

Introduction to the Proceedings of the Workshop “Natural Organisms, Artificial Organisms, and Their Brains” at the Zentrum für interdisziplinäre Forschung (ZiF) der Universität Bielefeld, March 8–12, 1998

Hennig Stieve

Institut für Biologie II, RWTH Aachen, D-52074 Aachen, Germany

Z. Naturforsch. **53c**, 439–444 (1998); received May 18, 1998

There have been many meetings in recent years on “Brain and Computer” topics under a number of different aspects. Although it is not my own field of experimental work, I have been discussing this subject with brain- and computer scientists, roboticists, psychologists, and scholars of the humanities for several years (Stieve, 1995). The conferences about mind and brain which I attended or read about were often characterized by a lack of true discussion, i.e. disputation of differing arguments, and did not seem to have an impact on ongoing research. People presented their data, findings and opinions in monologues without seriously discussing their grounding and the reasons for the disagreements with others. This was especially so in the those fields in which consciousness and feelings are involved, e.g. the question whether a computer can have feelings. In many cases convictions were mixed with confirmed facts, and plausibility or wide acceptance was taken as proof. Compared to conferences in biophysics and biochemistry this is quite a contrast.

Therefore I felt there is now – after the first wave of these “Brain and Computer” conferences – even more need for interdisciplinary discussion in certain parts of this field. I tried to assemble a group of people from neurobiology, computer science, robotics and philosophy for a serious discussion. With this aim I pursued the following strategy:

I contacted more than 50 people from the various fields concerned and received very useful advice. I only want to mention a few of them who were especially helpful. Peter Bieri, Berlin, Tobias Bonhoeffer, München, Valentin Braitenberg, Tü-

bingen, Hans Flohr, Bremen, Hans-Joachim Freund, Düsseldorf, Karl Georg Götz, Tübingen, Klaus Hausen, Köln, Wolfgang Prinz, München, Helge Ritter, Bielefeld, Jürgen Schnakenberg, Aachen, Werner von Seelen, Bochum, Jan-Dieter Spalink, Durham, Gerhard Vollmer, Braunschweig and Reto Weiler, Osnabrück.

At first I tried to organize a one year research group at the Zentrum für interdisziplinäre Forschung (ZiF) in Bielefeld. But this seemed not to work. Later I became convinced that an intense workshop would be a better frame for such a discussion.

For such a workshop it seemed necessary to

- Select topics which need and deserve interdisciplinary discussion and where such discussion promises to yield results, because there are many very interesting topics in this field where discussion does not promise results now.
- Choose a framework which ensures a successful discussion. I chose the Dahlem Workshop Concept with which I already had some positive experience.

A conference of the Dahlem type had never been held at the ZiF in Bielefeld. This appeared to be quite a challenging opportunity to demonstrate at the same time the efficiency of this concept. On the other hand, the directors of the ZiF were not easily convinced that this concept would work. For scholars of humanities a conference which is so different from a sequence of monologues and from the standard workshops of the ZiF may be difficult to accept. All the more, we wish to thank the ZiF which after some hesitation decided to house and support for the first time a Dahlem-type conference, our workshop. We also wish to thank the staff of the ZiF, especially Marina Hoffmann, Andreas Lueking, and Daniela Mietz, and its managing director Dr. Gerhard

Reprint requests to Prof. H. Stieve.

Fax: (0241) 8888133.

E-mail: stieve@rwth-aachen.de

0939–5075/98/0700–0439 \$ 06.00 © 1998 Verlag der Zeitschrift für Naturforschung, Tübingen · www.znaturforsch.com. D



Dieses Werk wurde im Jahr 2013 vom Verlag Zeitschrift für Naturforschung in Zusammenarbeit mit der Max-Planck-Gesellschaft zur Förderung der Wissenschaften e.V. digitalisiert und unter folgender Lizenz veröffentlicht: Creative Commons Namensnennung-Keine Bearbeitung 3.0 Deutschland Lizenz.

Zum 01.01.2015 ist eine Anpassung der Lizenzbedingungen (Entfall der Creative Commons Lizenzbedingung „Keine Bearbeitung“) beabsichtigt, um eine Nachnutzung auch im Rahmen zukünftiger wissenschaftlicher Nutzungsformen zu ermöglichen.

This work has been digitalized and published in 2013 by Verlag Zeitschrift für Naturforschung in cooperation with the Max Planck Society for the Advancement of Science under a Creative Commons Attribution-NoDerivs 3.0 Germany License.

On 01.01.2015 it is planned to change the License Conditions (the removal of the Creative Commons License condition “no derivative works”). This is to allow reuse in the area of future scientific usage.

Sprenger, and last, but not least Bettina Halbe, a student of the RWTH Aachen, for their sensitive, effective and friendly help in preparing and conducting our workshop.

The Concept of the Workshop

The concept of the Dahlem Conferences is a concept for an unusual – but to my mind very efficient – type of workshop. This Concept has its emphasis on the discussion of the meaning of new and old observations, models, and theories and aims to encourage new cooperations and the design of new critical experiments.

This concept includes some “tricks”:

- Small discussion groups which design their own agendas. This allows effective discussions in which each member of the group can participate. The jointly designed agenda allows to find room for the topics which the members find worth discussing and brings the responsibility to the participants.
- No lectures at the workshop, but background papers distributed in advance. Lectures tend to be monologues and could kill the path of the discussion. Background papers distributed in advance allow the members to prepare themselves to the discussion.
- Joint group reports which are published right after the workshop with all the members of a discussion group as the authors. This is an essential requirement because it forces the group to reach conclusions during the limited time available for discussion, and (since every member is a co-author) every member of the group has to agree on what is written down. The task to finish a group report implies that everybody tries to stay focused during the group discussion.
- A “Program Advisory Committee” consisting of representatives of the various disciplines relevant for the theme of the workshop. It proposes the topics and the participants of the workshop and their roles as moderators, rapporteurs and authors of background papers.

The Preparation of the Workshop

The Program Advisory Committee: For the Program Advisory Committee I won the cooperation of six persons from different backgrounds. So the Program Advisory Committee consisted of the

following seven: Niels Birbaumer, Tübingen (clinical and physiological psychology); Tobias Bonhoeffer, München (neurobiology, brain physiology); Yadin Dudai, Rehovot (neurobiology, memory); Rolf Pfeifer, Zürich (computer science; artificial intelligence and simulations); Jan-Dieter Spalink, Durham, NC (electrical engineering, communication networks); Hennig Stieve, Aachen (neurobiology, photoreception); Gerhard Vollmer, Braunschweig (philosophy, epistemology). Three of them already had positive experience with Dahlem conferences in Berlin.

The Committee met in Bielefeld in the Zentrum für interdisziplinäre Forschung (ZiF) on April 14 and 15, 1996 – almost two years before the workshop – selected the topics for the four discussion groups, chose a name for the workshop, and suggested the participants to be invited and their roles in the workshop.

The date of workshop, 8–12 March, 1998 was chosen to fit to the possibilities of the Zentrum für interdisziplinäre Forschung (ZiF) in Bielefeld and to be outside the teaching periods of German universities. However, this is the time of many other meetings and in addition it is during the teaching periods of universities in some other countries. Our time window happened to include the 70th birthday of the designer and founder of the Dahlem Conferences, the late Silke Bernhard, to whose memory the conference is dedicated.

With our program we closely followed the well-tested time schedule of the Dahlem-workshops, with one exception: we made it one day shorter. It turned out that it probably would have been better not to reduce it.

The participants: As the number of participants of the workshop should not exceed about 40, only a small number of the persons interested and competent in the concerned fields could be invited. We tried to assemble a variety of participants including younger and older, who were prepared to cooperate, could supply interesting contributions, were able to listen with an open mind, and would commit themselves for the entire duration. Not all who were invited could participate, but those who did come made up for these losses by working hard, inventively, and cooperatively.

The multidisciplinary of the workshop participants may be demonstrated by the following. Normally, a participant personally knows about seventy

or even eighty percent of the participants already before coming to a workshop. In ours, most participants had not met at least two third of the others.

Our aim, then, was to have an efficient discussion between workers in neurobiology, robotics, informatics (computer science), and philosophy (epistemology) in comparing certain properties of brains and computers of organisms and robots. Whereas the interdisciplinary exchange of views between computer people and brain researchers has already begun a few years ago, the inclusion of epistemologists seems an innovation. To make such different people talk to each other in an understanding way (patiently listening and exchanging arguments) is no easy venture. We hoped that the Dahlem concept would make this possible.

The topics: The suggested topics for the four discussion groups of our workshop were chosen because the Program Advisory Committee thought that they deserved an open-minded interdisciplinary discussion and were likely to lead to interesting results:

1. Representation of the environment in natural and artificial systems.

How is the information about the properties of the environment represented in the brain to ensure the generation of appropriate behavioral actions? Examples: The representation of space, motor action and faces in brains and in the analogous control “organs” of robots.

2. Functional advantages of organisms with brains in evolution.

The functional consequences of certain brain properties for the behavioral fitness of organisms. Examples: brains of birds and insects as instruments to cope with ecological pressures, the evolution of robots’ capacity to cope with ecological pressures, and the significance of the brain as a learning machine for success in evolution.

3. The behavior of natural and artificial systems: solutions to functional demands.

Comparing performances of systems with functional demands, including questions like behavioral decisions in conflicting situations. Suggested examples: How molluscs decide on taste preferences, recognition of optical objects in natural and artificial systems, resource allocation in distributed technical networks, and walking robots.

4. Emergent properties of natural and artificial systems.

The term “emergence” has become fashionable. It is used with quite different meanings by brain researchers, mathematicians, and philosophers. It seemed to be useful to understand the different definitions in which the term is used, try to find a definition of emergence which is useful for our purposes, describe supposed emergent phenomena in brains and computers and to discuss how far we today understand their origin. Suggested topics were non-linear mechanisms in brains and computers, brains as generators of emerging faculties, and the so-called binding phenomenon, i.e. the adequate handling of distributed brain activities which correspond to different properties of the same object.

It was suggested by the Program Advisory Committee that the discussions at the workshop should focus on the comparison of properties of brains and computers with an emphasis on the functional dependence of brains on their natural “bodies” and of computers in robots on their connected machineries. The significance of this interdependence for the behavior of an organism has only gradually become clear during the past few years. Comparing those properties of brains and computers which affect the behavior of organisms and robots helps to understand both better. Most of today’s robots are still relatively rigidly programmed to fulfill certain well-defined tasks in a straightforward “if – then” manner. There is, however, a development of more advanced robots which can make more autonomous decisions to be used e.g. in emergencies or on Mars. These may allow even more interesting comparisons to living organisms.

We know of relevant properties of computers which cannot be achieved by brains (e.g. speed of processing). Are there relevant properties of brains (accessible to natural science) which cannot be achieved by computers, now or ever?

Limitations: Natural science is limited to objectively observable phenomena, in other words, to consider only phenomena which are accessible to independent observers is a *conditio sine qua non* for “clean” natural science (Stieve, 1998: this issue, pp. 445–454). This confinement has certain pragmatical advantages: The scientific experiment is the acid test which decides whether an assumption or a theory can survive or has to be discarded.

Humanities do not have this independent judge. Therefore many old theories cannot conclusively be discarded (e.g. those of the ancient Greeks). They may be recycled over and over, according to the “Zeitgeist”. By contrast, in natural science many theories (some of them brilliant) have been falsified and discarded by the force of critical experiments.

This limitation sets boundaries to pure natural scientific brain research. It excludes consciousness and feelings as possible objects of research since these are only subjectively accessible and alternatives cannot be decided objectively. If beliefs and convictions come into play, this makes a discussion very difficult, if not impossible. Whether computers can have feelings, be creative or have intuitions like a human mathematician are questions of belief; they can not be scientifically decided.

Other brain problems which can not be solved today, may be solved in the future. Our present conjectures or assumptions may be right or wrong. An example is the historical controversy about the signal conduction in nerves: In the twenties and thirties of this century there was a vigorous controversy whether signal conduction in nerves is “electrical” or “chemical” (see e.g. Muralt, 1946). Both sides had powerful advocates. Finally, this controversy was solved experimentally in the late forties and fifties (see Hodgkin, 1951). Those 1920s scientists whose theories emerged as true had not had the better arguments, but better intuition and possibly more luck. Guessing the right answer is not in the end sufficient for the progress of our scientific understanding, but designing and making the critical experiments is, and even a wrong theory may lead to it. Of course, to look at the available data in the right way, which involves sensitive guesswork, intuition, and even aesthetics, is very important for designing good theories and important experiments.

For our understanding of brains and computers there may be problems involved which are insoluble in principle. Gödel has shown that there are certain mathematical problems which can not be solved: “All consistent axiomatic formulations of number theory include undecidable propositions” (Gödel, K. (1931), quoted after Hofstadter (1979)). Similarly, the so-called ‘strong emergence’ is defined as the occurrence of a novel phenomenon which cannot be explained on the basis of the un-

derlying structures and may play a part in brain function (see Stephan, 1998: this issue, pp. 639–656 and Walter, 1998: this issue, pp. 723–737). However, since today there is no clue to determine whether such a problem is unsolvable (e.g. whether a phenomenon is strongly emergent), the only successful strategy for brain research is to try to solve the riddles at hand until we know more about them.

Our Program Advisory Committee suggested that topics which today are unlikely to lead to much progress (e.g. the mind-body problem) must not occupy much of the group’s activities, but the arguments for the current stalemate situation should be stated in the discussion group report. It seemed, however, desirable to come to a joint statement on what the participants agree and where and why they disagree.

How Was Our Workshop?

Our groups worked in a self-organizing manner after the moderator and rapporteur and main topic had been determined by the Program Advisory Committee. They found their own agenda and way to handle their tasks. They acted somewhat like organisms. Sadly, our workshop was hit by unforeseeable losses. Apart from the usual drop-outs, an infectious flu deprived us of several important participants. As they had been carefully chosen to be part of our small number, this was a hard blow. But this is the beauty of organisms, they can compensate for losses. If a shore crab (*Carcinus*) loses one or two legs it can adapt its way of walking in a way that beautifully compensates for the missing limbs. And our discussion groups did exactly the same thing. They compensated for the missing members synergetically and did a complete job by modifying the suggested topics according to the expertise present. Actually the groups discussed the following:

Group 1: Representation in natural and artificial systems:

The concept of representation was in the focus of the discussion in this group. Examples can be found in descriptions of the visual system, the motor system, as well as in central systems like the hippocampus. Although the description of neuronal activity is often phrased in terms of representa-

tions, the concept is less well understood and has several implications which have to be considered.

Firstly, the relevant aspect of neuronal activity establishing a representation is subject to an intense debate – some researchers favor the mean activity of neurons as the relevant signal conveyed to other neurons, while others emphasize the precise temporal structure of neuronal activity. Thus, estimates of what is represented in a particular structure are dependent on the assumptions made on the relevant variables and might be different for an outside observer and other brain structures.

Secondly, the concept of neuronal representations is often used in a passive sense. The emphasis is placed on recreating properties of the external environment as faithfully as possible. Thus, the interaction with neuronal structures upstream and downstream have traditionally been assumed to occur in a feed-forward manner.

Thirdly, representations are not static entities but have to form during development and learning. This poses the questions of maintenance of representations, of the interaction of new experiences with previous ones already stored, and poses the problem of grounding a representation in the real world.

For these three aspects this group considered it timely to investigate the concept of neuronal representations within several fields of neuroscience research. Exploiting the broad range of expertise of their members, the group undertook to investigate the concept of neuronal representations from as many aspects as possible. In the group report a series of examples is presented in a comparative fashion. For each example an attempt is made to identify the relevant variables, the nature of the interaction with upstream and downstream structures and the influence of training and development. By this careful analysis of neuronal representations in different contexts a more precise and consistent use of this concept might be possible.

The most outstanding result was the commonality in a number of principles of representation that were common to all of these examples. The importance of the interaction of organisms with their environments was observed in each of the systems studied.

Group 2: Influence of brain and computer design on the performance of natural and artificial organisms:

The discussions were focused on the relations between *structures* (the materialistic implementation at the level of molecules, subcellular organelles, neurones, and networks) and *functions* defined teleologically for brains and computers (on a variety of different levels of explanation). Examples of evolutionary achievements in natural organisms were compared with the advancements of computer and robot design that improve their performance. Pivotal points of the discussions were the role of algorithms for the functioning of brains and computers, their embodiment in the internal architecture of natural brains and artifacts, the implementation of plasticity, and some possible limitations for the performance of computers, e.g. in the field of creativity.

Group 3: The behavior of natural and artificial systems: solutions to functional demands:

Natural organisms have different strategies for the adaptation to their environment. For example there is a great variety of eye types in natural organisms adapted to a certain environment and different tasks. The discussion group tried to focus on the similarities between natural and artificial systems and the benefits for the different fields of research. At the moment it is common sense that both fields could profit from each other, but the major problem is to give concrete examples for successful interactions and to raise new and interesting questions. The examples discussed ranged from sensory information processing to motor behavior, and some speakers gave nice examples of artificial systems incorporating biologically motivated strategies in sensory-motor integration. But these examples showed clearly that the development of a theory of the motivational system is a demanding task with a lot of unresolved problems both in natural and artificial systems.

The role of models, their usefulness and limitations, for brain and computer research was also discussed in this group.

Group 4: Emergent properties of natural and artificial systems:

The discussion of this group focused on the phenomenon "emergence". Due to the complexity of natural systems and the limitations of the related theoretical approaches to their properties, the term emergence originally was mainly used to label certain properties as somehow appearing from some underlying structures. As it became more and more technical, the need increased to relate this term to other theoretical approaches and to discuss it in terms which can be reduced to observable phenomena. Thus, suggesting that all observable phenomena can be reduced to a consistent set of underlying scientific laws, at least in principle, would imply that emergence has no intrinsically methodological significance as a technical term. On the other hand, the term emergence would play a significant role under the assumption that there exist properties that in principle are not reducible. The related question about the principally reducibility of systems phenomena is truly at the very border of science and appears to be impossible to answer.

As a practicable avenue, the term emergence was discussed in the frame of pattern formation, especially in the context of neurobiology. It turned out that pattern formation may provide a methodological and conceptual frame to discuss empirical findings in the fields of neurobiology and behavioral sciences on a rigid mathematical ground while opening clearly shaped ways to different but related fields.

In this group it took a major effort on the second day to resolve some initial misunderstandings between brain and computer scientists on the one hand, and philosophers on the other. But this paved the ground to an interdisciplinary discussion which apparently was satisfactory to all the members of the group.

Conclusions

The aims of our workshop were to exchange ideas and arguments and to evaluate them jointly, and to initiate and plan new experiments and co-operation.

Here we present the outcome of our efforts. As expected our workshop was hard work. But all participants seemed to enjoy to work so hard for satisfactory results. Many members, especially the rapporteurs, often worked until very late at night (sometimes until 4 a.m.). The longer the workshop proceeded, the more we all felt the time pressure to produce results jointly before the workshop ended. However, with high pressure you can make small diamonds from coal!

Our group reports are of course not the ultimate truth but represent the present state of understanding as determined by the joint expertise of the members of the group; they state the agreements and disagreements and reasons for both.

I hope that for all the tough work we all enjoyed this exercise in discussion culture, the intellectual pleasure of exchanging competing arguments.

My thanks go to those who read the manuscript and made helpful suggestions and contributions, namely: Bettina Halbe, Peter König, Hubert Preissl, Stefan Reimann, Rupert Schmidt, Tilman Stieve and Gerhard Vollmer.

The workshop was supported by the Deutsche Forschungsgemeinschaft, Ministerium für Wissenschaft und Forschung des Landes Nordrhein-Westfalen, Zentrum für interdisziplinäre Forschung (ZiF) der Universität Bielefeld, Daimler Benz AG, Georg Thieme Verlag, Schering AG, Siemens AG, and SmithKline Beecham Stiftung.

- Gödel K. (1931), Über formal unentscheidbare Sätze der Prinzipia Mathematica und verwandter Systeme. I. *Monatsh. Math. Phys.* **38**, 173–198.
 Hodgkin A. L. (1951), The ionic basis of electrical activity in nerve and muscle. *Biol. Rev.* **26**, 339–409.
 Hofstadter D. R. (1979), Gödel, Escher, Bach: An Eternal Golden Braid. Harvester Press Ltd., Hassox, Sussex, GB.
 Muralt A. v. (1946), Die Signalübermittlung im Nerven. Birkhäuser Verlag, Basel.

- Stephan A. (1998), Varieties of emergence in artificial and natural systems. *Z. Naturforsch.* **53c**, 639–656.
 Stieve H. (1995), Limits of natural science: Brain research and computers. *Z. Naturforsch.* **50c**, 317–336.
 Stieve H. (1998), Comparing properties of brains and computers. *Z. Naturforsch.* **53c**, 445–454.
 Walter H. (1998), Emergence and the cognitive neuroscience approach to psychiatry. *Z. Naturforsch.* **53c**, 723–737.