Figuring out How the Mind Works

—

The State of Affairs

Tino Schwarze

September 28, 2001

Proseminar "Language, Mind, Brain"

Christoph Haase

SS 2001

Contents

1 Introduction

The development of science has been driven by mysteries. Ever since, humans asked themselves where they came from and why the world is like it is. This curiosity was first satisfied by religions. But from an atheist point of view, religions were merely the first theories about the world. Religions almost always deal with unobservable and mighty powers.

By and by, people tried to find more objective descriptions and explanations for natural powers. They also started to bother about themselves. Philosophy arose, first as a supplement to religion, later as a replacement.

It is known that the ancient Greeks started to reflect on themselves. They were curious about how the mind works and what mind is. In the 17th century Descartes saw the mind as a metaphysical entity interacting with the physical body. Boole tried to explain thinking with a system of mathematical logic. He considered thinking to be manipulation of symbols representing entities in the world. His thoughts were later assimilated and extended by Frege and again by the philosophers Russell and Whitehead which formalised the system, added a proof system and created the "Principia Mathematica", a compendium of human reasoning.

It was just in the last century that mathematician Allan Turing worked on finite–state automata and proposed his theoretical "Turing machine". This machine was extremely simple[1] compared to real computers and of purely theoretical nature[2] but it laid the foundations for digital computers.

Shannon (1948) was one of the first to figure out that Boole's system could

[1] It consists of a tape, a set of symbols, a scanner to read and write to the tape and an internal state together with a table mapping the current state and the symbol recently read to a new state, a symbol to write and one of three actions: do nothing, go left, go right

[2] In computer science, it is usual to talk about Turing machines with infinite tapes.

be implemented by electronic circuits. He was also responsible for a major breakthrough: Allowing the quantification of information. Assumed that Boole was right and human thinking was based on Boole's logic, thinking could now be automated. Unfortunately, Boole was wrong in this sense.

Descartes, Boole and Turing are representatives of different efforts to explain mind. The essence of their views and an evaluation of these will be the subject of this paper.

2 An Overview of Approaches to Explain Mind

The approaches to explain what mind is and how it might work are numerous. Although there is no big picture containing all approaches and their relationships amongst each other, there are some points of view encountered often.

A fundamental distinction can be made between treating the issue bottom–up or top–down. Representatives of the bottom–up fraction start at very small constituents of the brain[3] which form the higher functions while the top–down people see the brain or the mind as a whole which needs to be decomposed into smaller parts. I will discuss theories of both kinds.

Since nowadays there are a lot of different sciences concerned with the mind and the brain I had to narrow the field a bit. Therefore I will almost leave out everything bothering with the biological aspects of the brain and mostly concentrate on the more abstract theories (which nevertheless sometimes have their foundations in biologic principles). I also will neither include the various psychological nor the philosophical views of mind.

What follows is an (incomplete) tour through some of the most important views of mind and brain.

2.1 Dualism vs. Materialism

Dualism and materialism are contrary views of the relationship between mind and brain. The dualists consider the mind to be a rather metaphysical phenomenon which is completely separated from the brain and "somehow" interacts with the body (see figure 1), Descartes therefore was a dualist. They solely rely on the observation of the effects to explain the processes. But this procedure is inherently flawed as Cummins (2000a: 3) explains: "Inferred

[3] like neurons and the chemical processes involved therein

$$\frac{\text{mind}}{\text{brain}} = \frac{\text{immaterial}}{\text{material}}$$

Figure 1: The dualist view of mind.

mechanisms, if they are to have any explanatory or predictive value, must be, to some extent anyway, understood independently of the effects they are rung in to explain."

The materialists on the other hand simply identify the mind with the brain which unfortunately does not help us on, since "... even if we are convinced that the mind is the brain, or a process going on in the brain, physical observation of the brain seems to give us data in the wrong vocabulary: synapses rather than thoughts." (Cummins 2000a: 4)

Both the dualists and the materialists have to rely on introspection — observing oneself while carrying out different tasks (like perceiving something, solving a problem or learning). But introspection in turn is also flawed for three reasons:

1. We do not get objective but only very subjective descriptions. Although there is strong evidence that there are major similarities between all human's minds it might not be the case.

2. The process of introspection will certainly influence the processes we are trying to introspect since it uses the same "machinery".

3. Introspection has it's limits: We simply cannot observe single synapses while firing. Therefore the result of introspection will be incomplete.

Although inherently flawed, introspection was also used by other approaches.

2.2 Structuralism

The structuralists believed that they would be able to figure out the basic parts of mind by introspection:

> The most significant introspectionist program in the United States was structuralism. The project was to discover the fundamental and introspectively unanalyzable elements of consciousness, determine their origins in sensation, and to formulate the principles of combination whereby these elements are synthesized into the complex and familiar experiences of ordinary life. Every compound mental or process state was to be explained compositionally, the characteristics of the whole derived from the characteristics of the parts and mode of combination.
>
> (Cummins 2000a: 4)

Because of the improper introspection technique and the lack of other means to analyse mental events or processes, this approach failed.

But structuralism was not defeated yet. A second effort tried to "explain the elements of consciousness as responses to perceptual stimulation" (Cummins 2000a: 5). The trouble was that although this is not based on introspection, the the subject's reports can never be validated. If there is a response which does not fit into the theory or law, it cannot be decided whether the subject misdescribed his / her sensations (by accident or even intentionally), he / she is abnormal or even both. The same applies to responses fitting to the hypothesis — they simply cannot be verified. Hence, only the theory can be used to verify it. Or, as Cummins puts it: "Once introspection is disqualified, we have no access to sensation intensity other than the very law that is supposed to explain it." (Cummins 2000a: 5)

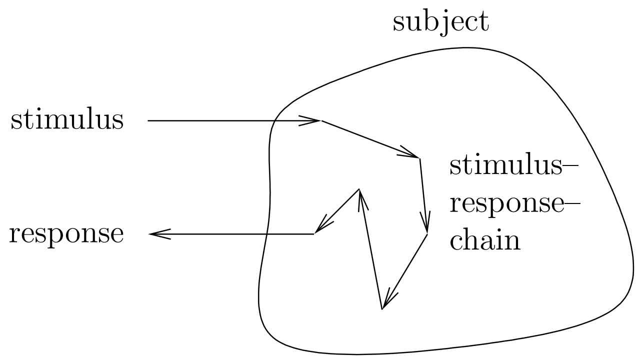

Figure 2: The behaviourist view — there is no mind.

2.3 Behaviourism

The behaviourists chose a rather radical position. Instead of trying to explain the mind, they simply eliminated it: ”Behaviourism seeks to avoid the problem about observing the mind by eliminating the mind from psychology.” (Cummins 2000a: 6) The behaviourist sees any responses from subjects as direct reactions to stimuli via stimulus–response–chains (see figure 2). The death of behaviourism was caused by the inability to explain the acquisition of novel behaviour and therefore learning.

Even if the behaviourists were successful, they would only have specified psychological behaviour instead of explaining it.

2.4 Gestalt Psychology

Gestalt psychology is a rather unusual flow. It does not bear a central, cohesive theory of cognition or behaviour but is instead defined by it’s opposition to behaviourism and structuralism. Gestaltists produced a lot of counterexamples for behaviourism and structuralism. They also believed that high-order psychological phenomena could not be decomposed, since ”an adequate explanation of intelligent behavior required reference to internal states and highly integrated cognitive structures." (Cummins 2000b: 12)

Apart from that they also presented strong evidence

> that our perceptual capacities shape our knowledge about the world. They showed that our visual system is capable of *augmenting and organizing* stimulation in reliable ways. A classic example is that of the phi phenomenon ... If light bulbs are lined up in a row, and each one in succession is quickly turned on and off, one sees an illusion of movement down the line of bulbs. In fact, nothing in the physical environment is moving [apart from plenty of electrons, electic fields etc. — at least nothing directly observable], but the pattern of light stimulation is "interpreted" by our visual system as movement.
>
> (Cummins 2000b: 13)

Cummins continues to point out that this illusion was problematic both for behaviourists and structuralists. The illusion is neither explainable in terms of responses nor decomposable by introspection. "Psychological phenomena such as these seemed to indicate that 'the whole is greater than the sum of its parts' ..." (Cummins 2000b: 13). It made the Gestaltists believe that perception is an active, constructive process. They also found out that problem solving is a top–down process and goal–oriented, since "subjects formed *plans*, generated *goals*, and developed *strategies* based on acquired *knowledge*." (Cummins 2000b: 14)

2.5 Representationalism

The base of representationalism is the assumption that "cognition is to be understood as disciplined transformation over states whose primary function is the representation of information relevant to the cognitive capacity being exercised." (Cummins 2000a: 172) Boole was clearly a representationalist. The representational view is not an approach of it's own but rather part of others and fundamental there. Representationalism brings along a rather fundamental problem called Leibniz' Gap (or Leibnizian Gap) which will be discussed in detail in section 3.1 (page 17).

2.6 Functionalism

Another fundamental view is functionalism which solves the problem that functions cannot, in general, be read off from form. The consequence is that there is no immaterial soul missing but a functional analysis of the brain and it's component structure and processes. Cummins describes functionalism as "the central idea … that mental concepts specify their instances in terms of what they do — in terms of functions — rather than in terms of their intrinsic structures." (Cummins 2000a: 6)

2.7 Computationalism

The main hypothesis of computationalism is that mental processes are computational processes. The brain is regarded as a computer running the "mind" software. Computationalism inherits both functionalism and representationalism and is a top–down approach. Viewing mind as the result of a computational process allows the to draw reverse conclusion that computational processes can "produce" mind.

Computationalism also assumes that the cognitive functions are actually computable[4]. Let me cite Cummins again for a nice summary of the consequences of computationalism:

[4] There exist non-computable functions, for example the function which, given an algorithm, determines whether this algorithm will terminate. As far as I know there is neither evidence nor counter evidence for the computability of cognitive functions, though I got the impression that cognitive functions are not computable. I will further explain my objections in section 3.3.

> The idea that the mind is essentially a functionally specified computational process running on the brain provides a bridge over Leibniz's Gap (functionalism), a supply of mental mechanisms with precisely specified properties (anything you can program), and *medium independence*: the possibility that though can exist in a non–biological computer, and hence can be investigated in the computer lab as well as in the psychological lab. It was a powerful vision. And though it shows signs of fading today, it was, and in some respects, continues to be, a hugely prolific vision, fueling the initial birth and development of what came to be called cognitive science.
>
> (Cummins 2000a: 7)

2.8 Connectionism

Connectionism also uses the ideas of functionalism and representationalism and shares some with computationalism. A connectionist program is made of a lot of similar units which are modelled after neurons. The result is called artificial neuronal network. Such a network (see also figure 3) consists of a number of nodes (usually of the same type) which are organised in layers. A typical neuronal network is a so–called feed–forward network where the output of each node is only connected to the input of nodes in the next layer or in the same layer. Recurrent networks allow all sorts of connections and might not have a clear layer distinction.

The connections between the nodes ”transport” an activation potential which is modified by a weight factor (a real number) and the nodes usually perform a quite simple task: They sum up all incoming potentials and calculate their output depending on that (negative values are allowed). It is important to note that the single unit does not know or depend on where the input came from. Input is sent to the network by setting the activation of the input nodes while output is essentially reading the activation of the input nodes after the input activation has spread across the whole network.

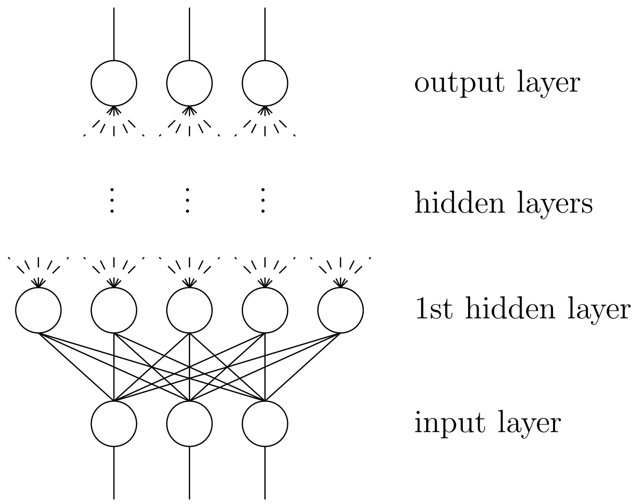

Figure 3: Structure of a feed–forward artificial neuronal network

In practice, mostly feed–forward networks are used since there is an efficient algorithm (backpropagation) to train these networks. Training in this way is achieved by starting with random connection weights, setting an input activation and comparing the output with the desired output. The backpropagation algorithm then adjusts the connection weights backwards (from output to input layer) to minimize the error. The process is then repeated with more training input. Backpropagation is a kind of supervised learning, that is, the connection weights are adjusted by processes external to the network. There are also non–supervised methods of learning like Hebbian learning or Kohonen maps.

Recurrent networks are more complicated and therefore not used so much. The problem with recurrent networks is that they might oscillate (it is therefore complicated to decide when to read the output) and backpropagation cannot be used for training. These networks are also far more difficult to understand — there are no formal and systematical means of analysis yet.

Artificial neuronal networks share some properties with the brain. At first, there is no central "place of representation", but the information is stored across the network (in connection weights and in activation patterns). This feature makes them virtually inunderstandable by humans. Second, neuronal networks degrade graceful. "'Graceful degradation' refers to the fact that degraded input, lesion connections, or damaged units, typically do not bankrupt a network, but lead to impaired but interpretable performance, with the degree of impairment depending on the degree of damage." (Cummins 2000a: 175)

Artificial neural networks are nowadays used in practice. They are able to compute very complicated functions[5] and can be successfully applied to pattern recognition, control of autonomous agents and other tasks too complicated for ordinary programming.

2.9 Dynamicism

Dynamicism is an approach which is similar to and different from computationalism at the same time. The main assumption is that cognitive processes are best described as the behaviour of dynamic, physical systems. The similarity to computationalism is that cognitive processes are considered to be the result of some physical device. The difference is that computationalism is strongly bound to the computer metaphor and therefore rather abstract while dynamicism only allows physical processes.

The point is that there are pretty simple physical systems carrying out tasks which are very complex for computers since they do not fit into the input–calculate–output principle. A simple example is a centrifugal governor as used

[5]It is often easier to write a program using a neural network than to write it in the ordinary way. Using a neuronal network has the disadvantage that after all the programmer does not really know how the task is accomplished.

in early steam engines. It operates in real time, regulating steam pressure according to engine load to keep a constant speed. This "simple" system does not measure–calculate–regulate, but, in a way, it does all at once, continuously. And although it is known that human vision, for example, "runs" at around 25 frames per second, cognitive processes are considered to run continuously and there is no central clock pulse synchronizing them.

Van Gelder (1996: 427) explains the problems more thoroughly:

> The heart of the matter is this. At all times, the speed of the engine influences the angle of the arms [of the governor]. Yet the arms are directly connected to the throttle valve, which controls the flow of steam to the engine. Thus, at all times, the angle of the arms is also influencing the speed of the engine. The quantities are thus simultaneously determining the shapes of each other's changes. There is is nothing mysterious about this relationship; it is quite amenable to mathematical description. However, it is much more subtle and complex than the standard concept of representation — very roughly, one thing "standing in" for another — can handle. In order to describe the relationship between arm angle and engine speed, we need a framework that is *more powerful*, with respect to this kind of situation, than talk of representations. That framework is the mathematical language of dynamics [Maxwell equations]; and, in that language, the two quantities are said to be *coupled*.

Later, Van Gelder (1996: 446) justifies the dynamic approach:

> One central fact about natural cognitive processes is that they always happen *in real time*, which means not merely that, like any physical process (including ordinary digital computation), they occupy some extent of actual time, but that details of *timing* — duration, rates, rhythms, and so on — are critical to how they operate in real bodies and environments. . . . dynamics is all about how processes happen in real time, whereas timing details are in a deep sense extrinsic to computational systems[6].

[6] Computational systems are always independent from time since there is no way to measure the time an abstract algorithm will require — the efficiency of algorithms is measured in other terms like amount of input data. Therefore one cannot say how long a process will take until one knows the exact hardware it will run on. However, computational theory is not about hardware but about abstract processes.

Dynamicism is an interesting view of mind. It's aim is to identify the control loops producing the observed behaviour. As dynamicism is pretty new, it has still to stand the test.

2.10 Subsumption Architecture

The subsumption architecture also uses some ideas from computationalism but differs from it notedly. It was brought into play by Brooks in 1986:

> Rodney Brooks, ... one of the founders of this new field, argued that the traditional approach to AI[7] was fundamentally flawed. He maintained that all of AI's ideas concerning thinking, logic, and problem solving were based on assumptions that come from our own introspection, from how we see ourselves. He suggested that we drop these assumptions, do away with thinking and reasoning, and focus on the interaction with the real world. ... He suggested [in a seminal paper in 1986] that intelligent behavior could be achieved using a large number of loosely coupled processes that function predominantly in a asynchronous, parallel way. (Pfeifer 1999: 25)

The subsumption architecture is astonishing simple and quite straightforward. It is an architecture of distinct layers which separate the "*activity* producing subsystems" (Brooks 1991: 403). Each of the subsystems individually connects sensing to action. Therefore, neither sensory input nor action output needs to be represented in a central, complex and clumsy manner. Brooks calls such systems "behaviour–based systems". These systems neither need a "perception system" nor a "central system" nor an "action system". They are built incrementally, starting from basic layers pursuing simple functions, adding more layers if higher functions are needed: "We wire finite state machines together into layers of control. Each layer is built on top of existing layers. Lower layers never rely on the existence of higher–level layers." (Brooks 1991: 410, caption of figure 15.2) Higher layers are able to influence parts of

[7] Artificial Intelligence, basically the "product" of computationalism.

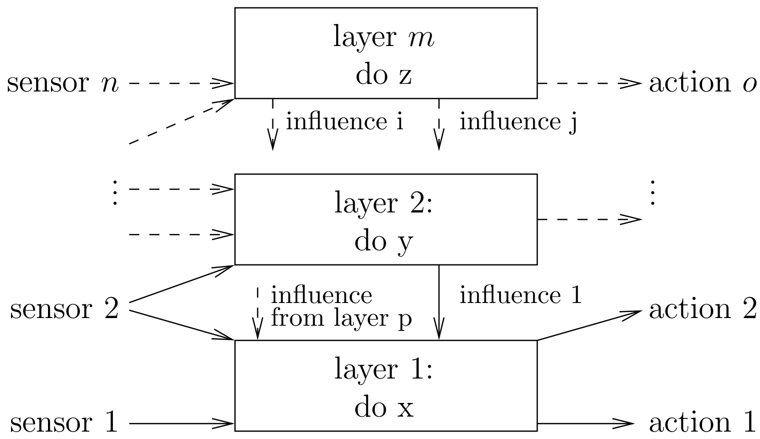

Figure 4: The rough structure of a behaviour–based system.

lower layers (see figure 4). All finite state machines run in parallel which is similar to how the human brain works.

Brooks also built several "Creatures" (as he calls them) which showed very complex behaviour. The funny thing is that these robots behaved in such a complex manner that an observer would certainly impute complex internal processes which in fact did not exist explicitly:

> The point of Herbert [one of the Creatures, driving around and picking up soda cans with it's arm] is two-fold:
>
> - It demonstrates complex, apparently goal–directed and intentional behavior in a system which has *no long–term internal state and no internal communication*[my emphasis]; and
> - It is very easy for an observer of such a system to attribute more complex internal structure than really exists — Herbert, for instance, appeared to be doing things like path planning and map building, even though it was not.
>
> (Brooks 1991: 413)

3 Unsolved Issues

Some of the weaknesses of the different approaches have already been pointed out. They are partly related to more generic issues which will be discussed in this section.

3.1 Leibnizian Gap

The Leibnizian Gap (illustrated in figure 5) is probably the most fundamental problem of cognitive science. I cite Leibniz by means of Cummins (2000a: 4):

> Here is Leibniz's formulation of the Gap:
>
>> It must be confessed, moreover, that perception, and that which depends on it, are inexplicable by mechanical causes, that is, by figures and motions. And, supposing that there were a mechanism so constructed as to think, feel and have perception, we might enter it as into a mill. And this granted, we should only find on visiting it, pieces which push one against another, but never anything by which to explain a perception. This must be sought, therefore, in the simple substance, and not in the composite or in the machine. (Leibniz, *Mondology* sec. 17)
>
> There is, as Leibniz points out in this famous passage, a gap between the concepts we use to describe the mind, and those we use to describe the brain.

The Leibnizian Gap is hard to cross. Especially the "simulation ↔ duplication" problem spoils the broth of computationalism. The problem is that, in theory, it should be possible to create a simulation of a human brain. When the simulation is run, it will show human–like behaviour, intelligence etc. But it is still only a machine. It is a simulation of intelligence, not real intelligence itself, one could argue.

mind ↔	brain
information ↔	formal symbols
information processing ↔	symbol manipulation
semantics ↔	syntax
duplication of mind ↔	simulation of mind

Figure 5: Dichotomies characteristic for the Leibnizian Gap.

To defend this argument, one either needs to declare that the human brain is simply a biological machine or that there is something very special, metaphysical going on in the brain.

It is also possible to deny that there is a difference between simulated and real intelligence. The machine running the simulation should then be granted human rights and it must not be halted or turned off since that would be murder.

The dualists claimed that mind would be immaterial and made the Gap part of their theory. The materialists are severely affected by the Gap, because, after all, there seems to be no way to avoid it and still identifying mind with brain. Computationalists narrowed the gap a bit by stating that the processes are computational and therefore do not depend on the underlying realisation. They consequently stayed on the "mind" side of the Gap. Connectionists on the other hand did the splits across the Gap by retaining to the computational view and using concepts found in the brain at the same time, though they still suffer from the "simulation ↔ duplication" problem.

The biggest consequence of the Leibnizian Gap is a question: How to figure out whether a machine actually thinks? It is too simple to declare that any machine exposing intelligent behaviour is just simulating intelligence. The question is hard to answer since there is no single acknowledged definition of intelligence. The famous Turing Test[8] is empirical and only examines the behaviour of the system. Currently, this seems to be the only way to ascribe intelligence to a machine.

3.2 The Frame Problem

The Frame Problem is purely computationalist and characteristic for representational approaches. Pfeifer (1999: 65) sums it up: "The central point concerns how to model change (Janlert 1987): How can a model of a continuously changing environment be kept in tune with the real world?"

If we assume that we have a (probably huge) set of logical propositions describing the world and want to infer on it, we suffer from "combinatorial explosion", since we need to find matching rules for successful inference. The more rules there are, the more time it will take to find matching rules. And typical inferences do not only include two or three rules — there might be hundreds involved in figuring out that an apple is in fact eatable[9].

Modelling changing environments usually leads to an explosion of propositions — things get extremely complicated if they are able to change. A lot of workarounds have been suggested but trying to represent the world with

[8]The Turing Test is set up as follows: There is an interrogator C in one room and two subjects A and B in another room. C is connected to A and B by some tele typing machine. The task for C is to figure out which of A and B is a human and which is a machine. C is allowed to ask any questions. The machine will try to fool the interrogator into thinking it is human while the human will try to give hints to the interrogator. See also Turing (1950). Since this test only focuses on Natural Language Processing, extended tests have been proposed, for example involving video and speech.

[9]There are more complicated problems, e. g. in connection with perception. Imagine a table with a cup and a ball on it. If a robot moves around the table, the visible relation between the cup and the ball changes although nothing on the table changed. Representing this in some sort of calculus would require several propositions and several rules which would have to be applied to every object known to the robot.

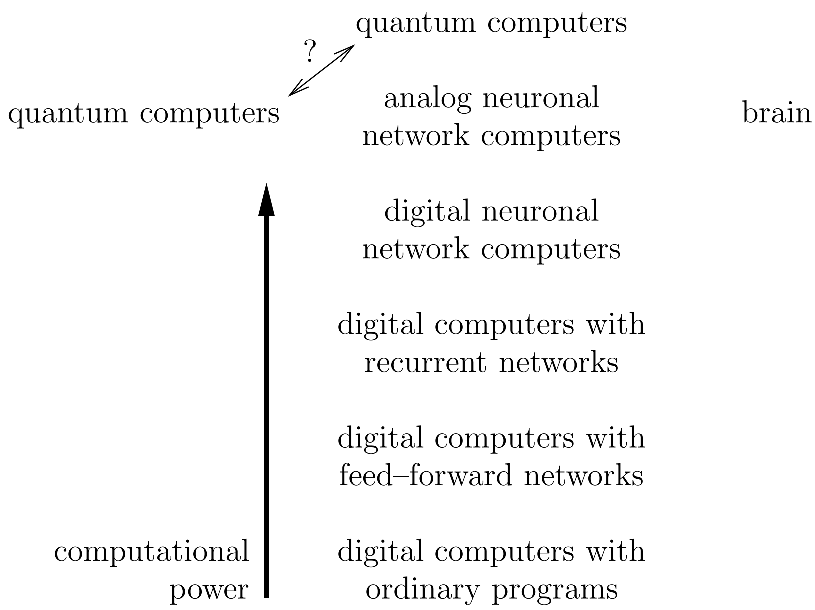

Figure 6: Different levels of computational power.

propositions seems a dead end to me. This approach appears to be inherently flawed and throws a bad light on computationalism.

3.3 Structural Differences Between Brains and Computers

It looks like there is an invisible, fundamental border which prevents successful and thorough analysis of mind. The computational paradigm is probably the most powerful so far. Alas, mind seems to be significantly more powerful. Artificial neuronal networks are plain compared to the brain. And even those simple networks are not yet fully understood — recurrent nets still resist sound understanding.

I'd like to sketch several levels of computational power (see figure 6) and explain what the differences are. This scheme is rather empirical although I present some evidence here. When I'm talking about computational power here, this is no strong scientific term and it is different from the notion of complexity classes (P, NP etc.) used in computer science.

The lowest level of computational power is equal ordinary digital computers with ordinary (classical) programs. There is one central processor, some memory and some input / output systems. The computer processes one instruction after another. The programs are written by humans and contain explicit instructions of what to do. Adding more processors only increases the speed and therefore allows more data to be processed in the same time — it does not change the big picture.

On the next level are computers simulating feed–forward artificial neuronal networks. Now, there is only a relatively simple program which is not directly responsible for solving the task but only for interfacing an artificial neuronal network with the outside and performing the calculations necessary for the operation of the network. The programmer only knows how to program the network and how to train it while the main task is solved by the network. Neuronal networks are able to perform tasks which would be extremely complicated for ordinary programs (thus, for programmers). Feed–forward networks are still bound to the input–calculate–output paradigm.

I suppose recurrent artificial neuronal networks to be more powerful than feed–forward networks since they introduce a timing component which resembles the patterns found in living systems closer. They might also produce phenomena similar to short time memory (parts of the network might start to oscillate and therewith keeping some information). Unfortunately, I don't know of any experiments with very large networks.

Artificial neuronal networks suffer from the limitations of digital computers. The network nodes cannot be simulated in parallel but change needs to be calculated serially, one node after another. For real applications, only networks up to a certain size are feasible because of time and memory constraints. No artificial network with approximately the complexity of the brain has been built yet.

I'm not sure whether a new level of computational power is reachable by building artificial neuronal network machines. I guess so, since current state of the art computer chips are not simulated with computers any more but using special hardware[10]. Unfortunately, there are a lot of technical issues with building such computers, wiring, for example, cannot be arbitrary but is constrained in certain ways.

The brain seems to be a palladium of chaos (anyone doubting that should take a look at an EEG). Chaotic systems have been and are subject to scientific study — a very small change might affect the whole system. It is therefore questionable whether an artificial neuronal network machine is as powerful as the brain if it only uses discrete values for calculations. It might be necessary to switch to an analog machine to gain computational power equal to the brain. But then, given enough capacity, we have a brain simulator. (I probably should mention that there actually might be aspects of the brain which cannot be simulated, for example the entity which is usually referred to as soul.)

Recently, a new paradigm of computing appeared: Quantum computing. Quantum computers are both difficult to create and difficult to program. They are based on principles of quantum physics and use superpositions to store data (that is, the QuBits are both zero and one at the same time, the real value only emerges by measuring). The essence of quantum computers is that they "calculate" all possibilities at once. If we had 8 QuBits (current quantum computers

[10]The circuit to be tested is transferred into the emulation hardware, which establishes appropriate wiring . The emulation hardware can then, in a sense, be used like the circuit it is emulating, which allows the emulated circuit to be evaluated and debugged.

have three or four), they represent all 256 possible states at once. Completely new algorithms need to be developed for quantum computers since the result of a computation needs to be filtered out of all possibilities. Theoretical computer science already discusses quantum computing and it is suspected that this is a magnitude more powerful than ordinary digital computing. I could only guess whether quantum computers are more powerful than the brain or artificial neuronal networks, but I won't. Maybe, time will tell us.

3.4 Social Functions of Intelligence and Evolutional Aspects

There is an interesting question left which has not been discussed yet: What is intelligence good for? Evolution does not produce luxury, there is no capability of any creature without a purpose, because superfluous capabilites do not increase fitness, they decrease it instead. Evolution only rewards changes which increase fitness. Therefore, intelligence has to fulfil a purpose which increases the fitness of individuals. Humphrey (2000: 513) says that "it is not her[nature's] habit to tolerate needless extravagance in the animals ...: superfluous capacity is trimmed back, new capacity added only as and when it is needed."

The intelligence of animals has been explored a lot, but it is seldom shown or asked how this intelligence contributes to the fitness of a particular individual. "What advantage is there to an anthropoid ape in being able to recognize its own reflection in a mirror (Gallup, 1970)?" (Humphrey 2000: 513)

Intelligence is only useful in an intellectually challenging environment. As a matter of fact, the life of the most intelligent species seems to be rather easy and does hardly show any demand for intelligent behaviour:

> During 2 months I spent watching gorillas in the Virunga mountains I could not help being struck by the fact that of all animals in the forest the gorillas seemed to lead much the simplest existence — food abundant and easy to harvest (provided they knew where to find it), few if any predators (provided they knew how to avoid them) — little to do, in fact (and little done), but eat, sleep, and play.
>
> (Humphrey 2000: 515)

This sounds like a contradiction — anthropoid apes have been shown to possess impressive intelligence which does not seem to be used in everyday life. What obviously is used and useful in practice is knowledge. Knowledge of where to find food, how to avoid predators and how to hunt successfully. Knowledge gives a real competitive edge.

Knowledge also leads to a sensible explanation for the need of intelligence, because knowledge has to be passed on from generation to generation. A social community is the best place to pass on knowledge. But a social group also bears immense difficulties:

> The life of social animals is highly problematical. In a complex society, such as those we know exist in higher primates, there are benefits to be gained for each individual member both from preserving the overall structure of the group, and at the same time from exploiting and out–manoeuvring others within it . . . .
>
> (Humphrey 2000: 516)

He further suggests that the society acts like a school for young animals and that it serves both to increase the 'school leaving' age and to keep older animals as experienced teachers.

Such a society produces a lot of intellectual challenges. "The presence of dependents (young, injured, or infirm) clearly calls at all times for a measure of tolerance and unselfish sharing . . . Squabbles are bound to occur about access to these scarce resources [subsistence materials and sexual partners]"

(Humphrey 2000: 517). The individuals need to maintain a complex network of relationships.

At some point, intelligence becomes a factor of fitness, viz if it increases social success and social success is connected to biological fitness. Now, intelligence is subject to evolution and will probably increase over time.

If we follow this argument, intelligence simply emerged to allow the development of complex social communities which in turn increases the fitness of the individuals participating in them.

4 Conclusion

After having discussed several views of mind, it is still not obvious how mind works and it is therefore difficult to find a suitable conclusion. The exploration of mind is an ample field. There are a lot of different approaches, both promising and disappointing ones. It does not look like the mind and associated concepts (perception, cognition, intelligence) will be understood tomorrow or the day after tomorrow.

During preparation of this paper, I always had the impression that something very important was still missing from all theories. Maybe the human mind is "designed" in a way that prevents it from understanding itself.

Cognitive science will probably not cease to be an interesting field of research. It is very interdisciplinary and will certainly produce more surprising ways of access to min in the future.

Bibliography

Bateson, R. P. G. / R. A. Hinde, eds. (1976). *Growing Points in Ethology.* Cambridge: Cambridge University Press.

Brooks, Rodney A. (1992). "Intelligence without Representation". In: Haugeland (1997: 395-420).

Cummins, D. D. (1988). "A History of Thinking". In: Sternberg / Smith, eds. (1988).

Cummins, Denise Dellarosa (2000). "A History of Thinking". In: Cummins / Cummins, eds. (2000: 8-19).

Cummins, Robert / Denise Dellarosa Cummins, eds. (2000). *Minds, Brains and Computers: The Foundations of Cognitive Science An Anthology.* Blackwell Philosophy Anthologies. Malden etc.: Blackwell Publishers.

Gallup, G. G. (1970). "Chimpanzees: self–recognition." *Science* 167: 86-7.

Haugeland, John, ed. (1997). *Mind Design II.* Cambridge etc.: The MIT Press.

Humphrey, N. "The social function of intellect" In: Bateson / Hinde, eds. (1976).

Humphrey, Nicholas K. "The Social Function of Intellect" In: Cummins / Cummins, eds. (2000: 513-522).

Minsky, Marvin (1985). *The Society of Mind.* New York: Simon & Schuster.

Minsky, Marvin L. (1990). *Mentopolis.* Translated by Malte Heim. Stuttgart: Klett–Cotta.

Pfeifer, Rolf / Christian Scheier (1999). *Understanding intelligence.* Cambridge etc.: The MIT Press.

Shannon, C. E. (1948). "A mathematical theory of communication". *Bell Systems Technical Journal* 27: 379-423, 623-56.

Sternberg, R. J. / E. E. Smith, eds. (1988). *The Psychology of Human Thought.* Cambridge: Cambridge University Press.

Turing, Alan (1950). "Computing Machinery and Intelligence". *Mind* 59: 433-460.

van Gelder, Timothy (1996). "Dynamics and Cognition". In: Haugeland, John, ed. (1997: 421-450).