Computers, Brains and Minds:

Where are the ghosts within the machines?

Anja Stemme [a,*], Gustavo Deco [b],
Stefan Büttner-von Stülpnagel [c]

[a] *Institute for Biophysics, University of Regensburg, Germany*

[b] *ICREA Research Professor, Computational Neuroscience, Universitat Pompeu Fabra, Barcelona, Spain*

[c] *Institute for Philosophy, Univsersity of Potsdam, Germany*

Abstract

This work opens a new perspective in the debate regarding the relationship between brain and mind. It puts the central question whether in our days, dominated by a scientific view of the world, we really do have an adequate idea of what could be actually meant by terms like "mind", "self" or even "perceptions". Based on the consideration that any "non-material" or "mind" stuff can never be able to move particles around the world we found ourselves forced to believe in a physical monism as the premise for all plausible theories of mind. However, now we face different but still persisting kinds of problems: *Why* do we have "phenomenal" states? Do we really have them at all? How could phenomenal states be related to neuronal states? The still ongoing heavy debates about these issues demonstrate at least one thing: So far there is no commonly accepted explanation detectable at the horizon.

To understand all scientifically possible relationships between the world, the brain, perceptions and what some people tend to call "self" we suggest the support of a "Complete Computer Analogy" (CCA). The CCA takes all relevant aspects of human and "computer" life into account: Neuronal states (or binary states for the computer), phenomenal states (or the screen for the computer) and the connections brains and computers have to "the world": Exemplarily visual and auditory interfaces i.e. eyes and ears for our brains, a scanner and a microphone for the computer. Empirical evidence (e.g. the famous "color phi" experiment) demonstrates in the first instance the necessary differentiation between phenomenal states, neuronal states and the "outside world" though they are of course correlated, as are the screen, the computer and "his" world. Further on we are able to consider different options for the realization of the "self" within this analogy and according empirical hints (i.e. neurological diseases, empirical findings around the issue of "free will" and "attention"). These investigations finally reveal an astonishing aspect: It is possible to understand an interactive substance dualism based on a scientific view of the world but beyond the idea of a ghost moving particles around the brain.

20 February 2009

1 Perspectives in the age of science

Following a scientific perspective we start with the question of potential relationships between the world, our brains and our perceptions. However, a substantial amount of empirical evidence demonstrates the necessity to differentiate, first of all, between our perceptions and the "outside world". Any optical illusion as the "color phi" phenomenon, for example,[1] shows us that we are confronted with two different things: Two circles "outside" but only one moving circle in my perception.

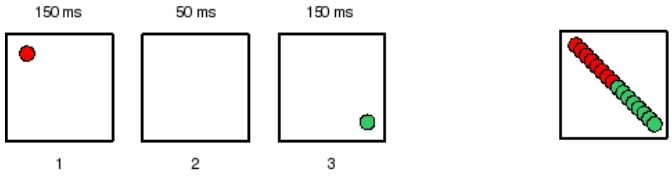

Fig. 1. The color phi phenomenon. Left: Two dots are presented on a computer screen. Right: The timing and spatial conditions of the presentation lead to the perception of a single moving dot.

There are two possible explanations for this phenomenon:

(1) We consider the perception of one moving circle as a kind of "story" told by the brain rather than a "true" perception.[2] However, as the neuronal explanations for color phi come very close to those neuronal explanations for the perception of a really moving dot (compare Stemme and Deco, 2008), we face the problem that then the whole world would have to be considered as a simple "story" told by the brain. Hence, we are to admit that we possibly do not know *anything* about this "outside" world. But our observations about this world (constituting the physical monism) constituted the premise for our theories of mind. With the falling premise the theories based on this premise consequently have to be discarded as well.

(2) We separate optical illusions and with them perceptions and a range of other phenomena from the "outside" world and consider our brains in the first instance as a kind of interface. To understand this relationship we suggest the support of a more *complete* computer analogy than previously used within the mind/brain debate.[3]

* Corresponding author. Email: anja.stemme@gmx.de

[1] An example initially raised by Dennett (1991) in this debate

[2] Dennett (1991)

[3] compare as well Stemme (2008)

2 Computers and Brains

When considering computer analogies for the purpose of understanding the mind/brain relationship we are to consider computers as autonomous actors in *their* environment as humans are considered as autonomous agents as well. One aspect of the primary environment of computers is of course constituted by other computers with which they communicate.

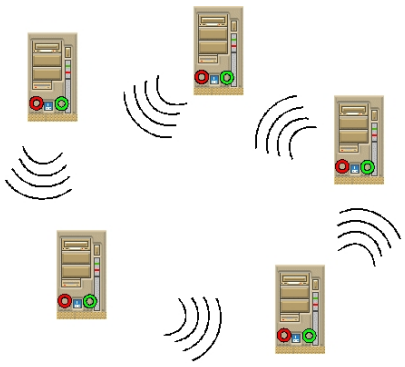

Fig. 2. Autonomous actors.

Within this communication they might debate as well about rather philosophical issues as far as the corresponding questions can arise *within* their communication and environmental access. For the purpose of this access we might consider them connected to something like "an outside world" via various interfaces. One of these interfaces might be a kind of simple scanner transforming pictures or "light waves" into the zeros and ones with which the computer operates. On this scanner we assume the presentation of "the outside world", in our example the two circles. Hence the binary codes reaching the computer and the following processing of these codes represent the neuronal states within our brains. To illustrate the *perception* of a single moving dot and to differentiate it from the two circles "outside", we now use the screen.

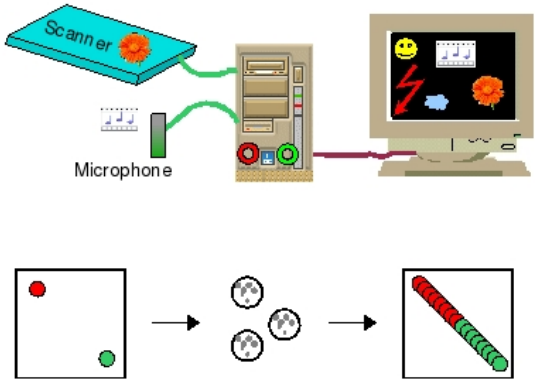

Fig. 3. "Color Phi" in a computer analogy.

We do not experience neuronal activities, we experience "the outside world" and somehow "inner" feelings. "Inner" feelings might be correlated with neu-

ronal activities but even if considered identical to those activities they are - somehow "additionally" - experienced. This experience is - as far as we can observe - not experienced by the neurons themselves and is, again following our observations, of no causal relevance for further neuronal activies. Also, between us and the outside world there are a couple of neurons. We do not have access to the "outside" world directly. An aspect responsible for a vast amount of philosophical literature. Thus we use the computer screen to illustrate these experiences which are in either theory of mind - based on a physical monism - of no causal relevance. As is the screen for the activities within the computer. The computer, considered as an autonomous actor, is determined by the "binary" activities within its cabinet as we consider us to be determined by the neuronal activities within our brains. For either activities, binary or neuronal, phenomenal states are unnecessary though existing (e.g. Color Phi).

3 Ghosts and maschines

Is there any advantage of this, lets call it "complete computer analogy" as it considers not only the equivalence of neuronal and binary or "computational" states but illustrates as well phenomenal states and their relationship to the brain and the world?

One of the major advantages is the ability to ask for the existence and role of a kind of "computer user". We are able to think about this computer user as being someone or something who experiences the "outside" world, someone who has perceptions and who experiences as well as "inner" feelings potentially generated solely by the brain.

We are able to reinvestigate the old question regarding the source of attention: Is it possible that we are able to direct our attention to certain issues amongst our experience? Is it possible that as soon as we direct our attention to our left hand or a forthcoming movement of our left hand, for example, is it possible that this attention somehow leads to an increased neuronal activity which would imply that a rather small object more at the border of the screen now "jumps" to the middle of the screen, in the focus of our attention? Is it possible that an increased neuronal activity following such changes in the focus of attention is measurable in terms of the famous readiness potential, for example? [4]

Thus now we are able to think about potential equivalents of the "self" in the CCA: Has the self to be considered as an element on the screen something generated by the binary states within the computer by a separate "process"? Or is the self to be considered as the sum of all experiences, the screen as a whole? Or is it possible that the self is "something" in front of the screen? Considering

[4] Libet (1985)

a range of neurological diseases we are at least not able to exclude this last option. Moreover, a range of phenomena allow far more plausible *neuronal* explanations if we do not have to search for the source of attention within the brain. [5]

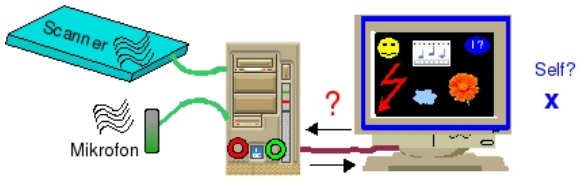

Fig. 4. The Complete Computer Analogy. Given the screen for the illustration of the relationship between phenomenal states, brains and the "outside" world, we are able to think about further relationships even including a formerly simply mysterious substance dualism.

However, in this last case we would have to consider our screen as a kind of "touchscreen" and the connection between the screen and the computer to be bidirectionally. Hence in this last case we are entering a domain which previously had to be considered contradicting a scientific view of the world. However, using the computer analogy we now are able to think beyond a funny ghost moving particles around the brain.

Almost all interactional theories searched for a certain *location* of or physical mechanism for this interaction. However, considering the circumstance that *if* we think dualism as a dualism of two very different "substances", should we consequently not look for an interface between them somewhere in the middle of the "way"? Considering phenomenal states in their relationship to the brain and the world with the help of the CCA enables us to consider them as a kind of "translation" provided by the neurons for some reasons or other. And these phenomenal states go somehow beyond pure "material" stuff. Obviously the brain does something more than absolutely necessary under the premise of a physical monism, something that goes beyond itself, hence they form a candidate for one part of the "interface". And *attention* is at least a hot candidate for the interaction in the "opposite" direction.

[5] E.g. the examples Daniel Wegner outlined concerning the "illusory" will, Wegner (2002), compare as well Stemme (2008)

References

Dennett, D. C., 1991. Consciousness explained. Little, Brown and Company, Boston, Toronto, London.

Libet, B., 1985. Unconscious cerebral initiative and the role of conscious will in voluntary action. The Behavioral and Brain Sciences 8, 529–566.

Stemme, A., 2008. Deus et machina: Der Geist und die Naturwissenschaften. PhD thesis, Ludwig Maximilian University Munich.
URL `http://edoc.ub.uni-muenchen.de/9413`

Stemme, A., Deco, G., 2008. Neuronal and cortical dynamical mechanisms underlying brain functions. Elsevier Handbook of Cognitive Science, 221–242.

Wegner, D. M., 2002. The illusion of conscious will. MIT Press, Cambridge, London.