Group Report: Influence of Brain and Computer Design on the Performance of Natural and Artificial Organisms*

Rupert Schmidt (Rapporteur)

Wolfgang Buschlinger, Martin Egelhaaf (Moderator), Michael Koch,
Jürgen Krüger, Rainer Malaka, Rolf Pfeifer, Gerhard Vollmer (Moderator)
and Hermann Wagner

Z. Naturforsch. **53c,** 752–764 (1998); received April 20, 1998

Structure, Plasticity, Algorithms, Levels of Analysis, Isocortex

Introduction

This group report concentrates on discussions of structure-function relationships in brains. It also attempts to compare some evolutionary achievements of natural organisms with the influence of computer and robot (artificial organism) design on their performance. Here, the term *"structure"* refers to the materialistic implementation at the level of molecules, subcellular organelles, cells or networks which enables the (natural or artificial) organism to achieve a teleologically defined function. Both, in neurobiology and in robotics, the meaning of the term *"function"* depends on the particular level chosen for analysis of performance and organisation. A given structure may have only one or many functions, and a given function may be implemented by only one or by several different structures.

Physical and chemical constraints have a large influence on the structure of organisms. Obvious examples are provided by *convergent evolution* (*e.g.*, in the mole and the mole-cricket). We also see effects of physico-chemical constraints in nervous systems. Receptor cells, *e.g.*, have two parts, an outer segment involved in signal transduction and an inner segment that has all the machinery to electrically process the transduced signal. The outer segment is specific for modality: Outer segments of photoreceptors are different from outer segments of mechanoreceptors or chemoreceptors. The inner segments, however, are rather similar in all cases. Since all signals are processed similarly in the brain, the question was raised whether more specific structure-function relationships can be seen in central parts of nervous systems than observed on the level of receptors.

To cut some vistas into the jungle of intermingled problems related to structure and function of brains and computers, the group decided to discuss five topics under the following headings:

- Influence of architecture on functional performance of brains and computers,
- Implementation of plasticity,
- Nature and role of algorithms,
- Enhancement of efficacy by internal reconstruction of brains and computers,
- Limitations of brains and artefacts.

Influence of Architecture on Functional Performance of Brains and Computers

Relations between brain structure and function may be analysed on different levels. They comprise the fine-structural organisation of individual neurones, the extent of synaptic connections onto and from neurones, the relationships to neighbouring glial elements, the lay-out of neuronal nets, the arrangement of neuronal populations in cell layers and the dimensions of distinct brain nuclei. Relations between structure and function

---

* This communication is a contribution to the workshop on "Natural Organisms, Artificial Organisms, and Their Brains" at the Zentrum für interdisziplinäre Forschung (ZiF) in Bielefeld (Germany) on March 8–12, 1998.

Reprint requests to Prof. Dr. Rupert Schmidt, Universität Gießen, Biotechnology Centre (ZBB), Leihgesterner Weg 217, D-35392 Gießen, Germany.
Fax: +49 (0)641-9916509.
E-mail: Rupert.Schmidt@strz.uni-giessen.de.

0939–5075/98/0700–0752 $ 06.00

have been described of small local elements, such as the number of synapses subserving transmission of specific signals from one set of central neurones to another and the number and form of dendritic spines receiving signals from presynaptic boutons. Much is known about the molecular characteristics of neurones, about enzyme activities involved in synthesis, release and degradation of neurotransmitters, and about the number and distribution of receptor molecules regulating the conductance of ion channels in synaptic membranes by direct coupling to ionophores or via second messenger cascades.

Just as in natural brains, also in artificial systems, functionality has to be analysed on different levels. We have to distinguish between a physical level (solid state physics), an electronic circuit level, a digital level, a program level, and a model level. If we have a complete agent, a robot, we finally have to relate the internal processing to behaviour. In such analyses, the so-called "frame-of-reference" problem must be taken into account: The relationship between observer, the object that he is attempting to model and the environment must be clearly specified. Likewise, the relation between the designer, the artefact that he tries to construct and its environment is important. Behaviour (of a natural or artificial organism) is the result of a system-environment interaction. Attempts to reduce the behaviour of an organism to internal mechanisms of brain processes only, would therefore constitute a category error.

Apparently, specific relations between structure and function are more obvious in peripheral parts of nervous systems, *i.e.*, close to sensory receptors and motor effectors. In central parts of the nervous system information processing makes use of neuronal generalists and stereotype small cell assemblies embedded into larger neuronal nets, in particular in the cortex of the telencephalon. For these parts of the brain it is more difficult to elucidate specific structure-function relationships: Very little is yet known, whether and how specific functions may be reflected in a specifically adapted architecture of neurones, nuclei or neuronal circuits, and where such relations have been decribed, they appear to be weak. In some specific elements, however, a high correlation between structure and function has been reported. Some of these specific adaptations were discussed in detail:

Example 1

In general, it can be observed that brain areas involved in the processing of biologically important information are enlarged. In those insects which rely very much on vision, the visual system comprises about 50% of the entire nervous system. In the avian brain, distinct nuclei, including area X, the nucleus robustus archistrialis and the hyperstriatum ventrale, pars caudale, are used to generate song. These nuclei are only found in songbirds and parrots. They are enlarged in males and comprise more and larger neurones than in females (Nottebohm *et al.*, 1976). Further reports referred to the number of dendritic bifurcations and the distances between them. These structural characteristics are under hormonal control, not only during ontogenetic maturation, but also in the context of seasonal requirements. Female canaries implanted with testosterone propionate develop larger vocalisation areas and begin to sing. The same effect was elicited in female zebra finches, if they had been pre-treated with 17β-estradiol as chicks.

Example 2

In 1948 Jeffress postulated the existence of a specific set-up of neuronal connectivities as a structural prerequisite for the measurement of small time differences. The predicted type of architectural arrangement was indeed observed some 40 years later in the nucleus laminaris of the barn-owl (Wagner and Luksch, 1998: this issue, pp. 560–581). The barn-owl represents the position of its prey in the brain in an acoustic map of the surrounding. Here, the horizontal deviation from the mid-sagittal plane, the azimuth, is coded as interaural time difference between the signals received by the two ears. To extract the value of a given interaural time difference, the delay of neuronal signals from both ears is compared by neurones in the nuclei laminares, which are the first neurones receiving ipsi- and contralateral auditory input. These neurones have partly lost their dendritic trees. The reduction in the dendritic arborisation is more pronounced in neurones involved in the measurement of high frequencies (7 kHz), than in neurones tuned to lower frequencies (1–2 kHz).

Example 3

In their mating behaviour, male flies chase females in virtuous flight manoeuvres in order to catch them. This conspicuous behaviour is mediated by neuronal circuits which, at least to a large extent, only exist in male flies. Already at the level of the retina and the first processing stages of the visual system, sex-specific specialisations have been found. Moreover, in the third visual neuropil, neurones have been characterised which are not present in females at all. These sex-specific specialisations reflect a high degree of correlation between structure and function (Egelhaaf *et al.*, 1998: this issue, pp. 582–592).

Example 4

A further example of structure-function relationships is the fine-structural specialisation of dendritic spines involved in the reception of neuronal signals in the nucleus arcuatus hypothalami: The mammalian brain develops primarily female characteristics, unless exposed to testosterone during a critical phase around parturition. When the (male) foetus produces testosterone, this hormone is biochemically converted (aromatised) by neuronal enzymes to estradiol, which reduces the number of dendritic spines on neurones in the nucleus arcuatus. Functional female characteristics of the mature mammalian brain are only preserved if the full amount of dendritic spines is expressed on these neurones. The neuroendocrine cells of the nucleus arcuatus release LHRH in the pituitary. In order to prevent a male differentiation of a genetically female brain, the mother's estradiol (a steroid passing via the placenta and through the foetus' blood brain barrier) is masked in the blood of the neonatal female rat by a special estradiol-binding protein, the α-fetoprotein. If this protein is not provided in a sufficiently high concentration, the individual will not retain its female organisation nor the ability to produce the hormones responsible for ovulation, in spite of its female phenotype.

Example 5

Long-term potentiation is a well known example of functional neuronal plasticity elicited at distinct sets of synapses, in particular in the mammalian hippocampus. The electrophysiological changes at these synapses are observed after short trains of tetanic stimulation, and they can persist from several hours to weeks. Long-term potentiation has been interpreted as a model system or even a partial mechanism of memory formation. Induction of long-term potentiation involves a sequel of partly understood biochemical reactions evoked by the neurotransmitter glutamate and by modulators at NMDA-receptors. Repeatedly it was reported in the literature that, some hours after induction of long-term potentiation, changes can be observed in the number and fine-structural characteristics of dendritic spines (spine length, bifurcations etc.) of those postsynaptic neurones which exhibit the long lasting functional adaptation.

Example 6

Teleosts (bony fish) lack the high degree of telencephalic differentiation typical of higher vertebrates, mammals in particular. In the teleost mesencephalon, sensory inputs from different modalities project on distinct classes of neurones, each situated in a particular cell layer of the "optic" tectum. Later in evolution, such layers have been rearranged to form distinct nuclei. As it is known on which tectal cell population each set of afferent axons projects and because some of the different types of synapses can be identified with the electron microscope, the tectum of teleosts offers an unique possibility to study modality-specific functional adaptations at the level of the synapses involved. Following an active shock avoidance conditioning in goldfish, *e.g.*, deposition of a functionally important cell adhesion molecule was observed at synapses of the type I neurones in the optic tectum, *i.e.*, at neurones involved in the integration of those excitations representing the stimuli used in the preceding conditioning (Schmidt, 1995).

Example 7

Specialised neurones in fish (Mauthner cells) mediate their escape response. These neurones have well myelinated axons to speed-up signal conduction to muscles of the tail, and they mainly use electrical transmission at mixed electrical/chemical synapses, thus avoiding delay of information processing caused otherwise by release, diffusion and binding of chemical neurotransmitters.

Quite obviously, for the more specialised brain regions, functions are more evident, whereas the cortex of the telencephalon can potentially "do anything" – it resembles a universal purpose device. Architectural differences, as far as they exist in the cerebral cortex at all, may be attributed to different degrees of connectivity, reflecting the specific demand for neuronal interaction partners. Neurones in layer V, *e.g.*, project to the brain stem and the spinal cord. Where numerous fibres have to be sent to these targets, thickness of layer V is increased. Cortical input-output transformations cannot clearly be recognised if descriptions in terms of outer-world phenomena are chosen, such as seeing a visual stimulus or hearing a sound. But it is essential, to always relate neuronal processing to the actual behavioural performance. This requires precise knowledge of how the components of the brain are embedded in a physical organism. Correspondingly it is necessary to define how algorithms are *"embodied"* in a robot (cf. Pfeifer and Scheier, 1998: this issue, pp. 480–503).

The continuous advancement of computer design has little in common with the evolution of natural brains. In technical systems, the same functionality can be achieved by computer chips of rather different structures, and, moreover, computational functionalities can be implemented in software or hardware solutions. As for natural brains, it is easier to understand the function of a particular unit in the periphery of an artefact than in its central parts. Chips in central parts of computers are often very general processors used in a wide range of applications. Specialised parts enable higher velocities of computation, but usually a compromise has to be found between speed of performance and flexible applicability of a computer chip. The closer to function a particular element is, the more specialised it may be designed (this holds true for neurones as well as for flip-flop computer elements).

From an engineering perspective, the following two terms seem to be helpful: *"engineering"* and *"reverse engineering"*. Engineering works from specifications, also called functional requirements, to structure, the actual device. Reverse engineering starts with the artefact and tries to work out, what the functional requirements might have been. Reverse engineering has been applied quite successfully in computer science, *e.g.*, to derive the specification from the program code, solving the questions "What was the function of this program?" and "What was the intention to write this program?"

A very successful principle of design in technical systems is *modularisation* and layering leading to modules for each function or each group of functions. Modules should be designed in a way that interactions between them are only possible through precisely defined *interfaces*, although usually there will also be unintended interactions. Such modules as boards for monitors, ports for network components etc. can be used again in other systems or can easily be exchanged, if the computer has to be repaired or if it is to be improved. Making use of standardised modules, hardware implementation is fast. Software implementation, on the other hand, is much more flexible. Finally, the modular construction of computers facilitates their reconstruction during reverse engineering, in particular, if one module is implemented for one function or at least one class of functions.

Even though natural brains seem to be highly structured and certain architectural features can be found in many brains, we do not know, whether this involves a kind of modularisation. In particular feed-back loops of connectivities indicate that it is not easy to identify well-defined interfaces between the building-blocks of brains. Perhaps, evolution leads to some degree of modularisation in natural systems, although there seems to always remain a certain amount of resource sharing between different functions. The task of the biologist who tries to determine the function of a particular neurone, a group of neurones or a brain area is – in a sense – one of reverse engineering.

Implementation of Plasticity

If we attempt to understand brain functions, we have to know how neuronal processing relates to behaviour (in the field of robotics the corresponding concept of "embodiment" refers to the implementation of algorithms in the robot). In the course of evolution, many brain specialisations developed subserving specific functional demands for the behaviour of the species. Still, many plastic changes take place on different levels of the nervous system in each individual to adapt it even better to functional requirements: Natural brains modify their internal organisation during ontogenetic development, guided by a genetically fixed program. Some of these structural modifications occur gradually in the course of maturation, others rather instantaneously, *e.g.* during puberty or metamorphosis. During ontogenetic development, some neurones lose their morphological and biochemical phenotype (*e.g.*, neurones derived from the neural crest) and become endocrine cells (adrenergic cells reaching their final destination in the adrenal medulla). Other examples include changes in the number of steroid receptors in the brain induced by pregnancy and parturition. Plastic changes also serve functional repair after injury. Further functional demand for plasticity depends on individual experience, and it is induced by repetitive exposure of the organism to (new) stimuli and by learning in particular. But learning is only one specific form of plasticity.

In principle, plasticity of the brain is implemented by rearrangements of neurones and their connections and by changes in the efficacy of neuronal information transfer. These plastic changes may occur on many different organisational levels of the brain. The function of a plastic change should always be defined with respect to the chosen level of explanation which, in turn, depends on the objectives of the particular investigation. Important levels of plasticity include

- Brain subsystems: Take-over of the function of one brain area by another.
- Network level: Changes in synaptic efficacy, such as homosynaptic depression as a mechanism for habituation and heterosynaptic facilitation as a mechanism for sensitisation.
- Groups of neurones: Establishment of new synapses.
- Neuronal level: Formation of new neurones by mitosis; this is a rare event, but it does occur, *e.g.*, in the brains of many fish and in songbirds.
- Subcellular level: Formation of new extensions, a process called sprouting, and division of dendritic spines (often including the spine apparatus); recruitment of synaptic vesicles.
- Biochemical level: Molecular mechanisms involved in transmitter synthesis and release, upregulation of specific receptors, notably the NMDA-glutamate receptor, alterations at ion channels, second transmitter cascades, posttranslational modifications, such as protein phosphorylation, induction of early and late genes, and many more.

Different forms of plasticity occur within different time ranges. In natural brains a major distinction may be made between temporary physiological changes in the excitability of neurones on the one hand, and long lasting changes in structural connectivities on the other. Changes in the concentration and conformation of molecular constituents, *e.g.*, undergo rapid turnover, and may favourably lend themselves for short-term events. Structural changes, however, may persist for a lifetime and are thus suitable for permanent adaptations. There is increasing evidence that not only the formation of new synapses, but also the degradation of some neuronal characteristics follows a predetermined series of cell biological reactions (apoptosis). Even the maturation of a nervous system does not necessarily only involve the setting-up of new elements, but may equally well involve mechanisms to select a subset of specific elements from an initially larger group of such elements (pruning). An intriguing biological question is, whether some of the mechanisms involved in development and in repair (*e.g.*, regeneration) of the nervous system may have been adopted to serve a new function for behaviourally induced neuronal plasticity in the sequel of learning (Schmidt, 1997).

Learning is a form of behavioural plasticity observed in most animals. It has to be differentiated between associative learning and non-associative learning (some neurobiologists prefer to call this non-associative plasticity). Examples of non-associative learning are habituation and sensitisation of behavioural responses. They are based on homosynaptic depression and heterosynaptic facilitation, respectively (Kandel and Schwartz, 1982).

Associative learning has been studied in a large variety of paradigms including classical and operant conditioning. The process involves a *short-term memory* phase, which is susceptible to physical interference, such as electroconvulsive shocks or cooling within seconds to minutes after acquisition of the information, and a *long-term memory* phase, which is not blocked by such events and may last for years. The mechanism by which short-term memory is transformed into long-term memory is called *"memory consolidation"*. It depends on transcription and translation, in particular on biosynthesis of cell adhesion molecules and other glycoproteins, and is of uttermost importance, both with respect to ageing and several diseases of the central nervous system of which Alzheimer's is but one well known example.

During memory consolidation, a temporarily adopted new behavioural program is finally implemented in form of persistent changes in neuronal ultrastructure and connectivities, presumbly including – at least in part – also those neuronal elements that were involved in the original processing of the new information during acquisition. Synaptic membranes in neuronal circuits become "primed" by the electrophysiological events during acquisition. Repeated similar, or even identical, learning-events (rehearsal) involving learning-rules, like the Hebbian condition (Hebb, 1949), may be sufficient to induce permanent structural alterations at these activated synapses. Higher organisms, however, evolved additional mechanisms to evaluate, which primed circuits are or are not to be consolidated. Here, an additional, delayed signal may decide, whether a long-term structural change is initiated. In fear-conditioning, *e.g.*, a primarily neutral stimulus gains aversive properties by repeated contingent presentation together with an aversive stimulus (Koch and Schnitzler, 1997). This can be achieved in form of a Pavlovian conditioning procedure that leads to a re-evaluation of the meaning of the formerly neutral stimulus by the brain. A structural basis of fear-conditioning has been described in detail (Koch, 1998: this issue, pp. 593–598).

It has to be kept in mind, however, that neuronal adaptations may not only be induced by neuronal activity, but can also be influenced by humoral signals, hormones in particular. Humoral signals from endocrine and neuroendocrine cells are widely distributed via the blood and/or extracellular brain fluid like a message "to whom it may concern". Glucocorticoids and related stress-hormones, may represent signals which indicate how important a preceding learning situation was for the animal. Furthermore, the decrease in glucocorticoid concentration resulting from the acquisition of a new behaviour that proved to be advantageous for the individual may be used as a signal to evaluate behavioural success. Such endocrine signals may trigger expression of protein factors which durably modify all those synapses that were primed during acquisition. In this way a new pattern of connectivity may be obtained that becomes a physical implementation of long-term memory in the brain. Cell adhesion proteins in particular, have been shown to participate in long-term memory formation (Schmidt, 1998).

In behavioural biology the term *"contiguity"* refers to the observation that the conditioned stimulus has to precede the unconditioned stimulus in order to achieve effective classical conditioning. Contiguity may be explained by conventional biochemical mechanisms of allosteric regulation via secondary transmitter molecules: An enzyme E, *e.g.*, may be activated by two different factors, A and B, which are produced in the sequel of synaptic input from two different neurones A* and B* acting on different receptors within the membrane of the same postsynaptic cell. If A binds first to the enzyme, the activated complex AE may become even more susceptible to further activation by factor B, eliciting a synergistic effect to form the enzymatically active complex ABE. However, if factor B binds first to E, the formed complex BE may become *less* susceptible to bind A, and no synergistic effect can be achieved. The suggested mechanism may, therefore, be used to implement contiguity in associative learning: Only if neurone A* (representing the conditioned stimulus) is activated prior to B* (unconditioned excitation), the postsynaptic cell is fully activated by the enzymatic complex ABE. It will be noted, that such a type of mechanism will allow implementation of temporality. On a higher level of functional analysis these conjunctures may be interpreted to represent causality.

Plasticity is a wide concept that can also be applied to artificial systems. Plasticity of performance, *e.g.*, is found in robots. There are various different possibilities for the implementation of plasticity in artefacts. In von-Neumann-computers hard- and software can easily be distinguished. Many changes in the "hardware" of brains are simulated by software changes in computers. The hardware of a computer can also be changed, *e.g.*, if more memory capacity is integrated. Usually, however, plasticity is restricted to software and programs, as in simulated neuronal nets (connectionism). Artificial evolution partly simulates biomechanisms involved in evolution, development and learning. Recently, attempts have been made to change artificial hardware in a way, so that it may operate without software and induce physical alterations in the hardware itself. However, these mechanisms implemented in artificial organisms have little in common with the mechanisms used by the brain.

Critical for artefacts is the input-output coding. To give but one example, it is very difficult to use a robot to manoeuvre a truck with a hanger backwards, if Cartesian co-ordinates are used. Once, the input-output system makes use of polar co-ordinates, the problem becomes a linear one solved by just one artificial neurone. In computers it is always *specified*, what and where information is processed and stored (discrimination between address and contents). In the brain, however, the neuronal connectivities define to which neurones an excitation is transmitted and to which part of the brain this information is conveyed. Some of these connectivities remain plastic – at least in some animals and in some humans for some time.

Nature and Role of Algorithms

The concept of algorithm

By algorithms we solve problems. There are many algorithms. A typical example is multiplication: At school we learn the rules to multiply *any* two integers. Other algorithms serve in adding numbers, in subtracting, dividing, in solving equations, constructing geometric figures, winning games, checking for logical truth and logical validity. Computer programs are special algorithms. Our brain also solves problems. Does it work algorithmically? In order to discuss this question, an explication of the concept of *"algorithm"* is given in three steps of increasing precision:

- In everyday language of scientists, an algorithm is a general problem solving procedure or, in other words, a formal description of the manipulation of information. In many cases, such a vague conception suffices. But sometimes we need more precision, especially when we investigate the scope of such procedures.
- More precisely, an algorithm for a problem class C is defined as a procedure yielding the solution to every problem from C in finitely many determinate steps. This definition exhibits three important traits of algorithms: *Generality*: An algorithm solves not just one problem, but all problems of a problem class. The data and range of single variables may belong to an infinite domain such as integer values, functions or even algorithms. *Finiteness*: The algorithm is laid down in finitely many prescriptions, and it solves the problem in finitely many steps. *Deterministic character*: In every step the next step is uniquely specified. The set of basic operators may vary deeply. Some algorithms are formulated in terms of equations, some in determinative constructs of a programming language and even others may incorporate operators that generate random numbers. Sometimes, *elementarity* of steps is required, but this is not necessary.

For some procedures, termination is not guaranteed. Should we call them all algorithms? Take square roots as an example. The usual root extraction terminates and the result is rational: $\sqrt{6.25} = 2.5$; $\sqrt{(16/9)} = 4/3$. It does not terminate, if the result is irrational: $\sqrt{2} = 1.4142\ldots$ If a procedure doesn't terminate, *i.e.*, if the problem is not solved in finite time, the original problem is not solved at all. *In practice*, then, we must restrict our claims to finite precision, *e.g.*, to four digits. The procedure then terminates, and we have an algorithm in the strict sense defined above. Thus it makes sense to restrict the concept of algorithms to terminating procedures.

- For some purposes mathematicians, metamathematicians and computer scientists need an even more precise definition, in particular, when they want to establish the *limits* of algorithms. Such a definition can be given by the concept of the *Turing machine* specifying the mechanical or automatic character of an algorithmic computation.

According to the *Church-Turing Thesis*, every computation executed by humans can also be done by a Turing machine. This claim cannot be proven, but there are good arguments supporting it. But may *all* functions exhibited by brains also be exhibited by machines? May even every mental property exhibited by a natural system be exhibited by an artificial system (as proposed by Buschlinger *et al.*, 1998: this issue, pp. 455–479)? Although this question cannot yet definitely be answered, there are interesting arguments about it. In particular, it is a controversial point whether it is possible to separate out a finite set of building blocks of computation, since the brain makes use of physical properties of the organism and the outside world in order to optimise its behaviour. Various versions of the Church-Turing Thesis and the most ambitious point of view of Artificial Intelligence have been discussed by Hofstadter (1979).

The problem of embodiment

In organisms as in artefacts, algorithms must be "embodied", or implemented. On the conceptual level, the distinction between the algorithm and its implementation is clear enough, but how about its relevance for our question? From the fact that the program embodied in a computer plays a causal role (the computer follows the program's instructions) we cannot conclude that this is also true for brains. By definition, algorithms are *formal* procedures, abstracting from the underlying properties of the physical system. Every algorithm can be implemented in various ways. Therefore, we must never take the formalism as the mechanism itself. The question, then, is not whether the brain *"is"* an algorithm (it is not), but whether algorithms are appropriate means to describe the behaviour and achievements of natural brains.

The role of the physical *environment* in which an algorithm is implemented was discussed by Thompson (1997): Whether a pure algorithmic description of the brain's work could realise (instantiate) the properties of the brain seems to be doubtful. As intimated by Thompson's experiments on *F*ield-*P*rogrammable *G*ate *A*rrays (FPGAs), there are differences between the simple description of algorithmic work *in terms of a model* and the realisation of algorithmic work in a physically real circuit. A FPGA is, structurally spoken, a chip with an array of certain components and wires. The components can be connected to the wires. The connections themselves can be determined by electronic switches each of which can be addressed by (software-) memory on board of the chip. After fixing the connections of a FPGA, it is a real-world hardware-chip comprising "software built into wires".

There are two major differences between hardware-realisation of software and software (algorithmic description) executed on a given hardware: (1.) The possibility to exploit real world properties (the physical medium) and (2.) the necessity to regard physical attributes of a hardware realisation not only as accidental, but as essential features of an implementation. In Thompson's experiments it became clear that the size, shape and location of the components are not just necessary for their existence (and, therefore, negligible in a pure software description), but they are crucial for determining the interactions between the components themselves. Probably, the most striking result was the observation that there were parts of the FPGA that influenced the FPGA's total behaviour, although they had no connectivities to the output. Possibly, there are more subtle physical effects at work that cannot be completely described by the wiring of the components. But this is exactly what is done in a mere software description.

Different levels

Is it useful or even necessary to distinguish different levels in the description of brains? On some levels, a description by algorithms might be appropriate, on others not. Probably most behavioural performances of brains can be simulated by algorithms. We also have to distinguish, what a neurone *can do* and what it *actually does.* A major function of a particular neurone is to perform certain computations. Its morphology puts constraints on the computation (*e.g.*, temporal constraints in integrating different inputs and generating output). These constraints determine what it can do – in principle. But what it actually does, is mainly determined by the context. Depending on the context, an individual neurone will perform different tasks and even different computations.

Discreteness

Algorithms are based on discrete data, and all standard computers work on discrete procedures. How about the brain? Just like computers, brains are systems with many discrete subsystems, and at any moment in time, they comprise a limited number of discrete states. Can analogue devices solve more problems than Turing machines? If not, are they more effective, *e.g.*, faster or less memory demanding? No consent was obtained on the question, of whether analogue and discrete systems are finally equivalent, nor whether conversion of discrete to analogue data is more difficult than the conversion of analogue data to a digital form (some properties of attractors may be changed by such transformations).

From the perspective of neurobiologists, an important question is, how algorithms are implemented into natural systems. In neurones, EPSPs and IPSPs (excitatory and inhibitory postsynaptic potentials) can be induced via activation of receptor molecules. Integration of PSPs in the postsynaptic neurone may, or may not, reach a threshold to elicit an action potential as a discrete all-or-nothing response. On the electrophysiological level of analysis, the PSP is usually conceived as an analogue signal, but it emerges from discrete events occurring on the molecular level (movement of ions). On the level of biochemical reactions and below, in some instances, it appears less appropriate to describe brain functions by those deterministic algorithms and basic operators used in today's programming languages.

Chance events

If we attempt to model behaviour in terms of algorithms, we have to take into account that the behaviour may ultimately be emergent from lower levels, *i.e.*, the performance and output of brains may be influenced by stochastic effects of quantum mechanics, mutations induced by radioactive irradiation in DNA molecules or by steric hindrance resulting from Brownian movement etc.. Causalities might arise from levels below those which can be formulated by means of abstract, deterministic algorithms. How important are chance events? Are they constitutive for the functioning of natural brains? Four possibilities have to be envisaged:

- There are no chance events in the brain. This possibility seems to be ruled out by our observations.
- There are chance events in the brain, but they do not influence its performance.
- There is an influence of chance events on the brain, lowering the accuracy of its performance.
- There is an influence of chance events improving the performance of the brain.

Having discarded the first possibility, we find examples for all three cases left. Most astonishing is the fact that chance may *improve* the performance of a system. An instructive example is given by Schulten (1987). It became evident from the discussion that for an adequate description of brain functions by algorithms, stochastic elements have to be introduced and are required for performance. When random noise is removed from a simulated central nervous system, this results in a deterioration of its functioning (see Fuster, 1998: this issue, pp. 582–592). Apparently, chance (stochasticity of events) is a constitutive entity for the functionality of both, natural and artificial systems. Chance elements in the strict sense are not deterministic. Thus, a system working with genuine chance is not strictly algorithmic. It is, however, very difficult to decide whether a seemingly random element is truly random. Strictly speaking, it is even impossible to prove indeterminism. In most cases, even chance generators in computers work deterministically. In any case, we might be forced to ask whether the brain works by algorithms combined with chance elements and even whether chaos may play a constitutive role in brains.

Further questions

How and why did the class of problems solved by brains increase in the course of evolution? If all ecological niches manageable by simple devices are occupied, more complex devices, including problem solving devices, are necessary to explore, to use or even to create, more demanding niches. This explains the overall increase of complexity in evolution.

The question was raised whether there could be a meta-algorithm telling engineers how to construct – to any specific mental property – an artificial system exhibiting this same property (Gierer, 1985).

Enhancement of Efficacy by Internal Reconstruction of Brains and Computers

Evolution proceeds gradually, but occasionally big advances are also made in a single step. The neurone and, later, the isocortex may be regarded as examples of such achievements. Once "invented", they were used over and over again for many different purposes. Nevertheless, even a brain without an isocortex may reach high levels of performance. For each such purpose, however, a specialised network has to develop which, thereafter, cannot easily be used for other tasks.

The elaborated isocortex with an almost identical architecture (hence its name) in all areas analysed so far is a characteristic of the mammalian brain. Why was the isocortex so successful in evolution? – The isocortex of the telencephalon has much in common with a general purpose calculator. Its structural regularity and uniformity may provide functional universality and indicate that each cortical "module" essentially performs one stereotyped computational step, which, however, may not be very elaborated. A special feature of the mammalian telencephalon is that the isocortical tissue can be applied serially, one module after the other performing the same algorithmic operations several times, each time with inputs mixed up anew by convergence and divergence (Krüger, 1998: this issue, pp. 599–609). Here, the output from one cortical module may serve as an input signal to the next. This is a property of the cortex not shared by brain nuclei (that represent the major evolutionary achievement of the avian brain) nor by other neural tissues.

The rapid increase in the size of the isocortex during evolution may perhaps result from the fact that it makes use of distinct modular units applied serially. This architectural design, makes it feasible to increase the complexity of processing by simply adding such modules. The latter idea is largely approved, although questions remain, *e.g.*, why some clinical cases have been reported, in which the damage of large parts of the cortex did not profoundly impair the performance of the individual. It was emphasised, that an increase in brain size is not the only important means to cope with environmental challenges. Social insects partly compensate for their small brains by the establishment of large communities ("states"). Here, some tasks or problems may be solved, because one individual – out of thousands – finds a solution by trial and error. The new behaviour is then imitated or learned by the others. The cortices of some orders in the class of mammals have even secondarily been simplified, *e.g.* in salamanders. Here, the number of neurones was reduced for further functional specialisation. Neuroanatomists point out that in higher vertebrates the number of inhibitory cortical neurones is particularly high. Their inhibitory function might provide an alternative to the pruning of neurites in ontogeny as a means to improve functional specificity.

The idea was discussed that the cerebral cortex might have developed as a general purpose structure for which no *particular* ecological pressure can easily be delineated (it was compared to the invention of money in human societies). It has been argued that no distinct function can be assigned to some parts of the cerebral cortex, but it would rather have to be seen as an option to guarantee flexibility and to solve rarely occurring problems in the future. On the other hand, selection is blind for history and future demands. Though the environmental pressures act on systems with a certain history, selection itself acts on the system without regarding this history. Selection favours only temporary advantages to improve the reproductive fitness, and it drives the evolution into relative optima, but not towards an absolute maximum.

On the other hand, the implementation of the cerebral cortex may be considered as just one evolutionary achievement amongst many others, such as cell membranes, captured micro-organisms developing to constitute intracellular compartments, ion channels, secondary transmitter cascades etc.. This may be compared to the improvement of computer design in different generations, first making use of electron tubes, then of transistors, semiconductors, integrated circuits, network technology and higher programming languages. At several levels the performance of a computer can be enhanced by integrating and multiplying certain components, which in some respects might be comparable to the enlargements of the cerebral cortex. With state-of-the-art-algorithms this is easily possible with memory, additional processors or hard drives, and with systems and programs operating in parallel.

Limitations of Brains and Computers

Problem-solving

In history, the most recent technical invention (the clock, the steam engine etc.) has always been taken as a metaphor to describe the brain. How are functions of the brain differently implemented in computers? The main problem in answering this question is probably our lack in knowledge of the algorithms working in brains. Which problems can be solved by algorithms, be it in principle or in practice, within realistic time or with realistic memory, is nowadays intensively investigated, and there are already some interesting results. From mathematics, especially from geometry, we know that there are problems which cannot be solved with specified means, *e.g.*, trisecting angles, doubling cubes and squaring circles only with ruler and compass. But there are also problems which *probably* cannot be solved by *any* algorithm (*e.g.*, Turing's halting problem). To know that a problem is unsolvable may be useful because this knowledge saves time, energy and money. There are even problems for which an algorithm is known, but it takes too much time, too many steps or too much memory to allow a solution in any realistic way. For some of these problems it has been proven that no elegant algorithm exists (take the game "roadblock" in Stockmeyer and Chandra (1979)). For some problems this has not been proven, but for good reasons mathematicians gave up hope to find an efficient algorithm for them (*e.g.*, NP-complete problems, such as the travelling salesman problem).

Can we judge the difficulty of a problem? For a *solved* problem we use the number or length of the steps that were necessary to solve it. In geometry the number of subsidiary lines is a good measure. But for *unsolved* problems? Some experts seem to have a feeling for the difficulty of problems. A striking example was the mathematician Paul Erdös (1913–1996). He used to pose unsolved problems and to offer prizes for their solution, reaching from one to 25,000 dollars. Sometimes he had indeed to reward a solution given, but never by more than 1000 dollars. Obviously, he must have had some intuitive feeling for the difficulty of the problems. Could this ability be made more explicit?

Decision-making

May an artificial system decide something? There are already interesting examples, that computers have found *mathematical proofs*, including the four colour theorem (1977), Mertens' conjecture (1984), Waring's conjecture (1986) and Robbins algebra (1997; cf. Buschlinger *et al.*, 1998: this issue, pp. 455–479). The problem is not with deduction, but with an informed guess on whether a newly derived theorem is interesting, useful or deep. It has been shown that so far deduction machines will get lost in true but trivial theorems (Ebbinghaus, 1992). Obviously, this is related to the problem, of how search strategies might be made more effective. The performance of computers depends on their fast processors and their large memories. If we engage a computer to look up all papers with the word "brain", it will in principle be able to do so. But, we don't have the algorithms to handle this amount of information so as to make sense of it (*e.g.*, to write a review). Computers are able, however, to find new and better solutions by combining more information from many different sources and by averaging over or by neglecting individual statements and demands.

The frontal cortex, which is functionally involved in response selection and decision-making, obtains innervation from a larger number of sources than any other part of the brain. The decisions reached by the frontal cortex may appear to represent free will, but actually they may be derived from an integration weighing the amount of all inputs. There may operate a "winner-takes-all" type of competition: Who ever wins by number of input signals takes over the decision; free will may be just an illusion. If we cannot obtain any additional information from introspection as opposed to inter-subjective observation, it appears likely that computers can in principle do all that human brains can.

Development of aesthetic categories and creativity

Computers can already develop categories of aesthetics. The fact that computer-art is not approved by everybody, certainly does not rule out, that it may (or once will) be considered as true art. Computers can, *e.g.*, use digital image analysis and averaging procedures to superimpose many photographs of female faces and construct an averaged face. As compared with each of the original faces, this artificial construct was preferentially chosen as the prettiest by a group of male test persons. Furthermore, a computer may store images of faces rated "attractive" or "more attractive" by the observers. Then, it can use the analysed features to extrapolate and construct a super-attractive face, comparable to a super-natural key-stimulus in behavioural science. Computers may also simulate emotions. But, do we want to develop computers with all natural features of human brains? Can we construct better ones?

As compared with children, computers are still rather limited in their ability to understand and use human language. Also creativity may demand more complex features than simple information processing, evaluation and decision-making. Could a computer be creative like a composer? Could it write poetry on its own, if provided (implemented) with a basic understanding of human (?) emotions, or would it fail because of a lack in "personal experience" of love and sorrow? But what is "personal experience", and what is a "basic understanding of human emotions" with respect to the attempt to build a robot? Simply adding interfaces acting on the outer world and receiving responses from it? – Surely not. This is rather easily achieved. *E.g.*, a robot may be constructed with some kind of mouth and hands enabling it to eat vanilla ice-cream. Although it will indeed "consume" the ice-cream, it will not know how vanilla ice-cream tastes and will be unable to distinguish the vanilla flavour from that of strawberry. But let us even assume, that we could provide the robot with the physical devices for those measurements that enable it to distinguish vanilla from strawberry ice-cream, does the robot then really know how vanilla ice-cream tastes? – Is there any further prerequisite besides interfaces and physical devices for measurements in order to implement taste perception (Nagel, 1974)? The question may be asked the other way round: Assuming that the brain works algorithmically, how does it achieve recognition of any specific taste? – No convincing answers are yet available. We also asked ourselves, whether a computer may exhibit humour or understand a joke made about the computer itself. – We did not discuss, whether computers have (self)consciousness.

Can we make some *predictions* on the further development of artificial systems? – Yes. On the future of brains? – Barely. And on the future of brain research? Could a computer write a grant proposal in a theoretical discipline, by looking up all successful grant proposals of the previous years? Could the gentle reader of this contribution possibly decide, whether it was composed by an intelligent computer using language recognition or by a less capable rapporteur?

Burgess N., Donnett J. G., Jeffery K. J. and O'Keefe J. (1997), Robotic and neuronal simulation of the hippocampus and rat navigation. Phil. Trans. R. Soc. Lond. **B 352**, 1535–1543.

Buschlinger W., Vollmer G. and Walter H. (1998), The mental life of artifacts; explications, questions, arguments. Z. Naturforsch. **53c,** 455–479.

Ebbinghaus H.-D. (1992), Maschinen und Kreativität: Metamathematische Argumente für das menschliche Denken. Philosophia naturalis **29**, 1–30.

Egelhaaf M., Kern R. and Warzecha A.-K. (1998), Sensory ecology and neural coding in arthropods. Z. Naturforsch. **53c,** 582–592.

Fuster J. M. (1998), Cellular dynamics of network memory. Z. Naturforsch. **53c,** 670–676.

Gierer A. (1985), Die Physik, das Leben und die Seele. Piper, München, Chapter VIII.

Hebb D. O. (1949), The Organization of Behavior. Wiley, New York.

Hofstadter D. R. (1979), Basic Books. Gödel, Escher, Bach. Chapter XVII.

Jeffress L. A. (1948), A place theory of sound localization. J. Comp. Physiol. Psychol. **41**, 35–39.

Kandel E. R. and Schwartz J. H. (1982), Molecular biology of learning: modulation of transmitter release. Science **218**, 433–443.

Koch M. (1998), How can an adaptive behavioural plasticity be implemented in the mammalian brain? Z. Naturforsch. **53c,** 593–598.

Koch M. and Schnitzler H.-U. (1997), The acoustic startle response in rats – circuits mediating evocation, inhibiton and potentiation. Behav. Brain Res. **89**, 35–49.

Krüger J. (1998), How it is to be the brain of a monkey. Z. Naturforsch. **53c,** 599–609.

Nagel T. (1974), What is it like to be a bat? Philosoph. Rev. **83**, 435–450.

Nottebohm F., Stokes T. M. and Leonard C. M. (1976), Central control of song in the canary, *Serinus nanarius.* J. Comp. Neurol. **165**, 457–486.

Pfeifer R. and Scheier C. (1998), Representation in natural and artificial agents: An embodied cognitive science perspective. Z. Naturforsch. **53c,** 480–503.

Schmidt R. (1995), Cell-adhesion molcules in memory formation. Behav. Brain Res. **66**, 65–72.

Schmidt R. (1997), Regulated expression of the CNS-specific cell adhesion molecule ependymin after acquisition of an active avoidance behaviour provides a possible mechanism for memory consolidation. In: Neurochemistry, Cellular, Molecular and Clinical Aspects (A. W. Teelken and J. Korf, eds.). Plenum Press, London, pp. 869–876.

Schmidt R. (1998), Learning and memory, neurochemical aspects. In: Encyclopedia of Neuroscience (G. Adelman and B. Smith, eds.), 2^{nd} edit. www.Elsevier/ens/articles/00000738/ tx1.htm.

Schulten K. (1987). In: Ordnung aus dem Chaos (B.-O. Küppers, ed.). Piper, München, pp. 264f.

Stockmeyer L. J. and Chandra A. K. (1979), Intrinsically difficult problems. Scientific American **240**, 124–133.

Thompson A. (1997), Artificial evolution in the physical world. In: Evolutionary Robotics: From Intelligent Robots to Artificial Life (ER '97) (T. Gomi, ed.). AAI Books, pp. 101–125.

Wagner H. and Luksch H. (1998), Effect of ecological pressures on brains: examples from avian neuroethology and general meanings. Z. Naturforsch. **53c,** 560–581.