom, in the plane normal to the vertical beam, the "shoulder" joint. Stumpy's vertical axis is made of aluminum, while both its horizontal axes and feet are made of oak wood.

Although Stumpy has no real legs or feet, it can locomote in many interesting ways: it can move forward in a straight or curved line, it has different gait patterns, it can move sideways, and it can turn on the spot. Interestingly, this can all be achieved by actuating only two joints with one degree of freedom. In other words, control is extremely simple – the robot is virtually "brainless". The reason this works is because the dynamics, given by its morphology and its materials (elastic, spring-like materials, surface properties of the feet), is exploited in clever ways. There is a delicate interplay of momentum exerted on the feet by moving the two joints in particular ways (for more detail, see Paul et al., 2002a, b).



a b

Figure 3. The dancing, walking, and hopping robot Stumpy. (a) Photograph of the robot. (b) Schematic drawing (details, see text).

Let us briefly summarize the ideas concerning ecological balance, i.e. 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 while maintaining a certain level behavioral diversity. Tird, materials have intrinsic control properties. And fourth, because ecological balance is exploited, Stumpy displays a surprisingly diverse behavior (dancing walking, and hopping in different ways). In this sense, Stumpy also illustrates the diversity-compliance principle: on the one hand, it exploits the physical dynamics in interesting ways and on the other it displays high diversity.

In section 2 we postulated a set of design principles for adaptive motion. The principle of ecological balance, for example, tells us that given a particular task environment, there is an optimal task distribution between morphology, materials, and control. The principle of emergence asks the question of how a particular "balance" has emerged, how it has come about. In the study of biological systems, we can speculate about this question. However, there is a possibility of systematically investigating this balance, namely artificial evolution and morphogenesis. Pertinent experiments promise a deeper understanding of these relationships. The remainder of this paper will be devoted to this question.

4. Exploring "ecological balance"—artificial evolution and morphogenesis

Using artificial evolution for design has a tradition in the field of evolutionary robotics. The standard approach 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 problem with including morphology and materials is that the search space which is already very large for control architectures only, 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. An example that has stirred a lot of commotion in the media 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, or in other words, the types of morphologies, 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. If we want to explore different types of morphologies, we want to introduce as little designer bias as possible. This can be done using ideas from biology, i.e. genetic regulatory networks.

The mechanics of artificial genetic regulatory networks

We provide a non-technical introduction, for details, see, e.g. Bongard and Pfeifer (2001), Bongard (2002). It should be stressed, that although this computational system is biologically inspired, it does not constitute a biological model. Rather, it is system in its own right. Also, when we use biological terminology, e.g. when we say that "concentrations of transcription factors regulate gene expression", this is meant metaphorically.

The basic idea is the following. A genetic algorithm is extended to include ontogenetic development by growing agents from genetic regulatory networks. In the example presented here, agents are tested for how far they can push a large block (which is why they are called "block pushers"). Figure 4a shows the physically realistic virtual environment. The fitness determination is a two-stage process: the agent is first grown and then evaluated in its virtual environment. Figure 4b illustrates how an agent grows from a single cell into a multicellular organism.

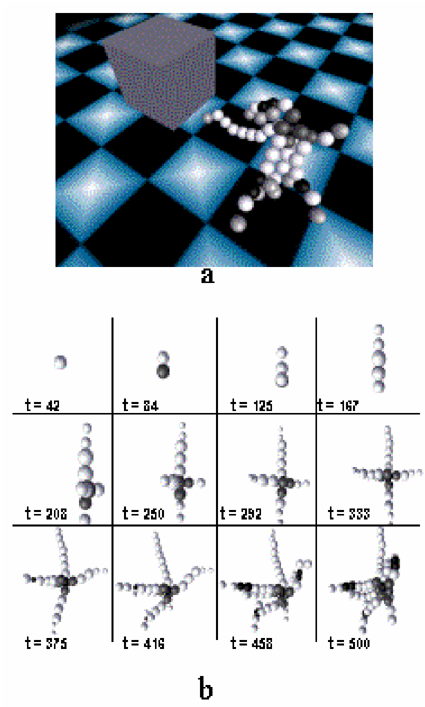


Figure 4. Examples of Bongard's "block pushers". (a) An evolved agent in its physically realistic virtual environment. (b) growth phase starting from a single cell, showing various intermediate stages (last agent after 500 time steps).

The algorithm starts with a string of randomly selected floating point numbers between 0 and 1. A scanning mechanism determines the location of the genes. Each gene consists of 6 floating point numbers which are the parameters that evolution can play with. They are explained in figure 5. There are transcription factors that only regulate the activity of other genes, there are transcription factors for morphology, and for neuronal growth. Whenever a gene is "expressed", it will diffuse a transcription factor into the cell from a certain diffusion site. The activity of this genetic regulatory network leads to particular concentrations of the transcription factors to which the cell is sensitive: whenever a concentration threshold is exceeded, an action is taken. For example, the cell may increase or decrease in size, if it gets too large, it will split, the joint angles can be varied, neurons can be inserted, connections added or deleted, structures can be duplicated, etc. The growth process begins with a single unit into which "transcription factors" are injected (which determines the primary body axis). Then it is left to the dynamics of the genetic regulatory network. The resulting phenotype is subsequently tested in the virtual environment. Over time, agents evolve that are good at pushing the block.

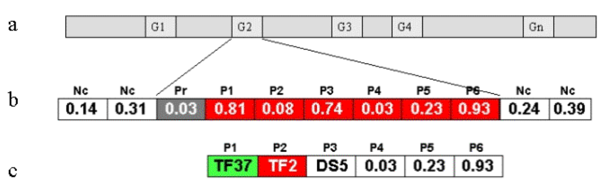


Figure 5. The mechanisms underlying the genetic regulatory networks. (a) Genes on the genome. Which regions are considered to be genes is determined by an initial scanning mechanism (values below 0.1 are taken as starting positions). (b) and (c) An example of a particular gene. Nc means "non-coding" region, Pr is a promoter site (start of gene), P1 through P6 are the parameters of the gene. P1: the transcription factor (TF) that regulates the expression of this gene [0,19]. P2: the TF the gene emits if expressed [0,42]. P3: the diffusion site, i.e. the location in the cell from which the TF is diffused. P4: the quantity of TF emitted by this gene, if expressed. P5, P6: lower and upper bounds of the concentrations within which the gene is expressed.

Emergence – the achievements of artificial evolution and morphogenesis

Although simple in their basic form, these mechanism lead to an interesting dynamics and produce fascinating results. Here are some observations: (1) Organisms early on in evolution are typically smaller than those of later generations: evolution discovers that in order to push a block of large size, it is necessary to have a large body. In other words, evolution had to manipulate morphology in order to achieve the task. (2) Evolution comes up with means of locomotion. In small creatures, these are very local reflex-like mechanisms distributed through the entire organism. Larger creatures tend to have additional tentacles that can be used to push against the block, which requires a different kind of control. Because they have been created by artificial evolution and morphogenesis, they are, in some sense, ecologically balanced (for this particular task environment). (3) There is no direct relation between genotype length and phenotypic fitness – the two are largely dissociated. (4) There is functional specialization, i.e. cells differentiate into units containing both sensors and actuators (the white colored cells in figure 4), cells that only contain sensors but no actuators (gray coloring), and cells not containing anything, only providing structural support (black coloring). (5) There is repeated structure, i.e. some combination of cells occur in slightly modified form in various places on the agent. An example from biology are fingers that are similar but differ individually. (6) Some genes specialize to become "master regulatory genes", i.e. they regulate the activity of other genes. Thus, to an outside observer, it looks as if a hierarchical structure were evolving in the regulatory network. Note that this hierarchy is emergent and results from a "flat" dynamical system. Thus, it can change at a later point in time, unlike "structural" hierarchies. Again, metaphorically speaking, artificial evolution has discovered how to manage complexity, i.e. by evolving a hierarchical organization. It is important to mention that this has all been "discovered" by simulated evolution and has not been programmed into the system. Or stated differently, it is emergent from the mechanisms of simulated evolution and genetic regulatory networks.

The work of Eggenberger (1997, 1999) is among the first to employ genetic regulatory networks to model growth processes in computational systems. He succeeded in evolving three-dimensional shapes. As in the case of

Bongard's system, the resulting shape (or organism) is emergent from a complex dynamical system.

4. Discussion and conclusions

We have argued that there is still a lack of consensus in the field of adaptive behavior on the theoretical foundations. By employing the form of design principles we have attempted to make a first step in the direction of providing a coherent framework for design. In the present form we have proposed the principles and have argued why they are plausible. The passive dynamic walker and Stumpy provide illustrations of the principles of cheap design and ecological balance.

While this is acceptable and interesting, the design principles would be much more compelling and powerful if they could be demonstrated to emerge from an evolutionary process (which is one of the messages of the principle of emergence). Using the principles of genetic regulatory networks, we have worked out methods by which entire agents can be evolved, including their morphology, their material properties, and their control systems.

There are a number of limitations of this approach that we will put on the research agenda for the coming years. One is the incorporation of interaction with the environment during ontogenetic development. Moreover, the "rewrite rules" for neuronal growth will be replaced by more biological mechanisms. Third, instead of defining a fitness function, we will turn to "open-ended evolution" where the survival of the individual is the sole criterion. This requires the definition of pertinent resources that need to be maintained. Fourth, we need to incorporate the variation of material properties into the evolutionary algorithm, so that this aspect can be studied as well. And last but not least, we need to be able to increase the complexity of our task environments which requires much higher computational power.

At the moment we are confined to 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 (e.g. gravity, impact, friction), 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 big 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 do hope that researchers will take up the challenges posed by embodiment.

Let us conclude by raising an issue that is always in the air when working with relatively simple systems (such as block pushers), the one of scalability. By scalability we mean in this context whether the methods proposed (simulated genetic regulatory networks) will be sufficiently powerful to evolve much more complex creatures capable of many behaviors in very different types of environments. This question, we believe is still open as it is not clear to what extent the real world plays an essential role in evolution, or whether simulated environments can be made sufficiently complex.

One hope is, for example, that as the environments and agents get more complex, involving not only one or a few tasks, but perhaps hundreds or thousands, we will begin to see a certain centralization of the neural substrate which in the very simple creatures is largely distributed through the entire agent.

Acknowledgments

I would like to thank the members of the Artificial Intelligence Laboratory for many discussions, in particular Josh Bongard for his patience in explaining evolution to me and Gabriel Gómez for discussing the manuscript. Credit also goes to the Swiss National Science Foundation for supporting the research presented in this paper, grant # 20-61372.00.

References

Braitenberg, V. (1984). Vehicles: Experiments in synthetic psychology. Cambridge, MA.: MIT Press.

Breazeal, C.L. (2002). *Designing Sociable Robots (Intelligent Robotics and Autonomous Agents)*. Cambridge, Mass.: MIT Press.

Bongard, J.C. (2002). Evolving modular genetic regulatory networks. In *Proc. IEEE 2002 Congress on Evolutionary Computation (CEC2002)*. MIT Press, 305-311.

Bongard, J.C., and Pfeifer, R. (2001). Repeated structure and dissociation of genotypic and phenotypic complexity in artificial ontogeny. In L. Spector et al. (eds.). *Proc. of the Sixth European Conference on Artificial Life*, 401-412.

Collins, S.H., Wisse, M., and Ruina, A. (2001). A three-dimensional passive-dynamic walking robot with two legs and knees. *The International Journal of Robotics Research*, 20, 607-615.

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

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.

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

Ishiguro, A., Ishimaru, K., Hayakawa, K., and Kawakatsu, T. (2003). Toward a "well-balanced" design: a robotic case study. How should control and body dynamics be coupled? This volume.

Hara, F., and Pfeifer, R. (2000). On the relation among morphology, material and control in morpho-functional machines. In Meyer, Berthoz, Floreano, Roitblat, and Wilson (eds.): *From Animals to Animats 6. Proceedings of the sixth International Conference on Simulation of Adaptive Behavior* 2000, 33-40.

Kornbluh, R. D., Pelrine, R., Eckerle, J., and Joseph, J. (1998). Electrostrictive polymer artificial muscle actuators. *Proceedings of the IEEE International Conference on Robotics and Automation* 1998. New York, N.Y.: IEEE, 2147-2154.

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

Lichtensteiger, L., and Eggenberger, P. (1999). Evolving the morphology of a compound eye on a robot. *Proceedings of the third European Workshop on Advanced Mobile Robots (Eurobot'99).* IEEE, Piscataway, NJ, USA; 1999; 127-34 .

Manzotti, R. (2000). Intentional robots. The design of a goal-seeking environment-driven agent. Unpublished Doctoral Dissertation, University of Genova.

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

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

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

Paul, C., Dravid, R. and F. Iida (2002a) Control of lateral bounding for a pendulum driven hopping robot. to appear in *Proceedings of the International Conference of Climbing and Walking Robots* , Paris, France (to appear)..

Paul, C., Dravid, R. and F. Iida (2002b) Design and Control of a Pendulum Driven Hopping Robot. *Proc of the IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS-2002*, Lausanne, Switzerland (to appear).

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 4. Proceedings of the fourth International. Conference on Simulation of Adaptive Behavior.* Cambridge, Mass.: A Bradford Book, MIT Press, 3-12.

Pfeifer, R. (1999). Dynamics, morphology, and materials in the emergence of cognition. In Burgard, W., Christaller, T., Cremers, A. B. (eds.): *KI-99 Advances in Artificial Intelligence. Proceedings of the 23rd Annual German Conference on Artificial Intelligence*, Bonn, Germany, 1999, Lecture Notes in Computer Science, Springer, 1701, 27-44.

Pfeifer, R. (2000a). On the role of morphology and materials in adaptive behavior. In Meyer, Berthoz, Floreano, Roitblat, and Wilson (eds.): *From Animals to Animats 6. Proceedings of the sixth International Conference on Simulation of Adaptive Behavior 2000*, 23-32.

Pfeifer, R. (2000b). On the role of embodiment in the emergence of cognition and emotion. In H. Hatano, N. Okada, and H. Tanabe (eds.). *Affective minds.* Amsterdam: Elsevier, 43-57.

Pfeifer, R. (2002). Robots as cognitive tools. *Int. Journal of Cognition and Technology*, 1, 125-143.

Pfeifer, R. (2003). Morpho-functional machines: basics and research issues. In F. Hara, and R. Pfeifer (eds.). *Morpho-functional machines: the new species.* Tokyo: Springer, 2003.

Pfeifer, R., and Glatzeder, B. (in preparation). *How the body shapes the way we think: the embodied revolution in artificial intelligence.* Cambridge, Mass.: MIT Press.

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

Picard, R.W. (1997). *Affectictive computing.* Cambridge, Mass.: MIT Press.

Reeke, N.G., Finkel, L.H., Sporns, O., and Edelman, G.A. (1998). Synthetic neural modeling: a multilevel approach to the analysis of brain complexity. In G.M. Edelman, W.E. Gall, W.M. Cowan (eds.). *Signal and Sense: local and global order in perceptual maps.* New York, N.Y.: Wiley, 607-706.

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

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

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

Sims, K. (1994a). Evolving virtual creatures. *Computer Graphics,* 28, 15-34.

Sims, K., (1994b). Evolving 3D morphology and behavior by competition. In R. Brooks, and P. Maes (eds.). *Artificial Life IV Proceedings,* 28-39.s