Some Thoughts on Computation and Simulation in Cognitive Science

Markus Peschl (*) and Matthias Scheutz (**)

(*) Department of Philosophy of Science and Social Studies of Science
University of Vienna
Sensengasse 8/10
A-1090 Wien, Austria
Franz-Markus.Peschl@univie.ac.at

(**) Department of Computer Science and Engineering
University of Notre Dame
Notre Dame, IN 46556, USA
mscheutz@cse.nd.edu

1. Introduction

Computation plays a double role in cognitive science: conceptually, on the one hand, it underlies the main foundational tenet called *computationalism*—the claim that mental states are computational states—while, on the other hand, being applied in cognitive science's central research method, that of (computer) *simulation*.

In this paper, we investigate the role of simulation in cognitive science in oder to determine whether the methodological foundations of (computer) simulation are indeed compatible with computationalism, as is standardly assumed by cognitive scientists. We show that simulation might force computationalists to modify their credo.

2. Simulation and The Process of Theory Development

The goal of any scientific theory or model is to construct mechanisms that account for observable phenomena which can be detected by our sensory systems (possibly mediated by gauges of scientific apparati). In the case of cognitive science we are confronted with behaviors, neural activities, etc.; some of them can be observed directly, but often times one only observes the *effect* of a cognitive process (e.g., being realized as neural process), which is not directly accessible. Hence, constructing possible mechanisms which could account for these "hidden processes" is the main scientific aim. Their explanatory value consists in the *causal relations* established between (visible) phenomena which are only seen as a more or less coherent succession of states or behaviors over time.

How are these causal links constructed and which methods should one use, in order to accomplish this task in an efficient manner? Normally one thinks of the traditional *empirical* approach as the standard means for developing a scientific theory about a certain aspect of reality. In the natural sciences the classical epistemological feedback loop between the phenomenon (explanandum) and its theory is applied (see Figure 1, lower part): this cyclic process is based on the "epistemological tension/discrepancy" between a phenomenon in reality and its theoretical description. The goal of any scientific endeavor consists in closing this epistemological gap by applying highly sophisticated (empirical) methods for exploring the regularities of the phenomenon under investigation. The theoretical knowledge finds its expression in an experiment in which the environmental dynamics is penetrated in such a way that a certain—hopefully desired and predicted—effect/state is provoked. These results of the experiment are transformed into values being associated with variables of the original theory in the process of observation and interpretation. This transformation enables us to compare the actual values with the predicted values. A difference between them indicates that it might be necessary to apply some changes to the original theory. Combining results from other experiments and applying statistical methods leads to the inductive construction or adaptation of an alternative theory which acts as a starting point for the next cycle in this feedback loop. This cycle is repeated until a sufficient fit between the phenomenon and its theoretical description is achieved.

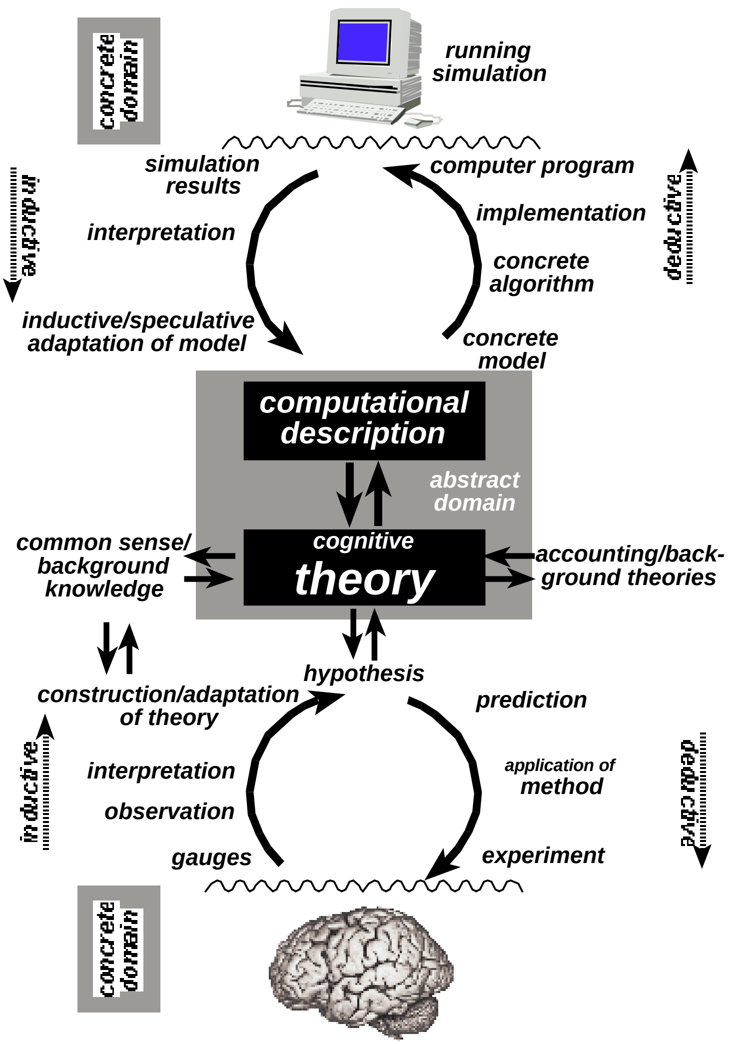

Figure 1 The process of theory construction: the classical feedback loop between a cognitive phenomenon in reality and its description in a theory, the “classical *empirical* loop” (lower part); the method of *simulation* is an

extension establishing a second feedback loop for “virtual (simulation) experiments”, the “virtual loop”.

Contrary to many other disciplines studying cognitive phenomena which are applying only this classical empirical approach, one of the methodological characteristics of cognitive science consists in the fact that its theories/models are often the result of *simulation experiments*. As simulation is an integral part of cognitive science, . its theories have a different character and follow a different strategy of construction. The method of simulation introduces a second feedback loop having a direct influence on the development of the particular theory (Figure 1, upper part). This extension to the classical method is a kind of mirror loop; i.e., the empirical loop gets extended/mirrored in the domain of virtuality and computation. The (empirically constructed) theory is transformed into a computational model[1] and the empirical experiment is replaced by a *virtual experiment* being realized by running a simulation of this model on a computer. The result of this cyclic process is a possible change in the computational model which might suggest the necessity of changing the original theory. This rewritten version of the theory acts as the starting point for a new cycle of empirical and/or simulation experiments. This implies that the simulation of a cognitive model does not only contribute to the development of the computational model, but also influences the construction of the (empirical) cognitive theory and, hence, essentially contributes to the understanding/explanation of the phenomenon under investigation.

3. Simulation and Computationalism

Computationalism can be defined as “the hypothesis that cognition is the computation of functions.” (Dietrich, 1990, p. 135) For many computationalists the virtue of computations (and hence their potential as descriptive vehicles for cognitive functions) is that they can be viewed as being descriptions of the distilled causal structure of the systems that implement them. If computations *qua* computations are to explain the capacity of a system to exhibit certain behaviors, they can do so only by virtue of being able to stand in a certain relation to (some of) the physical states of the system, the behavior of which they are supposed to explain. Since computational states are abstractions over space and time in that they only retain identity and difference of the physical states of which they are abstractions, they can only describe behaviors which *do not depend on timing* nor any other real-world fact (besides causal order). Put differently, computations are not able to tell physical systems apart that differ with respect to any physical quality, yet share the same computational structure.

With respect to cognition then it is conceivable that while one system might be practical and serve as a basis for cognition (because it is fast enough, gets the timing right, etc.), another with the same computational structure is of no practical use (maybe because it is too fast or too slow, think of Block’s example of the People’s Republic of China implementing the functional architecture of a mind). In short: computations can at best distinguish causally different physical systems, and thus the question remains *if this*

[1] We will focus on this crucial process in the following sections – for the moment it is sufficient to remain in this rather superficial description.

kind of distinction is sufficient for cognitive systems, i.e., if computations can distinguish cognitive from non-cognitive physical systems.
An answer to this question will depend on whether mental or cognitive properties can be individuated solely in causal terms, or whether additional physical aspects (such as time, etc.) enter the definition of cognitive functions. Obviously, there needs to be a tight relationship between environment and perceptual system, for to be able to recognize events in the environment where timing and/or the quality of the stimulus matter the perceptual system needs to be in tune with the physicality of these processes, otherwise it would fail to recognize them as such. As a consequence, (higher) cognitive functions also need to be embedded in the temporal structure imposed by perceptual and motor processes, which in turn depend on the predetermined temporal metric of causal network of our world (e.g., see Port and van Gelder, 1995). Yet, it seems that this temporal metric of our physical world does not matter for Turing computation. The merely computational, purely causal level is too abstract a level of description (in that it leaves out too many details about timing, etc.) to be able to capture cognition entirely.

While computation might be too abstract to capture all relevant aspects of a physical system, it is still possible to simulate the physical system by turning the mathematical description (e.g., a dynamic system) into a computational model (e.g., an algorithm that computes solutions to that dynamical system). In more mathematical terms, as a first rough cut one could view “simulation” as a relation between two physical systems (or possibly types of systems), such that the one system, which is to be simulated, is described by an abstraction (i.e., theory, architecture, function) *f*, which in turn is modeled by a computable (theory, architecture, function) abstraction *g* implemented by the other system. That leaves open the question as to the exact relation between the two abstractions. It is clear that they need to be *similar enough* (in some sense of “similar”), otherwise the system implementing *g* would not count as a simulation of the system described by *f* (see Figure 2). Obviously, everything hinges on the notion of similarity that underwrites the notion of simulation.[2]

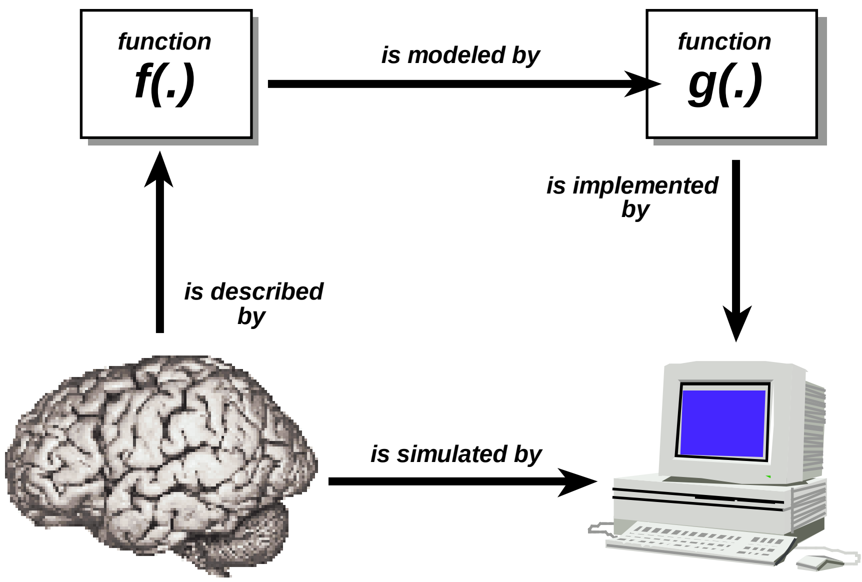

Figure 2 The relation between a system under study (the "brain" on the left), its mathematical description f and a simulation of the system (on the "computer" on the right), which implements the computable abstraction g.

Note that there are two special cases, one in which $f = g$ and another in which there is no physical system of which f is a description. In the former case, since one abstraction is implemented in two different physical ways, the two systems are effectively *the same* with respect to f (even if the two implementations might not be "perfect" or f is just an idealization of a physical/causal aspect of the systems). In the latter case, what is being simulated is a V(irtual)R(eality) system, one that is governed by the description f regardless of whether it is physically plausible or possible. The reason why we can simulate virtual worlds in our actual world is that the physical states in the simulating system are merely used as the implementing states of the computational states of g, that is, they are used for their causal role, not for their physical makeup.

This last point is central to all simulations: physical states of the *simulating system* do *not* have to resemble the physical properties of the physical states of the *simulated system* as their only purpose is to serve as the causal implementations of the computational states modeling the description of the behavior of the simulating system. It should be clear then that computer simulations of thunderstorms are not thunderstorms (e.g. Searle, 1980) because the physical qualities of the implementations of the computational states that model the dynamical system describing the behavior of physical weather states are different from weather states: at the chemical level the former are made of silicone, whereas the latter are made of air.

Why then the fuss about whether simulated cognition is the same as cognition? The difference between cognition and physical phenomena is that cognitive functions are usually not defined physically. Rather, they are defined at a very abstract functional level, which might have led computationalists to believe that this level is *all that matters* to cognition. And, in fact, there are even very recent arguments that are intended to establish that mental properties are so-called "organizational invariants" (see Chalmers, 1997), which are true of a physical system by virtue of the system's causal organization alone, regardless of any other physical qualities.

It is an easy consequence that if mental properties are organizational invariants, computational descriptions are sufficient to capture cognitive functions, as computations capture (part of) the causal organization of systems. Consequently, simulations of computations are computations, and hence simulating cognitive functions means performing them (in terms of properties this means that simulations of mental properties duplicate or instantiate these properties). Another, more suggestive way of putting the above is that computations are too abstract to allow for a distinction between the simulation and the simulated.

---

[2] Note that the Latin "simul" and "simile" are synonymous, and thus betray the common ancestor of the English nouns "simulation" and "similarity".

4. Conclusion

For the computationalist there is no distinction between computations that give rise to cognition and computational simulations thereof. At the very least computationalists will have to accept that causal explanations will have to take into account the temporal metric imposed by physics to do justice to real-world cognitive processes regardless of whether or not some cognitive functions might intrinsically depend on some of the physical qualities of their implementing systems. Yet, this concession does not obviate computer simulations as means of elucidating the nature of cognitive functions even if simulations do not implement cognitive function themselves.

5. References

Chalmers, D. J. (1997) “A computational Foundation for the Study of Cognition”. (published on the internet)

Dietrich, E. (1990) “Computationalism”, *Social Epistemology*, vol 4, no 2, 135-154.

Port, R. and T.v. Gelder (Eds.) (1995). *Mind as motion: explorations in the dynamics of cognition*. Cambridge, MA: MIT Press.

Searle, J (1980) “Minds, Brains and Programs”, *The Behavioral and Brain Sciences* 3, 417-424.