Euro Jnl Phil Sci (2012) 2:395–434
DOI 10.1007/s13194-012-0049-7

ORIGINAL PAPER IN PHILOSOPHY OF SCIENCE

How can computer simulations produce new knowledge?

Claus Beisbart

Received: 12 December 2010 / Accepted: 11 March 2012 / Published online: 24 April 2012

Abstract It is often claimed that scientists can obtain new knowledge about nature by running computer simulations. How is this possible? I answer this question by arguing that computer simulations are arguments. This view parallels Norton's argument view about thought experiments. I show that computer simulations can be reconstructed as arguments that fully capture the epistemic power of the simulations. Assuming the extended mind hypothesis, I furthermore argue that running the computer simulation is to execute the reconstructing argument. I discuss some objections and reject the view that computer simulations produce knowledge because they are experiments. I conclude by comparing thought experiments and computer simulations, assuming that both are arguments.

Keywords Computer simulations · Knowledge · Arguments · Thought experiments · Reasoning · Extended mind hypothesis

1 Introduction

These days, science news often report that computer simulations have shown this and that. To take one of many examples, Soter (2007) writes:

> Extensive computer simulations show that the eight planets greatly disturb the motions of test particles placed on circular orbits at most locations in the solar system.

C. Beisbart (✉)
Institute for Philosophy and Political Science, TU Dortmund 44221, Dortmund, Germany
e-mail: Claus.Beisbart@udo.edu

The idea is obviously that scientists have obtained new knowledge by running computer simulations. That computer simulations (CSs, for short) can produce knowledge is also claimed by philosophers (see e.g. Norton and Suppe 2001, p. 88; Küppers and Lenhard 2005b, Sec. 1.2; Winsberg 2010, p. 6; Gramelsberger 2011a, p. 215).

If the science news and the philosophers have it right—and I will assume so in this paper— we face the following question: *How can scientists gain new knowledge by running computer simulations?*[1] It seems puzzling that scientists can obtain knowledge about a real-world target system without actually observing it. The aim of this paper is to address this puzzle and to defend a certain answer to the question. This answer will also justify the assumption that one can obtain new knowledge by doing computer simulations.

To address the question of this paper, it is useful to look at a method that parallels computer simulations in important ways. *Scientific thought experiments* (TEs, for short) bear a pre-theoretical similarity to computer simulations. Like CSs, they are claimed to produce new knowledge about the real world although the system of interest is not observed.[2] As Humphreys (2004) notes,

> many simulations are examples of what would have been, in technologically more primitive times, thought experiments (p. 115).

The question how thought experiments can provide knowledge has recently attracted some attention.[3] It is therefore useful to draw on the related philosophical discussion. In particular, I will follow an example set by Norton. He defends what he calls the *argument view* about thought experiments (Norton 1991, 1996, 2004a). According to this view, thought experiments are simply arguments. I will argue that computer simulations are best considered as arguments too.

To do so, I will first summarize Norton's argument view concerning TEs as well as related discussions in the philosophy of science in my Section 2. The aim of Section 3 is to expound the argument view concerning computer simulations. To argue for this view, it will be necessary to reconstruct computer simulations as arguments, and I will do this in Section 4. Since it is often thought that rational reconstructions are boring and too far from actual scientific practice, I will go one step further and show in Section 5 that the reconstructing argument is in fact gone through when a computer simulation

---

[1] The importance of this question is also stressed by Winsberg (1999), p. 277, Stöckler (2000), p. 366 and Barberousse et al. (2009), pp. 558–559.

[2] That thought experiments can produce new knowledge is the majority view in the philosophical literature; see e.g. Gendler (2004), p. 1153 and Cooper (2005), p. 328. A dissenting voice is Atkinson and Peijnenburg (2004). Moue et al. (2006) and Brown and Fehige (2010) review the current philosophical research literature about scientific thought experiments, see also Kühne (2005).

[3] See Brown (1991, 2004), Norton (1996, 2004a) for the most important contributions; see also Gendler (1998, 2000, 2004) and McAllister (2004).

is run. In Section 6 I will assume that both TEs and CSs are arguments and compare them on this basis. I draw my conclusions in Section 7.

The recent philosophy of science has seen various attempts to locate CSs on “the usual methodological map” (Galison 1996, p. 120)[4] and a number of philosophical characterizations of CSs try to account for their ability to produce knowledge.[5] Morrison (2009) argues that simulations can provide knowledge in the way experiments do. Barberousse et al. (2009) disagree and argue that CSs produce information about a target not because the computer is a real physical system as is the target, but rather because the states of the computer can be understood “as computational states, then as values of variables, and finally as representing the target system’s properties” (p. 562). They further stress that a CS can only produce knowledge if it is based upon a good model of the target (ibid.). This is in accordance with the views of Humphreys who thinks that CSs can produce knowledge because they solve what he calls computational models (Humphreys 2004, p. 110; see his Section 3.14 for computational models). Winsberg seems to disagree when he suggests that simulations do not form abstract models, but rather furnish “direct representations” of their targets (Winsberg 2001, p. S450). According to Winsberg, the credibility of simulations derives partly from theories (e.g. Winsberg 2010, pp. 64–65; see also Winsberg 2001, particularly p. 448), but also relies on substantive background knowledge about the target system and about appropriate computational techniques (Winsberg 2010, p. 65; the importance of subject-specific knowledge in the construction of computational models is also stressed by Humphreys 2004, pp. 91–95). For Winsberg, computational techniques and tricks can be self-vindicating (e.g. 122), the idea being that they prove successful and become entrenched in traditions of research. A similar position is taken by Lenhard and Küppers. They deny that computer simulation is only a number-crunching technique to solve the equations of prior models and propose to say that a CS produces knowledge by imitating the dynamics of its target or of a model thereof (Küppers and Lenhard 2005a, p. 305; see also Lenhard 2007). Accordingly, they stress that simulations have to be empirically successful (pp. 319–320, see also Küppers and Lenhard 2005b), which gives them a quasi-empirical character (Küppers and Lenhard 2005a, p. 319). Gramelsberger (2010) goes even further than Winsberg, Lenhard and Küppers when she claims that simulations and algorithms constitute a new symbolic form (Ch. III.3, see also Gramelsberger 2011b). The algorithms underlying CSs are supposed to form a new language with symbols that can be executed (Gramelsberger 2010, p. 264).

Although this literature has led to important insights, it is not as illuminating as it could be, I believe. Some suggestions, for instance

---

[4] See Humphreys (1990), Hartmann (1996) for pioneering philosophical works about CSs, Humphreys (2004) and the essays collected in Frigg et al. (2009, 2011) and the monograph Winsberg (2010) for more recent philosophical contributions about CSs.

[5] Works that state this aim very clearly are Winsberg (1999), Stöckler (2000) and Barberousse et al. (2009).

Gramelsberger's proposal or Winsberg's idea that computational techniques become self-vindicating are fairly opaque from the outset and do not advance our understanding of how simulations produce knowledge. It is correct when many authors stress the links and parallels between CSs and models, but it is a much-debated question how we can learn from models and how models represent to begin with.[6] I think we can give a more straight-forward explanation how scientists can gain knowledge when they run computer simulations. To unfold this explanation is the task of this paper. What I propose is compatible with many insights from the literature, and I will explain this later in the paper.

Computer simulations may be conceived of in different ways. In a narrow sense, to run a computer simulation is to run a suitable simulation program. In a broader sense, we may can think of programming, testing the program and evaluating the results as part of a CS too.[7] I will follow Parker (2009), p. 488 in calling simulations in a broader sense *(computer) simulation studies*. Now it is clear that computer simulation *studies* typically include arguments. As Winsberg (1999) puts it, simulations

> involve a complex chain of inferences that serve to transform theoretical structures into specific concrete knowledge of physical systems (p. 275).

He seems to get even closer to the argument view when he says that

> simulation *is* a rich inferential process (Winsberg 2001, p. S442, my emphasis).

But while Winsberg's papers focus on the construction of a computer simulation and its validation, I will here restrict myself to computer simulations in a narrower sense. I will focus on one run of a computer simulation program. This is certainly a key part of every simulation, and I will argue that this process can be regarded as argument. This is also what Stöckler (2000) thinks:

> Given the model and the alleged initial state of the system, a simulation has the form of an argument [...] Simulations are arguments, not experiences (pp. 367, 369).

Stöckler takes this to be obvious and he does not work out in detail how simulations have the form of arguments. I will fill this gap in this paper. Apart from Stöckler's paper, I only know of occasional remarks that relate computer simulations and arguments.[8] It is thus time to think more thoroughly about computer simulations and arguments. By doing so, we will see that the argument view about computer simulations is subject to certain limitations.

---

[6] See Bailer-Jones (2003), Suárez (2003, 2004) and Giere (2004) for related discussions.

[7] Cf. Frigg and Reiss (2009), p. 596 for a similar distinction.

[8] For instance, Humphreys (2004), p. 71 draws a comparison between argument schemes and what he calls computational templates.

I nevertheless think that this view is a step forwards to understand the work that computer simulations do.

Before I start, a few clarifications are in order. *First*, I will often say that computer simulations produce knowledge. This is only meant to be a shorthand way of saying that scientists gain knowledge by doing computer simulations. I do not mean to imply that simulations produce knowledge in the way coffee machines produce coffee. *Second*, this paper focuses on knowledge and its acquisition through simulations. This is not to deny that, at least at first sight, scientific computer simulations are used for other purposes as well. But it is plausible to think that CSs perform other tasks *because* they provide knowledge or information. For instance, CSs are often used to design experiments and to validate the results of the latter.[9] They do so because they make experimentalists learn how their instruments would react to certain signals and how certain factors could disturb the experiment. Whether or not each possible task of CSs can be reduced to their obtaining information, does not matter for the purposes of the present paper though, and my focus will be on knowledge. *Third*, this paper does not aim to argue in favor of Norton's argument view *about TEs*. I have in fact some reservations about this view, as I explain in footnote 15 on page 8. Rather, I take Norton's view to be an interesting point of departure for my discussion. *Fourth* and finally, the discussion of TEs in this paper is confined to *scientific* TEs. Famous examples of scientific TEs include Galileo's TE with the falling bodies or Stevin's analysis of the forces acting on a chain. Philosophical thought experiments will not concern me in this paper.[10]

2 The argument view concerning thought experiments

The argument view concerning scientific TEs is supposed to answer the question of *how* one can gain new knowledge by running a thought experiment.[11] As I am only interested in new knowledge, I will often drop the qualification "new" in what follows. The rough idea behind Norton's argument view is this: TEs can give us new knowledge about nature because we draw on prior knowledge when we run a TE: Starting from prior knowledge, an argument leads us to new beliefs, and to the extent that the argument is sound, the new beliefs provide knowledge.

---

[9] See e.g. Perret-Gallix (2002), p. 488 for evidence; cf. also Galison (1997), particularly p. 689.

[10] Galileo's TE can be found in Galilei (1933), Vol. VIII, 107–109 (see Galilei 1974 for a translation into English); see Brown (1991), p. 43, Norton (1996), pp. 340–345, Gendler (1998), Atkinson and Peijnenburg (2004), McAllister (2004) and Kühne (2005), pp. 41–57 for the philosophical discussion about Galileo's TE. See Brown (1991), pp. 3–6 and Norton (1996), pp. 349–351 for Stevin's TE. For philosophical thought experiments see e.g. Cohnitz (2006).

[11] See Norton (1991, 1996, 2004a, b) for articulation and defense; see Irvine (1991), pp. 149–150 for a similar view.

The argument view about TEs can be summarized using the following slogan:

A_{TE} “Thought experiments are arguments” (Norton 1996, p. 354).

To spell out the argument view in more detail, Norton makes two claims (Norton 1996, p. 354):

R_{TE} Reconstruction thesis: Each thought experiment can be reconstructed as an argument such that the epistemic power of the TE is that of the argument.

P_{TE} Practice thesis: To run a thought experiment is to execute an argument, viz. the reconstructing argument.

The *reconstruction thesis* R_{TE} concerns the context of justification (ibid.). It explains the power of TEs to produce knowledge by saying that each of them can be reconstructed by an argument. The argument is supposed to fully capture the epistemic power of the thought experiment: The conclusion of the TE is justified if and only if (iff) the related argument is sound. An argument is sound iff its premises are true and support the conclusion. The reconstruction thesis entails that knowledge obtained via a TE may in principle be gained by running through an argument.

The *practice thesis* concerns the context of discovery (ibid.) and the practice of doing TEs. The thesis implies that, whenever somebody gains new knowledge by going through a thought experiment, she executes the reconstructing argument.[12]

The argument view does not imply that every argument is a TE. TEs are very special arguments. In Norton’s view, TEs characteristically concern “hypothetical or counterfactual states of affairs”; they also refer to particulars that do not figure in the conclusion of the thought experiment (Norton 1991, p. 129, emphasis deleted). That TEs consider imagined scenarios with particulars is also affirmed by opponents of the argument view, see e.g. Gendler (1998), pp. 398–399 and Cooper (2005), pp. 328–329.

In Norton’s view, the reconstruction and the practice theses substantiate the slogan that thought experiments are arguments. But it is arguable that the theses fall short of fully justifying the slogan. It might be that a TE *is* different from the reconstructing argument even though conducting the TE is in fact to go through the argument. Some evidence for this is given by Bishop (1999) who argues that thought arguments are not individuated in the same way in which arguments are. The individuation and “ontology” of thought experiments does

---

[12] Here the “context of discovery” has to be taken with some grain of salt. The context of discovery concerns what is done when a TE is run. It is not about the construction and first discovery of a TE. Norton’s argument does not say much about the construction of TEs. In this paper, I understand “context of discovery” in the sense in which Norton uses the term.

in fact raise very difficult questions (as does the ontology of other products of human creativity, e.g. of symphonies). In this paper, I bracket such questions. My reason is that they are not important for the epistemological question that interests me. I think that we have a satisfying explanation of how one can gain knowledge using TEs if we can establish the reconstruction and the practice theses. I will therefore take the slogan A_{TE} to abbreviate the conjunction of the theses R_{TE} and P_{TE}.

But how can the argument view be supported and what are its strengths? In his papers, Norton has assembled a number of arguments in favor of the view.

A first strength of the argument view is that it explains a simple way how TEs can produce knowledge. It is uncontroversial that we can obtain new knowledge by running through an argument, and if thought experiments are arguments, we have a simple explanation how TEs provide new knowledge.[13]

Second, Norton has provided an inductive argument for the reconstruction thesis: He has reconstructed a number of TEs in terms of arguments (see e.g. Norton 1996, pp. 338–353). He has in fact considered thought experiments that were presented to him as a challenge since they were supposed to resist reconstruction.[14] The reconstruction thesis seems thus well supported. In order to make a case for the practice thesis as well, Norton (1996) points out that Galileo explicitly calls his famous thought experiment a “demonstration” (p. 341). And in textbooks etc., TEs are mostly presented in written text that runs through an argument (Norton 2004b, p. 51). Norton also argues that it would be strange if arguments fully accounted for the production of knowledge due to TEs, but if TEs relied upon a completely different epistemic capacity (ibid. and Norton 1996, p. 356).

A third line of support for the argument view starts from the idea that thought experimenting is a reliable method of scientific inquiry (Norton 2004b, pp. 52–55). But there are thought experiments that support contradicting conclusions (ibid., pp. 45–49). Which one is right? Norton argues that, to answer this question, we have to find out which TE boils down to a sound argument.

In spite of all this, the argument view has also been subject to criticism. For the following, it is useful to mention two objections that have been levelled against the argument view about TEs.

---

[13] But can we really gain *new* knowledge by running through an argument? Some might want to deny this because the conclusion of the argument was already entailed or supported by prior knowledge. If knowledge is closed under deduction and under sound inductive argument, we cannot obtain new knowledge by going through an argument. But the assumption that knowledge is closed in this way is not very plausible to begin with, and when we run through a sound argument, the conclusion can at least be new in the sense that we did not believe it before. If this psychological sense of novelty is not sufficient at this point, it should be pointed out what the stronger sense of novelty is and why thought experiments provide new knowledge in this stronger sense (cf. Norton 1996, p. 346).

[14] Such TEs include Newton’s bucket experiment, Stevin’s thought experiment (ibid., pp. 347–351) and an argument concerning the continuum hypothesis (see Norton 2004a, pp. 1147–1148).

First, it has been suggested that TEs may be reconstructed in terms of arguments, but that the reconstructions do not fully capture the epistemic power of the thought experiments. Gendler (1998) provides a good example for this kind of criticism. She argues that a particular reconstruction of Galileo's thought experiment with falling bodies (pp. 403–404) does not fully account for the force of this very TE (pp. 404–408). If every possible reconstruction of the TE fails to explain what is essential about the TE, then there is clearly a problem for the argument view. On the constructive side, Gendler (2004) points out that pictorial imagination is essential for thought experiments.

Second, Cooper (2005), p. 332 has objected against the practice thesis that it does not capture the phenomenology of going through a TE. She writes:

> Simply put, constructing a thought experiment feels quite different from producing a logical argument. Thought experiments are often fun and easy, arguments are usually not. (ibid.).

These kinds of criticism instantiate the following scheme (Norton 2004b, pp. 55–56): Arguments lack some X that thought experiments have and that is essential for them. Thus, TEs cannot be arguments. If X is closely related to epistemology, an objection against the reconstruction thesis arises. If X is not epistemologically significant, we have an objection against the practice thesis.

This is not the place to comment on these objections and to balance the strengths and weaknesses of the argument view.[15] Also, I will not here discuss rivals to Norton's argument view. This short review suffices to set the stage for a discussion of computer simulations to which I now turn.

3 Computer simulations as arguments

Can we establish an argument view concerning CSs? What would it claim and how can we support it?

The argument view concerning CSs is supposed to answer the question of how computer simulations can obtain new knowledge. The rough idea is this: CSs can give us new knowledge about nature because we draw on prior knowledge when we run a CS: Starting from prior knowledge, an argument leads us to new beliefs, and to the extent that the argument is sound, the new belief is justified belief and thus at least a candidate for knowledge.

---

[15] Very briefly, my own view is as follows: I agree with Norton that TEs can be reconstructed as arguments in some way. But it is naïve to think that we reason from the premises to the conclusion of this argument when we conduct a thought experiment on our own (i.e., if we think what may happen in a certain counterfactual scenario as pictured by a TE). At least in some kinds of TEs, we “know” more intuitively what will happen. This point has implications for both the context of justification and the context of discovery.

The argument view about CSs can be summarized using the following slogan:

A_{CS} Computer simulations are arguments.

To spell out the argument view in more detail, we follow Norton's example and make two claims:

R_{CS} Reconstruction thesis: Each computer simulation can be reconstructed as an argument such that the epistemic power of the CS is that of the argument.

P_{CS} Practice thesis: If a computer simulation is run, the reconstructing argument is executed.

As in Norton's argument view, the *reconstruction thesis* R_{CS} concerns the context of justification. It claims that the epistemic power of a CS can fully be captured in an argument: The use of the CS is justified iff the premises are true and support the conclusion of the argument. The reconstruction thesis entails that knowledge obtained via a CS may in principle be gained by running through an argument.

There are many CSs though that are built upon assumptions that are thought to be false. Some CSs rely on such assumptions because they are not intended to trace a real-world target, but rather meant to probe a counterfactual scenario or a model. Other CSs are meant to refer to a real-world target, but use highly idealizing assumptions that are thought to be false literally speaking. The reconstruction thesis can easily be generalized to such simulations. A CS of this sort can still be reconstructed using an argument, but the use of the CS is already justified if the premises of the argument *would* support its conclusion if they were true. In the rest of this paper, I will understand the reconstruction thesis in this sense, but since my interest is in the acquisition of knowledge of the real world, I will most often consider simulations with true premises. I will come back to idealizing assumptions later in this paper.

The *practice thesis* concerns the context of discovery. The thesis implies that the reconstructing argument is executed if a CS is used to obtain new knowledge.

The argument view concerning CSs does not hold that every argument is a CS. To be a CS, an argument must be executed using a computer; otherwise it will not be a *computer* simulation. It must also go through a calculation that traces the values of empirical characteristics of a target system in time; otherwise it will not count as a *simulation*.

In the following, I will read the slogan A_{CS} as being fully justified by the theses R_{CS} and P_{CS}. Issues that concern the ontology, particularly the individuation of CSs will be bracketed.

How can we support the argument view concerning CSs and what are its strengths? In the next section, I will start with an example of a computer simulation in order to establish the reconstruction thesis R_{CS}. My strategy parallels Norton's inductive case for his reconstruction thesis R_{TE}.

4 Reconstructing computer simulations

Can computer simulations be reconstructed as arguments? The answer may seem yes quite obviously because CSs are based upon algorithms and programs, which are usually communicated using statements. But if algorithms and programs consist of statements, the statements are instructions and commands, and there is no part of them that supports the other in the way premises support an argument's conclusion. Furthermore, the algorithm upon which a computer simulation is based is most often not exactly followed in the simulations. Therefore, the reconstruction thesis about CSs is not trivial. To argue for the latter, I will proceed in four steps. I will first reconstruct a simple example of a computer simulation and second amend this reconstruction. Third, I will argue that a large class of computer simulations can be reconstructed in the same way. Fourth, I will explain why I think that the reconstructions fully capture the epistemic power of the underlying computer simulations.

4.1 A simple example of a reconstruction

Imagine a little piece of gum hanging from the ceiling. The piece is free to move back and forth and subject to several forces: gravity attracts the piece downwards, there is an external periodic force driving it (for instance, somebody pushes the piece), and friction tends to slow down its motion. For simplicity, we assume that the motion is to a good approximation confined to a plane.

Under some standard idealizations, which I will not discuss here, we obtain a mathematical model of this system. The core of this model is a dynamic equation that allows one to trace the motion of the piece (see Pang 2006, pp. 90–91 for the following). Call the model "driven and damped pendulum". In the model, our piece of gum is modeled as a pendulum of mass m. Let l be the length of the pendulum (which models the length of the thread between the piece of gum and the ceiling), and $\varphi(t)$ denote the angle between the thread and a plumb line at time t. Let furthermore g denote the gravitational field at the surface of the earth, and assume that the externally driving force results in a torque of $lf_0 \cos(\omega t)$ with e and ω being constant in time. f_0 encodes the strength of the external force, while ω reflects its period. Assume finally that $-\gamma l\dot{\varphi}$ is the damping force, where γ is a constant and where a dot denotes the derivative with respect to time t. $\dot{\varphi}$ is thus the angular velocity. The motion of the pendulum can then be described using the following equation:

$$\ddot{\varphi}(t) = -\frac{g}{l}\sin(\varphi(t)) - \frac{\gamma}{m}\dot{\varphi}(t) + \frac{f_0}{ml}\cos(\omega t)\ . \qquad (1)$$

For simplicity, we set

$$f(x, y, t) := -\frac{g}{l}\sin(x) - \frac{\gamma}{m}y + \frac{f_0}{ml}\cos(\omega t)\ , \qquad (2)$$

such that our dynamic equation, Eq. 1 reads

$$\ddot{\varphi}(t) = f\left(\varphi(t), \dot{\varphi}(t), t\right). \tag{3}$$

This equation and the initial values of angular position and angular velocity,

$$\varphi\left(t_0\right) = \varphi_0, \quad \dot{\varphi}\left(t_0\right) = V\varphi_0, \tag{4}$$

define an initial value problem. I shall assume that the initial value problem has a unique solution.

For our purposes, it is useful to assume that units, a reference frame and coordinates have been fixed. In our example, we assume polar coordinates. l, m etc. can then be regarded as real numbers, and the function $\varphi(t)$ maps real numbers to real numbers. This interpretation of the equations is not without alternative, but it is most natural because we are interested in CSs, and computer simulation programs return numbers. To keep the terminology simple, I will say that coordinates C have been fixed; this means that a full and unique empirical interpretation of the functions and equations is adopted.

For certain choices of the parameters and the initial values, the dynamic equation, Eq. 1 cannot be solved analytically (for some choices, the model predicts chaotic behavior). This is why one has to run computer simulations to analyze the model.[16] Simulations of the driven and damped pendulum and a few remarks about the model can be found under

http://www.myphysicslab.com/pendulum2.html.[17]

An extensive theoretical analysis of a very similar system is provided by Bartuccelli et al. (2001).

A very simple method to simulate the behavior of a driven, damped pendulum is based upon the so-called *Euler method* (see Pang 2006, pp. 81–83). To introduce the Euler method, we note that the second-order differential equation Eq. 3 is equivalent to the following system of coupled first-order differential equations:

$$\begin{aligned} \dot{\varphi}(t) &= V\varphi(t), \\ \dot{V\varphi}(t) &= f\left(\varphi(t), V\varphi(t), t\right). \end{aligned} \tag{5}$$

We have here introduced a new function $V\varphi(t)$; according to the first equation, it is angular velocity.

The Euler method approximates these first-order differential equations using difference equations. Time is discretized, i.e., we only consider the times $t_j = t_0 + j \times \Delta t$ with a fixed time interval Δt for $j = 0, \ldots, N$, where

---

[16] Cf. Weisberg (2007)

[17] Checked 12/2011

N is a natural number. The following recursive scheme solves the difference equations: For $j = 1, \ldots, N$,

$$\begin{aligned} \varphi\left(t_j\right) &= \varphi\left(t_{j-1}\right) + \Delta t \times V\varphi\left(t_{j-1}\right), \\ V\varphi\left(t_j\right) &= V\varphi\left(t_{j-1}\right) + \Delta t \times f\left(\varphi\left(t_{j-1}\right), V\varphi\left(t_{j-1}\right), t_{j-1}\right). \end{aligned} \tag{6}$$

We start this scheme with setting $\varphi(t_0) = \varphi_0$ and $V\varphi(t_0) = V\varphi_0$ and obtain a trajectory of the pendulum in $(\varphi, V\varphi)$-space, sampled at times t_j for $j = 1, \ldots, N$.

The Euler method is too inaccurate for many applications (Pang 2006, p. 82). But it can easily be improved by using a predictor-corrector method (Pang 2006, Section 4.3, particularly p. 84). The Euler method approximates time derivatives of φ in an interval $[t_{j-1}, t_j]$ by evaluating $V\varphi$ at t_{j-1}, i.e., at the left boundary of the time interval (and it proceeds similarly for $V\varphi$). A predictor-corrector method uses instead $\left(V\varphi\left(t_{j-1}\right) + V\varphi\left(t_j\right)\right)/2$ to approximate the time derivative of φ, where $V\varphi\left(t_j\right)$ is estimated using the Euler method. However, for the purposes of this paper, it suffices to reconstruct the Euler method as an argument. Our reconstruction of the Euler method can be generalized to other, more complicated methods, e.g. to a predictor-corrector method.

A simulation based upon the Euler method will carry out the following algorithm. I assume that φ_0 and $V\varphi_0$ are numbers given to the program as initial values.

```
set phi[0] = $\varphi_0$
set vphi[0] = $V\varphi_0$
for i = 1,...,N do:
      set t[i] = t0 + dt × i
      set phi[i] = phi[i-1] + dt × vphi[i-1]
      set vphi[i] = vphi[i-1] + dt
                     × f(phi[i-1],vphi[i-1],t[i-1])
      print t[i], phi[i], vphi[i]
```

I have used a slightly different notation than before to indicate how the algorithm is implemented in a programming language such as C++. For instance `dt` replaces Δt. `dt`, `t0` and `N` are supposed to be constants in the program. The function `f` is the same as f above. To implement this algorithm on a computer, I have written a program in java, called motion1.java. The code of the program is a slight modification of the java program provided in textbook about computational physics (Pang 2006, p. 11).

We can reconstruct this run of the simulation program in terms of the following argument. I use "φ_i" and "$V\varphi_i$" as place-holders for real numbers, and "$\varphi(t_i)$" and "$V\varphi(t_i)$" to denote the values of empirical characteristics in some coordinates.

P_1 The initial angular position of the pendulum p at time t_0, $\varphi(t_0)$, takes the value $\varphi(t_0) = \varphi_0 \in \mathbb{R}$ in coordinates C.

P_2 The initial angular velocity of the pendulum p at time t_0, $V\varphi(t_0)$, takes the value $V\varphi(t_0) = V\varphi_0 \in \mathbb{R}$ in coordinates C.

P$_3$ For each $i = 1, \ldots, N$, the angular position at time $t_i = t_0 + i \times \Delta t$, $\varphi(t_i)$, takes the value $\varphi(t_i) = \varphi(t_{i-1}) + \Delta t \times V\varphi(t_{i-1}) = \varphi_{i-1} + \Delta t \times V\varphi_{i-1}$ in coordinates C.

P$_4$ For each $i = 1, \ldots, N$, the angular velocity at time $t_i = t_0 + i \times \Delta t$, $V\varphi(t_i)$, takes the value $V\varphi(t_i) = V\varphi(t_{i-1}) + \Delta t \times f(\varphi(t_{i-1}), V\varphi(t_{i-1}), t_{i-1}) = V\varphi_{i-1} + \Delta t \times f(\varphi_{i-1}, V\varphi_{i-1}, t_{i-1})$ in coordinates C.

Co The angular positions of pendulum p at times $t_i = t_0 + i \times \Delta t$ for $i = 1, \ldots, N$, in coordinates C, $\varphi(t_i)$, take the values $\varphi(t_1) = \varphi_1 \in \mathbb{R}, \ldots$; and the angular velocities of the pendulum at times $t_i = t_0 + i \times \Delta t$ for $i = 1, \ldots, N$, in coordinates C, $V\varphi(t_i)$ take the values $V\varphi(t_1) = V\varphi_1 \in \mathbb{R}, \ldots$.

The premises and the conclusion of this argument refer to a pendulum with proper name "p". I assume that the simulation scientist knows what that pendulum is. She may think of our piece of gum, but also of a merely imagined system. We would not speak of a simulation if the working scientist did not think that her simulation is about some system. If the simulations have a real-world target, it is natural to say that the variables in the simulation program and thus the argument directly refer to this target and that p is this very target. This is natural because angular position and angular velocity are characteristics that a real-world pendulum system has.

Matters are different in other simulations. For instance, so-called N-body simulations of cosmic structure formation follow the trajectories of artificial particles to trace the distribution of the matter in the universe.[18] It is agreed that the artificial particles do not exist and that they are only used to model the matter distribution in the universe. The idea is that the distribution of the artificial particles is similar to that of the matter in the universe. I propose to say that, in such a case, the reconstructing argument refers to a model system distinct from the target.

The premises P$_1$ and P$_2$ specify the initial conditions of the pendulum, i.e., concrete values of physical characteristics that describe p at the initial time.[19] What the computer program takes as input are of course mere numbers. But for the working scientist, these numbers have empirical meaning; they encode the values of physical characteristics of the system in certain units. If we want to capture this empirical meaning in our reconstruction, it is useful to mention the numbers in the reconstruction, but to take them to be the values of physical characteristics in some units, given some reference system and suitable coordinates. To simplify the reconstruction, I use again the expression "in coordinates C" as an abbreviation for the conventions used to interpret mathematical objects empirically. Given coordinates C, the

---

[18] See e.g. Peebles (1980), Part II, Efstathiou et al. (1985), Bertschinger (1998), Klypin (2000) and Dolag et al. (2008) for such simulations. Hockney and Eastwood (1988) provide a general introduction to particle simulations.

[19] Balzer (2009), p. 324 assumes likewise that the inputs to, and the outputs of, simulations can be represented as sentences. –Since my example is from physics, I am speaking of *physical* characteristics such as mass etc. But nothing hinges on the characteristics being from physics, and in other examples we would be concerned with chemical characteristics etc.

statement that a certain physical characteristic (e.g. the mass of the pendulum) takes a certain number as its value, can be translated into a statement with empirical content.[20]

What the premises claim could equally be said using different coordinates. We do need coordinates to endow the numbers with empirical meaning, but what the resulting statements say does not depend on the specific coordinates used.

The other premises P_3–P_4 specify how the values of the relevant physical characteristics at later times follow from those at earlier times. These premises are general, they do not fix the value of a physical characteristic in some units and coordinates, but rather state how the values of these characteristics are connected to each other. To make this more explicit, premise P5 may be formulated like this:

P_3 For each $i = 1, \ldots, N$: If, at time t_{i-1}, in coordinates C, the angle takes the value $\varphi\,(t_{i-1}) = \varphi_{i-1}$ and if, at that time and in these coordinates, the angular velocity takes the value $V\varphi\,(t_{i-1}) = V\varphi_{i-1}$, then the angle φ, at time t_i and in the same coordinates, takes the value $\varphi\,(t_i) = \varphi\,(t_{i-1}) + \Delta t \times V\varphi\,(t_{i-1}) = \varphi_{i-1} + \Delta t V\varphi_{i-1}$.

In the conclusion of the argument, Co, the φ_i and $V\varphi_i$ are again place-holders for numbers for $i = 1, \ldots, N$. In the simulation, these numbers (i.e., the values stored under `phi[i]`, `vphi[i]` for $i = 1, \ldots, N$) arise, as the instructions of the algorithm are carried out. In the reconstructing argument, premises take the instructions to trace certain facts. These premises together with the other premises imply which values φ_i, $V\varphi_i$ the physical characteristics take at the times t_i ($i = 1, \ldots, N$). The argument is thus deductively valid.

We could add additional conclusions to the reconstruction. These conclusions would specify the values of characteristics that are not finally printed out as output, but that arise during the course of the calculations as values of variables in the program. For the purposes of this section we can bracket such additional conclusions because they will not play any role in what follows.

So far, I have assumed that the working scientist refers to a particular system and that she uses specific coordinates to translate the numbers that constitute the in- and outputs of the simulation into statements about this system. But this is not always so. For instance, if a scientist considers the dynamics of an imagined particle in some purely hypothetical gravity field, she is not interested in particular places, but only in the type of motion that arises. Likewise, in many simulations, it does not at all matter what the initial time t_0 is.

This is manifest from the way the system p and the coordinates C enter my reconstruction. Each statement refers to both of them, and if we replaced

---

[20] By stressing the empirical significance of the premises I do not mean to exclude statements about unobservables and their characteristics (provided we can refer to unobservables). Rather, what I call “empirical meaning” has to extend beyond purely mathematical significance and to refer to concrete systems in some however indirect way.

p with a different system *p*', and/or the coordinates C with different ones C' consistently, we would obtain another reconstruction that is equally fine, at least if the system *p*' and the coordinates C' are each taken from appropriate reference classes. Different polar coordinates may work as well, but we may not switch from polar coordinates to Cartesian coordinates because the dynamic equations are cast in polar coordinates.

This suggests that scientists often refer to a general *class* of systems and *classes* of coordinates. We can capture this by taking our reconstruction to define an *argument scheme* rather than one single argument.[21] The idea is that the expressions "system *p*" and "coordinates C" are place-holders for expressions that refer to single systems and to single coordinates from appropriate classes of systems and of coordinates, respectively. By inserting expressions that refer in this sense, we obtain a specific argument. What the appropriate classes of systems and coordinates are forms part of the knowledge of the working scientist.[22]

I have so far considered one single execution of an algorithm. But of course, an algorithm can be followed using different initial conditions. We can capture this by moving to more general argument schemes. The idea is that φ_0 and $V\varphi_0$ are place-holders for values of the physical characteristics in a suitable coordinate frame. We obtain different arguments (or argument schemes) if we insert different values.

This then is the picture we have come up with: The execution of an algorithm by a computer simulation program using concrete input values can be reconstructed using an argument or an argument scheme, depending on how the scientist thinks of the simulation. The whole computer program can be conceptualized as a more general argument scheme. In what follows, I will for simplicity assume that the single run is properly reconstructed in terms of one argument rather than an argument scheme.

4.2 Amending the reconstruction

So far, I have only reconstructed the *algorithm* and the *program* upon which a computer simulation is based. But if a scientist runs computer simulations on a specific machine, the algorithm is not followed exactly. Every computer on which the program is run will produce *roundoff errors*. For instance, if the computer is to add the numbers 10^{10} and 10^{-10}, it may end up with 10^{10} rather than with $10^{10} + 10^{-10}$.[23]

This has a severe consequence for my reconstruction. If the reconstructing argument is deductively valid as suggested, then the numbers $\varphi_i, V\varphi_i \in \mathbb{R}$ that figure in the conclusion have to be the numbers that would arise if the

---

[21] I take the notion of an argument scheme from Kitcher (1981), Section 5.

[22] Note that we cannot generalize our initial reconstructing argument by quantifying over *p* and C in each premise. It is important that we first fix one system and coordinates and then go through the whole argument.

[23] See Press et al. (2007), Section 1.1 for an introduction.

algorithm were followed in an exact way. Due to the roundoff errors, these numbers are different from those that the computer actually outputs, call them φ_i^{CP}, $V\varphi_i^{CP}$ for $i = 1, \ldots, N$. In real-world simulations, the scientists typically only have the numbers φ_i^{CP}, $V\varphi_i^{CP}$, and it would be very tedious work to derive the φ_i, $V\varphi_i$ by actually going through the argument on foot if this can be done at all. It is exactly the point of computer simulations to avoid this work.

Now our task is to reconstruct the computer simulations as they are actually done, and these yield the φ_i^{CP}, $V\varphi_i^{CP}$ as outputs. These numbers should then appear in the reconstruction, and it seems appropriate to replace the conclusion of my reconstruction by:

Co' The angular positions of pendulum p at times $t_i = t_0 + i \times \Delta t$ for $i = 1, \ldots, N$, in coordinates C, $\varphi(t_i)$, take the values $\varphi(t_1) = \varphi_1^{CP} \in \mathbb{R}, \ldots$; and the angular velocities of the pendulum at times $t_i = t_0 + i \times \Delta t$ for $i = 1, \ldots, N$, in coordinates C, $V\varphi(t_i)$ take the values $V\varphi(t_1) = V\varphi_1^{CP} \in \mathbb{R}, \ldots$.

Here, φ_i has been replaced by φ_i^{CP} and so on. Strictly speaking, the label "CP" refers to one run of the simulation program on a specific machine. Another program that executes the same algorithm on a different machine may yield different numbers and thus a different conclusion.

Replacing the conclusion Co by Co' has important consequences for the reconstructing argument. Whereas the argument was valid when Co was its conclusion, it will in general not be valid any more if we replace Co by Co'. More seriously, if Co and Co' assign different values to one and the same physical characteristic, they are inconsistent. Thus, if the reconstructing argument consists of P_1–P_4 and Co', the premises of the argument are inconsistent with its conclusion and do not support it. As a consequence, we cannot reconstruct the simulations by an argument that runs from P_1–P_4 to Co' because we cannot capture the epistemic power of a computer simulation using an "argument" in which the premises do not support the conclusion. The reconstructing argument has to be repaired with further amendments.

One option is to change the premises P_3–P_4. So far, they reflect what the algorithm prescribes, and we may replace them by premises that describe more accurately what the computer program does, as it is implemented on a concrete machine. To this end, we could replace the "$+$"-sign in the premises P_3–P_4 by a "$+_{CP}$"-sign that characterizes how additions are performed on the computer. The $+_{CP}$-addition is not associative any more, so we would have to make sure that the premises fix the order in which the additions are done on the computer. While it is possible to change the reconstruction along these lines, we would have to know for this how the algorithm is implemented and how the program is compiled in detail. Not even the working scientist will usually know this. We would thus have to reach out beyond the scientist's knowledge in reconstructing the argument. This may not be appropriate because, presumably, our reconstructions should reflect how working scientists think about their CSs. As a matter of fact, scientists commonly take it that roundoff errors do not make a decisive difference to the results of their simulations.[24] Further, the reconstruction is likely to become more intransparent if we describe how exactly the algorithm is implemented.

An alternative option is to replace Co’ by a weaker conclusion Co”. According to Co”, the values of the physical characteristics in the coordinates C do not take exactly those numbers that the computer outputs, but only approximately so:

Co” The angular positions of pendulum p at times $t_i = t_0 + i \times \Delta t$ for $i = 1, \ldots, N$, in coordinates C, $\varphi(t_i)$, take *approximately* the values $\varphi(t_1) = \varphi_1^{CP} \in \mathbb{R}, \ldots$; and the angular velocities of the pendulum at times $t_i = t_0 + i \times \Delta t$ for $i = 1, \ldots, N$, in coordinates C, $V\varphi(t_i)$, take *approximately* the values $V\varphi(t_1) = V\varphi_1^{CP} \in \mathbb{R}, \ldots$.

If the numbers that the computer actually outputs, φ_i^{CP}, $V\varphi_i^{CP}$, are not too far from the numbers that the algorithm would yield if it were executed accurately, $(\varphi_i, V\varphi_i)$, an argument that runs from P_1–P_4 to Co” seems reasonable again: The premises entail the conclusion.

By moving from Co’ to Co” we do capture the understanding that the working scientist has of the simulations and the use that she makes of them. In most cases, she is not interested in the exact numbers that the computer outputs. In fact, very often, the output printed by the computer is rounded off in some way. Thus, when the computer outputs the numbers $(\varphi_i^{CP}, V\varphi_i^{CP})$, the scientist thinks that, given P_1–P_4, the values of the physical characteristics in coordinates C take *something like the values* $(\varphi_i^{CP}, V\varphi_i^{CP})$.

Co” is vague in that it does not specify the degree to which the computer outputs approximate the $\varphi_i, V\varphi_i$, which we would obtain if we followed the algorithm exactly. We can Co” render more precise by specifying ranges around the numbers $\varphi_i^{CP}, V\varphi_i^{CP}$. The values of the physical characteristics in coordinates C are then supposed to be in the ranges $\left[\varphi_i^{CP} - \epsilon_{\varphi,i}, \varphi_i^{CP} + \epsilon_{\varphi,i}\right], \left[V\varphi_i^{CP} - \epsilon_{V\varphi,i}, \varphi_i^{CP} + \epsilon_{V\varphi,i}\right]$, where the $\epsilon_{\varphi,i}, \epsilon_{V\varphi,i} \in \mathbb{R}$ are *bounds on the errors* that the program produces comparing to the algorithm.

However, such a precisification does not really make progress unless we know what the bounds $\epsilon_{\varphi,i}$ and $\epsilon_{V\varphi,i}$ are. These numbers are not part of the output and can only be derived by analyzing the implementation of the computer program. A related derivation is often impossible in practice; the bounds can at most be estimated.[25] Nevertheless, the working scientist will have some understanding of what the order of magnitude of the bounds is. In particular, she may know how the error depends of the step size, i.e., `dt` in the program. We can read this understanding into Co”. I will not try to make this explicit by reformulating Co” and propose that we stick to an argument that runs from P_1–P_4 to Co” when we reconstruct the computer simulation.

---

[24] See e.g. Press et al. (2007), Section 1.1 again.

[25] See Press et al. (2007), p. 10 for an estimate.

To justify my reconstruction of the CS, I have referred to the knowledge of the working scientist. Some may suggest to go even further with taking the perspective of the working scientist. In our example, both the algorithm and the computer program follow not the differential equation, Eq. 1, or, equivalently, Eq. 5, but rather the Euler method and the difference equations, Eq. 6. These difference equations are stated in the premises P_3–P_4. But the working scientist will not think that the difference equations are true of the target.[26] She may nevertheless learn about the target through the simulations. This suggests that the premises P_3–P_4 should be replaced by premises that state the original differential equation, Eq. 1. If we follow this suggestion, we have to change the conclusion too because the latter is supposed to be supported by the new premises. Since the computer solves the difference equations instead of the differential equations, the outputs of the simulations display additional errors with respect to the differential equations, errors that are not yet taken into account in the bounds $\epsilon_{\varphi,i}$ and $\epsilon_{V\varphi,i}$. Co” has thus to be weakened further using larger error bounds. It is a non-trivial task to obtain appropriate error bounds because the latter have not only to reflect the so-called approximation errors[27], but also how approximation errors interact with the roundoff errors mentioned earlier in this section.[28]

This suggestion to change the reconstruction is not unreasonable, but nevertheless, I do not think that we have to follow it. If we did so, we would move too far away from what the CSs actually achieve. If the scientist does in fact think that the simulations teach her something about the original equations she is interested in, then we can also capture this by saying that the scientist runs an additional inference over and above the computer simulation.

There is a more general dialectic behind the discussion of possible amendments to my first reconstruction. If we want to reconstruct computer simulations as arguments, we face a dilemma. On the one hand, there are good reasons to characterize in an exact way what the computer actually does. This suggests that we describe the calculations of the computer at the level of the machine code. On the other hand, there are also good reasons to provide a reconstruction that is apt to characterize the uses that are made of computer simulations. In the simplest case, a scientist trusts the assumptions upon which a simulation is built and comes to believe the result. But if this is to happen, the premises of the argument must be something that the scientist believes, and she will not believe that the calculations carried out by the computer provide literally true descriptions of nature because the calculations are subject to errors of various types, e.g. roundoff errors and approximation errors.

---

[26] She will not even think that the original differential equations are literally true of the target because they rest upon idealizations. For the time being, I will bracket this fact, and return to it later on page 24.

[27] Approximation errors quantify the differences between the solutions to the difference equations and to the exact differential equations.

[28] See Press et al. (2007), p. 11 for this interaction.

This steers us away from what the computer actually does to hypotheses that working scientists are interested in.

There are two radical strategies to deal with this dilemma. The first is to get as close to the computer and to what it actually does as one can in reconstructing the argument. The idea is to say that further uses that are made of the simulations involve additional arguments on the part of the scientist. These additional arguments are not taken to be proper part of the simulation (in the sense of one run of the program), but they may be part of a related computer simulation study. The argument that reconstructs the one run of the simulations will be deductive, while the further inferences run by the simulationalist are likely to be inductive. The other radical strategy is to move as close to the scientist as one can and to reconstruct the simulations in terms of statements that come as close as possible to scientific hypotheses that the scientist could at least take serious. In this way we use *one* argument to reconstruct what was captured as two or more distinct arguments under the first strategy. One can of course also take a middle course between these radical strategies, as did I when I proposed to reconstruct the simulations in terms of P_1–P_4 and Co".

Which of these strategies is most appropriate? What exactly is the argument that underlies a specific simulation?

It is important to note that this question is mainly a conceptual one. The question is: What strategy is best suited to capture what we call a simulation? An answer to this question will not prejudge more substantive questions, for instance the question what some scientists concluded from their simulations as a matter of fact, and the question what they should have concluded.

The use of the term "simulation" gives us some guidance about how to reconstruct simulations. For instance, if somebody runs a simulation, we can always ask what was simulated, and the answer should refer to some target system. It is thus not appropriate to reconstruct computer simulations as purely mathematical arguments. In everyday talk, we can also say that a computer simulation has shown this and that. This suggests that a computer simulation should be reconstructed as a sound argument and this puts certain constraints on the reconstruction as well. I believe that the middle course that I have suggested fully takes into account such constraints.

But our concept of a CS is not sufficiently clear-cut as to single out one particular type of reconstruction. I conclude that a simulation may be reconstructed using alternative, slightly different arguments. What reconstruction we choose depends on the purposes of the inquiry. This is nevertheless sufficient to underwrite the reconstruction thesis as applied to our example. The reconstruction thesis has thus been supported using an example. But can we always proceed in this manner?

4.3 Can every computer simulation be reconstructed in this way?

For a large class of deterministic computer simulations the answer is clearly yes. A computer simulation is deterministic iff the output is a function of the input only (and not of so-called random numbers). In the following, I will focus on deterministic simulations. I bracket non-deterministic simulations not because I do not think they can be reconstructed as arguments, but rather because their treatment requires additional argument.

In many deterministic computer simulations, the input consists of numbers that are supposed to specify the values of characteristics of some system (or some class of systems) in some coordinates at some time—call it the initial time t_0. These characteristics may be physical, chemical etc., but for simplicity I will call them physical. We can even take the notion of a "characteristic" with some grain of salt here; labels that denote objects or types of events could count as characteristics for our purposes too. We can translate the input into statements about a system and consider the statements as premises. In a similar way, we can translate the output into statements that form the conclusions of the argument. This is so because the output typically specifies the values of physical characteristics of the system in the coordinates at other times. Finally, the algorithm is a set of interconnected instructions that specify operations which act on the input and produce the output.[29] We add a premise that states the following: The system behaves as if the algorithm is used to act on the values of the characteristics at the initial times to produce the values of the physical characteristics at the other times.

If we follow this recipe, we end up with a deductively valid argument of the following form:

I_1 At initial time t_0, the physical characteristic Y^1 of system S takes the value y^1_0 in coordinates C.

... ...

I_m ...

A The physical characteristics $\{Y^\alpha\}$ of the target system evolve in time as if their values in coordinates C were manipulated using the algorithm.

Co At time t_1, the physical characteristic Y^1 of the target system S takes the value y^1_1 in coordinates C ... [specifications of other characteristics, also at other times follow.]

This type of reconstruction is feasible for many simulations. It can even be generalized. For instance, t_0 need not be an *initial* time; we need not assume that premises and conclusion use the same coordinates etc.

If the algorithm consists of a sequence of instructions, the premise that reflects the algorithm can be unpacked into a number of distinct premises. Such premises state that the system behaves in some respect as if some step in the algorithm was followed. For instance, the algorithm may contain an instruction to add up three component forces to obtain the total force. This can be translated into a premise stating that the total force arises from adding the three component forces. We would probably not speak of a simulation if we could not unpack the algorithm in some such way. The slogan may be: no

---

[29] See e.g. Kronsjö (1979), pp. 1–2 for a definition of algorithms.

simulation with some kind of underlying mechanism. A program that predicts the behavior of a target system using a black box would probably not count as a computer simulation.

The argument that I have provided reflects the algorithm. In the previous section, I have argued that this may not be the most appropriate reconstruction of the simulations. But it is clear how to change the schematic reconstruction above to arrive at a more appropriate reconstruction.

What does my reconstruction of the computer simulation presume apart from there being a deterministic algorithm that maps inputs into outputs? The main presumption is that simulations refer to a real or imagined process of a real or imagined system, viz. the target system of the simulations. This is an uncontroversial trait of computer simulations.

Results of computer simulations Our reconstruction has an additional benefit as a byproduct. Start with the following observation: Scientists often claim that computer simulations have results. But what are these results? Our reconstruction suggests a standard form of stating the result that a scientist obtains by running the simulation program with specific initial values. If the reconstruction yields an argument that runs from premises P_1–P_4 to Co, the result can be stated as follows:

R If P_1 and ... and P_4 hold true, Co obtains.

If the simulation is more appropriately thought of as going through an argument scheme, we state the result by suitably quantifying over the system and the coordinates using appropriate domains.

The statement R is hypothetical and concerns the transition from the premises to the conclusion of the reconstructing argument. This transition is made by the simulations. There is no commitment here to the truth of either the premises or the conclusion; neither the premises nor the conclusions are entailed by R. This is as it should be because the working scientist who runs a simulation may not be committed to take the premises or the conclusion to be true. For instance, she may think that the premises are *possibly* true of some system and just want to know what the premises imply for a certain setting. What she is minimally committed to if she trusts her simulations though is that, *if* the premises hold true, so does the conclusion.

Often, at least a subset of the premises of a computer simulation are taken to be uncontroversially true. In stating the results of the simulations, a scientist can then suppress those premises. This is often mandated by the rules of pragmatics.[30] In the limiting case in which the scientist takes all premises to be uncontroversially true, she may fully detach the consequent, and simply say that, according to the simulations, Co obtains. In this way, simulations may be

---

[30] As every utterance, the statement of results from simulations is subject to rules of pragmatics; see e.g. Grice (1989) for such rules.

said to have shown that a particular asteroid will be trapped by a planet at some time t_1. Such a statement is clearly stronger than R.

4.4 Do the reconstructions capture the epistemic power of the computer simulations?

The reconstruction thesis does not only require that computer simulations can be reconstructed as arguments. The reconstructing arguments should also have the same epistemic power as the simulations. What about this part of the thesis? Is the epistemic power of a CS fully captured if it is reconstructed in the way suggested?

There are at least two ways to support an affirmative answer to this question. The first way is to carry out a brief check. If we compare computer simulations and their reconstructing arguments, is anything missing? The answer seems to be no. At this point, it is interesting to compare to the discussion about *TE*s. Some authors (e.g. Gendler 1998) have claimed that reconstructions of certain TEs via arguments miss a crucial point of their TEs. Would a similar claim make sense in the case of simulations? Could there be an X that is present in the CS, but missing in the argument (cf. p. 8)? I do not think so. It is plain that the computer runs certain calculations and that the numbers are taken to trace the values of empirical characteristics. This can be represented using an argument. Is there anything more to computer simulations?

The second line of argument is inductive. As already mentioned in the introduction, at some surface level, CSs perform various sorts of epistemic tasks. The idea then is to show that, for every CS, a reconstructing argument can perform the task that the corresponding CS completes. The working hypothesis thus is

D Every epistemic work that a computer simulation does could also be done by executing the reconstructing argument without computers (e.g. by a human being thinking through the argument).[31]

According to D, the use of computers is in principle dispensable. It is only because some arguments are too difficult to devise why scientists run simulations on computers.[32]

The focus of this paper is on the acquisition of knowledge, and for this purpose it suffices to show that, when computer simulations produce new

---

[31] Cf. Stöckler (2000), p. 368. Note that D is *not* an analogue of what Norton (1991), p. 131 calls the elimination thesis and Gendler (1998), p. 401 calls the dispensibility thesis. The latter theses claim that thought experiments could be replaced using arguments *that are not thought experiment-like*.

[32] D does not imply that the reconstructing argument itself is dispensable. Also, the use of the computer is only dispensable in principle, but not in practice: Many epistemic tasks cannot be completed without computers in practice. As Humphreys (2004, Section 5.5; 2009, pp. 623–624) and Winsberg (2010), p. 5 point out, the perspective of what is possible in principle is often not relevant in the philosophy of science; and to the extent to which the perspective is irrelevant, D is not applicable.

knowledge, the reconstructing argument could do the same. This is obvious. Suppose, for instance, that meteorologists use a computer simulation to cast a prediction about tomorrow's weather, and suppose that the prediction constitutes knowledge. Obviously, this will only be so if the input for the simulations reflects actual initial conditions and if the model upon which the simulations are built reflects the dynamics of the weather. But once we know the initial conditions and the model, we could in principle execute the reconstructing argument to obtain the output of the simulations. In this sense the simulations could be dispensed with.[33]

However, our example from meteorology also points to certain limitations of the argument view. We can understand how the simulations can produce knowledge about tomorrow's weather if the premises of the argument are true. Since some premises reflect the algorithm and the program, the program has to succeed in tracing the values of the relevant characteristics in the target system. But what is the justification to think that the program succeeds in doing this? This question is often difficult to answer because the targets of many simulations are not well understood. Simulation scientists have thus to show that their simulations faithfully reflect the target in the intended respects. In the framework developed in this paper, this is to show that the premises of the reconstructing argument are true. To show this is often thought to be an important part of a computer simulation study. This part is called *validation* of the simulations. Validation is supposed to make a case for the results of a computer simulation study in relation to its target.

Our reconstructions abstract from the problems of validation. They do not explain how the working scientists come to think that the algorithm used in the simulations reflects the dynamics of the target, at least in some respects. They also do not cover the very construction of simulations, a process that is often intimately related to activities of validation. In this sense, our reconstructions abstract from important epistemological problems about computer simulations.

Note though that validation is typically inferential too. Unless it is completely trivial in that the conclusions of a CS are compared to data from the target, validation has to make a case that the conclusions are true. It must thus be possible to reconstruct validations using arguments. It is admittedly not easy to reconstruct validations in terms of arguments because validation is often thought to be messy and unsystematic (e.g. Winsberg 1999, p. 276). But if important activities of validation can be reconstructed using arguments, we may expand the reconstructions obtained in this paper to cover validation too.

---

[33] If we executed the reconstructed argument accurately, we would not obtain the values that the computer outputs, but values that do not suffer from any roundoff errors. That our values differ from the output by the computer does not pose any problem; for epistemological purposes, the simulation becomes certainly dispensable.

The upshot is that validation does not pose a systematic threat to the argument view.[34]

It is also not unreasonable to restrict the focus of a philosophical account of computer simulations to the runs of a simulation program on a machine. This seems the key element of every computer simulation and what is novel about computer simulations too. In the remainder of this paper, I will thus not consider validation and abstract from related problems. My focus will remain on one single run of a simulation, as I announced at the beginning of this paper.

There may seem to be another problem about my claim that the reconstructing argument has the same epistemic power as the corresponding CS. Many simulations are based upon assumptions that are obviously wrong. Almost every mathematical description of a system is based upon abstraction, idealization and approximation. For instance, certain forces are neglected, or the mathematical expression for some force is approximated using a simpler expression. Simulations of fluids assume a so-called eddy viscosity, but people do not think that there is this eddy viscosity in reality.[35] N-body simulations of cosmic structure formation trace the positions of artificial particles, but it is agreed that there are no particles of this type.[36] If we adopt the proposal made in this paper, these simulations should be reconstructed using arguments, and it seems as if the reconstructing argument would have premises that are obviously false. It would then seem that the related arguments cannot teach us anything about the real world. Often, the conclusions of the arguments will be false (or at least not true) in an obvious way. For instance, conclusions about artificial particles from N-body simulations cannot be true of the real world because there are no such particles.[37] Simulations that draw on premises known to be false are nevertheless often believed to produce new knowledge about a target. This seems to be a problem for the argument view, for how can we explain the epistemic power of simulations that are reconstructed using premises and conclusions that are obviously false?[38]

I nevertheless think that the simulations mentioned in the last paragraph do not provide counterexamples against the argument view. One way to

---

[34] Some simulation scientists, e.g. Oberkampf and Roy (2010), Chs. 10–11 recommend that validation include certain experiments, which they call validation experiments. Clearly, such experiments are not merely inferences. But if such experiments are run, it is not very plausible to take them to be literal part of a computer simulation study.

[35] This is a term that is introduced in CSs based upon the Navier-Stokes equations to smooth out shocks. Without eddy viscosity, the simulations can be caught in numerical problems. See e.g. Winsberg (2010), pp. 13–15.

[36] See page 13 above.

[37] Statements that refer to non-existing particles rest upon false presuppositions. They are either wrong, as suggested by Russell (1905) or have a truth gap (Strawson 1950). In either case, they are not true and it seems spurious how they should constitute knowledge.

[38] An argument with false premises and conclusions can of course generate new knowledge if the conclusion is known to be false and used to reject the premises. But we are here concerned with a different case; we refer to arguments the premises and conclusions of which are obviously wrong, e.g. because they assume the wrong type of ontology. Also, the focus of this paper is on simulations the results or conclusions of which produce new knowledge.

explain the epistemic power of the simulations is to say that the reconstructing argument is about a model system distinct from the target, e.g. the system of artificial particles to which N-body simulations refer. The point of the simulations then is to analyze this model system, i.e., to obtain information about the model system.[39] This information can then be used to learn about the target. The production of knowledge is model-based, and the argument view explains how we can obtain knowledge about the model. Admittedly, the argument view does not explain why scientists are justified to transfer knowledge from the model to the target. Additional considerations are necessary to justify this inference. But it should be possible to present these considerations as arguments too. Further, even if it is not, this does not mean that the argument view stumbles. The argument view can explain how scientists gain new knowledge about a model. It can thus explain how computer simulations make a contribution to new model-based knowledge about the real world. This seems sufficient to account for the epistemic power that simulations have in the alleged counterexamples.

An alternative strategy to handle the supposed counterexamples is to change the premises and the conclusions of the reconstructing arguments. In the example of cosmological N-body simulations, the premises would not claim that artificial particles have these and these positions, but rather that the matter distribution in the universe is *as if* it was produced by artificial particles with these and these positions. The conclusion claims in a similar way that the target behaves as if ... In other cases of simulations, the premises may be that the physical characteristics of the target system follow certain equations *to some approximation*. The advantage of the new premises and the new conclusion is that they may be true. If the new premises are known to be correct, we obtain a sound argument the conclusion of which may represent new knowledge. Using this argument, we can explain how the simulations produce new knowledge. This shows again that the simulations can be reconstructed using an argument that has the same epistemic power. The argument does admittedly not include a justification that shows the premises to be true. No reason is given why the target behaves as if ... But this does not provide a reason to reject the argument view. In many cases, a justification of the premises may be represented as an argument too. All in all, simulations that are based upon obviously false assumptions, but that nevertheless produce knowledge do not provide counterexamples against the argument view.

5 The practice thesis

The second part of the argument view is what I have called the practice thesis. According to the thesis, the reconstructing argument is executed when

---

[39] According to Weisberg (2007), the analysis of a model is a crucial step in modeling.

a computer simulation is carried out. My aim in this section is to show that the practice thesis is true.

There is a problem in establishing the practice thesis. In paradigm examples, in which we run through an argument or in which we reason, we consciously move from the premises to the conclusion (we may even speak to us: "If the book is on loan, I have to make a reservation. The book is in fact on loan, so I have to make a reservation."). Clearly, nothing like this happens, when a scientist runs a computer simulation. There is also unconscious reasoning, but when a computer simulation is run, neither the working scientist nor the computer engage in unconscious reasoning.

There is nevertheless a case for the practice thesis. The following considerations are based upon two assumptions that have at least some plausibility. The first assumption is the *extended mind hypothesis* (Clark and Chalmers 1998). I will second assume a certain *view of what reasoning is* (Wedgwood 2006).

In a famous paper, Clark and Chalmers (1998) argue that cognitive tasks can be achieved by systems that extend beyond "skin and scull" (p. 7) of a human being.[40] Suppose, for instance, that Peter uses a pocket calculator to find out what the product of 134 and 4.5 is. Clark and Chalmers (1998) propose to say that Peter and his calculator form a "coupled system" (p. 8) and that there is one cognitive process that achieves the task. This process cannot be located inside Peter's head.

Clark and Chalmers (1998) do not only focus on cognitive *processes*. They also consider mental *states* such as belief and claim that beliefs

> can be constituted partly by features of the environment, when those features play the sort of role in driving cognitive processes (Clark and Chalmers 1998, p. 12).

Their example is a person called Otto who always takes a notebook with him in which he can look up the address of the MOMA. The authors submit that Otto has a belief as to what that address is even if he would not be able to recall the address without using the notebook (ibid., pp. 12–16).

Clark and Chalmers (1998) are not crystal clear as to who or what exactly the bearer of the belief is. They always attribute the belief to Otto, while they also stress that the belief is not just in Otto's head. It may therefore be more appropriate to attribute the belief to the coupled system formed by Otto and his notebook. In the following, I will adopt this viewpoint and assume that a working scientist who runs a computer simulation and a computer form a coupled cognitive system.

I will now argue that this very coupled system executes the reconstructing argument, when a computer simulation is run. As I will also put it, the system reasons through the reconstructing argument. To argue this point, I will draw on an account of reasoning that has been put forward by Wedgwood (2006).

---

[40] See Clark (2007) for another defense of the view. Consult the essays collected in Menary (2010) for more discussion.

I have chosen this account for two reasons. First, the account was not devised to deal with computer simulations and mind-machine interaction. There can thus be no suspicion that the account was tuned to mind-machine interaction from the outset. Second, the account is designed to handle a number of problems (in particular the problem of deviant causal chains; cf. (Frankfurt 1978)). But I do not think that the details of Wedgwood's account matter too much for my purposes, and I should think that my argument would also go through if I used a different account.[41]

Wedgwood's account is Davidsonian in spirit and thus a causal one, the idea being that reasoning is a causal process in which some mental events or states produce other mental events or states (Wedgwood 2006, p. 660).[42] Following Wedgwood (ibid., p. 664), I will leave open whether the relata of the causal relationships under consideration are mental events, mental states or facts. As a matter of convention, I will speak of mental states. Wedgwood's broad idea is that the states upon which a piece of reasoning is based do not only cause a new mental state, but also rationalize it, and that they cause it *because* they rationalize it (ibid., pp. 662, 670).

In more detail, Wedgwood splits each chain of reasoning into *basic steps*. Such basic steps cannot be split into other bits of reasoning any more (ibid., p. 668). In each basic step, some antecedent mental states cause another mental state in virtue of an ability on the part of the reasoner (ibid., p. 671). This ability is a disposition to respond to the fact that the antecedent states rationalize the new mental state (ibid., pp. 670–677).

Wedgwood's example is the formation of belief, and I will also consider belief formation in what follows. My reason is that I am interested in the production of knowledge, and the most straightforward case in which computer simulations lead to knowledge and thus to belief is such that the output of the simulation is thought to contain information about a target system. This will only be the case if the assumptions that enter the simulations reflect knowledge on the part of the scientist. My argument would also go through, if I considered the formation of other types of mental states, for instance, the formation of conditional acceptance. Here conditionally accepting a hypothesis means to accept it for the sake of an argument (cf. ibid., p. 668).

Let us now consider one run of a deterministic computer simulation. This run can be split up into steps in which the values of the variables in the simulation program are updated following the program (there are other steps as well, for instance, when the input is read in, but we can bracket such steps). Most variables have empirical meaning for the scientist. Between the steps, the information is held constant. It does not matter for our purposes how exactly

---

[41] For other very interesting reflections about reasoning see Grice (2001).

[42] Reasoning can also lead one to abandon a mental state, but we can safely bracket this case for our purposes.

the updates are realized in the computer. For instance, the value of a variable may be stored somewhere in the computer.

The extended mind thesis suggests that the states between the updates are belief states of the coupled system. The reason is that the scientist could in principle read out the information and would endorse it. Each update forms a basic step of reasoning in the terms of Wedgwood. This is so because micro-processes that constitute an update cannot be reconstructed as a series of updates of information from the reconstructing argument. And even if they can, this would only mean that the basic steps appear at a more fine-grained level.[43]

Let us now look at a basic step in which information somewhere in the computer is updated. This gives rise to a new mental state with the content that a certain characteristic has a certain value. That this new mental state arises can be explained using features of the earlier states of the programmed computer hardware and using the beliefs of the scientist. The earlier computer states themselves jointly with the beliefs of the scientist constitute mental states of the coupled system. So we can say that the new mental state together with its content is explained using earlier mental states of the coupled system. Moreover, the causation manifests a general disposition to proceed from some states to other states according to some rules. These rules are implemented in the program and endorsed by the scientist. Finally, the scientist certainly thinks that, whenever the computer updates the value of a variable and if the coupled system thus arrives at a new mental state, this new state is rationalized by the earlier mental states (i.e., the inference is thought to be legitimate), and the transition would not be made at the level of mental states if the scientist would think that it yields an illegitimate inference. The disposition is thus responding to the fact that some mental states rationalize others.

My conclusion is that running a computer simulation can be seen as a process in which a coupled system reasons through the reconstructing argument.[44] This is not to deny that all kinds of complications may arise. For instance, it is not always realistic to assume that the information in the memory is always updated sequentially. This assumption is obviously false for parallel computing. Also, the working scientist may have false beliefs about how the simulations work. In such cases, the account given thus far faces problems. I cannot here discuss these problems, but I am confident that they can be solved and I leave them for future work. For the purposes of this paper, it seems

---

[43] Wedgwood only introduces basic steps of reasoning to exclude external deviant causal chains (ibid., pp. 660–665). I do not think it to be absolutely necessary that the steps are “atomic” in that no analysis into other steps of reasoning is possible. We have only to guarantee that the basic steps entirely belong to the cognitive system. If this is right, it is immaterial whether the basic steps are “atomic” or not.

[44] In this way, we obtain an argument in which the algorithm corresponds to many premises. I have suggested on p. 20 that we can always unpack the algorithm into a larger number of assumptions or premises.

sufficient to show in broad outline how running a computer simulation may be thought of as the execution of an argument.[45]

Note that, under my account, the scientist is absolutely crucial at every basic step. The reason is that there would not be mental states if there were no scientist. The content of the mental states is entirely borrowed from attitudes that the scientists has. In this sense, my account is not very speculative. Others may go further and suggest that computers themselves can think and reason. Whether this is so or not is immaterial for the purposes of my account. I take this to be an advantage.

Let me now consider two objections against the practice thesis concerning computer simulations. They do not concern the details, but rather its spirit.

According to the *first objection*, the practice thesis P_{CS} suffers from the problem that, according to Section 4, many reconstructions of a computer simulation can be given. The worry is that, if we cannot tell which of the reconstructing arguments is the right one, the practice thesis cannot be true. And it seems often difficult to say that one of the reconstructions is the right one.

To rebut the objection, note first that a coupled system may run through several reconstructing arguments of one and the same simulation at the same time. There is thus no need to pick exactly one argument. I further think that, at least in paradigmatic cases of computer simulations, there are facts of the matter as to which arguments are executed and which are not. These facts are largely determined by the way a working scientist thinks about her simulation. There are facts as to what the scientist takes to be the results of the simulation and how large she takes to be the error bounds, for instance, and such facts help to fix the argument. Further, our notion of a CS puts mild constraints on the argument that is executed qua simulation.[46]

There are admittedly hard cases. For instance, a simulation may primarily be run with an eye to a particular target system, but the working scientist may be aware of the fact that the simulation may be thought to refer to a larger class of systems. The question then is which argument was executed, when the simulation was run. Was it an argument concerning the single target system, or an argument concerning the larger class? Or were both arguments executed at the same time? This question may be difficult to decide.

But similar problems arise for everyday reasoning as well. Suppose, Peter reasoned to find out how much it would cost him to go to Oxford next Sunday. Was his reasoning about a particular trip to Oxford next Sunday or about a trip to Oxford on *a* Sunday or about both? There may not always be a

---

[45] One may object that Wedgwood (2006) take pains to isolate the reasoner from her environment in order to rule out external deviant causal chains (ibid., pp. 665–666). This may suggest that Wedgwood's account is only applicable to human reasoners. I do not agree with this objection. What Wedgwood has to achieve if the account is to work, is to isolate the *cognitive system* that is the bearer of some chain of reasoning. If the extended mind thesis is true, then cognitive systems may extend beyond a human being.

[46] Recall page 19.

clearcut answer to this question. I am therefore inclined to think that some indefiniteness as to what the argument is does no harm.

The *second objection* is as follows. When a scientist does simulations, she will typically not just start the simulations and wait for the results; she has first to program the computer, and she may rerun the simulation with different inputs and different parameter values. Some simulations allow one to change certain parameters by hand while the simulation is running. This rich practice seems important for simulation science, but it is neglected in my reconstructions.

As a reply, I wish to stress that my focus is on one single run of a simulation and not on a CS study (cf. page 4 above). When I restrict myself in this way, I do not deny that rerunning and varying simulations is important. I also admit that the term "doing simulations" may cover more than a single run of a simulation. But I do think that this richer practice of doing simulations should be analyzed in terms of a couple of interconnected, though distinct steps. One type of step is to run a simulation program once. In this paper, I want to understand what this amounts to. As already indicated, this step is crucial for every CS.

Despite of all this, in the final analysis, it may turn out that running a computer simulation is only to execute an argument *with some grain of salt*. Maybe, the practice thesis is only true if we stretch our notions of reasoning or running through an argument a bit. But recall that my ultimate interest is in the epistemology of computer simulations. I would like to defend the practice thesis or some variety of it to make sure that the reconstructing argument is not an idle ex-post reconstruction, but that the argument explains how we can acquire new knowledge using computer simulations. I would be in trouble if it could be shown that a completely different way of acquiring new knowledge would be at work. I do not think that this is so.

It is interesting to compare computer simulations to thought experiments in this respect. Some authors, e.g. Cooper (2005) think that to run through a thought experiment feels very different from reasoning through an argument.[47] Often, the result of a thought experiment is not consciously derived, but comes to mind without any efforts. And if we read through a thought experiment, we often feel immediately that the result is right. This is a close parallel to experience, and TEs have often been described as quasi-experiential.[48] These observations about the phenomenology of TEs may be taken to indicate that TEs draw on other epistemic sources than those that Norton mentions. For instance, TEs may make explicit prior, but inarticulate knowledge (cf. Mach 1926, 186–189), or they may rely on pictorial thinking (Gendler 2004, pp. 1156–1162).

Matters are different in the case of computer simulations. CSs do not have a phenomenology suggesting that the argument view about them misses

---

[47] See the quotation on my page 8 above.

[48] E.g. Gendler (2004), p. 1154.

something of epistemological significance. In particular, the result does not come to mind in a quasi-experiential manner.

Nevertheless, some authors think that the epistemic power of computer simulations derives from the fact that they are literally *experiments*. For instance, Morrison (2009) argues for this claim. Computer simulations are in fact often called computer experiments. It is also true that, to the working scientist, running a computer simulation feels very much like doing an experiment: In both cases the result cannot be anticipated by other means and may even come as a surprise to her. Finally, computer simulations may take the role of real experiments.

This is not the place to criticize the subsumption of CSs under experiments in detail. But let me briefly make a few points. I do not wish to deny that computer simulations may be regarded as a kind of experiment. However, I do not think that their epistemic power derives from their being an experiment. The epistemic power of an experiment derives from the fact that the system of interest is observed, after it has been set up or manipulated. In this way experiments gather experience, as the name suggests. This is not so with computer simulations. To run a simulation about a certain fluid is not to observe this fluid nor to manipulate it. It is of course true that the experimenter sets up and manipulates the computer and then "observes" it at a later stage. But the computer itself is not the target system, and experiments require an observation of the target. An inference from the computer to the target system can only be made if a lot of assumptions concerning the target system are adopted. I conclude that the epistemic force of computer simulations does not derive from the fact that they are experiments (if they are so at all), and that the argument view is superior since it brings to the fore the assumptions and models upon which simulations are built.[49]

I conclude that a case for the practice thesis P_{CS} has been made and that the argument view about computer simulations has been established.

6 A comparison between computer simulations and thought experiments

If both thought experiments and computer simulations are arguments, we have some common floor to compare them. In this section, I will assume that TEs and CSs are arguments and note commonalities as well as differences.

Let me stress once more that I have not here established the argument view about TEs and that I have some reservations about it. If the argument view about TEs is not the full truth about TEs, our comparison will miss certain aspects. But as even opponents of this view have to admit, many thought experiments can be reconstructed as arguments that capture the power of their TEs quite well (see above), and this will be sufficient for much what follows.

---

[49] I am here in substantial agreement with Barberousse et al. (2009). See also Giere (2009) for a criticism of Morrison's argument.

In comparing TEs and CSs we have to distinguish between *essential and necessary* characteristics of TEs and CSs, on the one hand, and traits that TEs and CSs only have *typically*, on the other hand. When we focus on the latter traits, we exclude TEs and CSs that could in principle be done but that are not in fact done. For instance, no working scientist would run a computer simulation to obtain a result that can immediately be derived in thought.

Qua arguments, TEs and CSs can be compared in three important respects: We can look at their *content*, their *form* and the *use* that is made of them.

Let us start with *content*: It is essential to both TEs and CSs that they involve concrete scenarios (real or imagined) with objects that have properties and that stand in certain relations to each other. TEs and CSs refer to real or imagined target systems. We would not speak of a thought experiment if no concrete system were posited or assumed. Likewise, computer simulations must refer to a target if they are to count as simulations. It is true that some TEs and CSs refer to *classes* of target systems rather than to one single target system, but in this case the systems from the class are very similar in that they contain similar objects that bear the same relationships to each other or so. A general derivation of the uncertainty relation, by contrast, is an argument, but would neither count as a thought experiment nor as a potential computer simulation. The concrete scenario with particular objects and relationships is missing.

Typical simulations are much more specific in content than typical TEs. When a scientist runs computer simulations, she has to specify the values that certain empirical characteristics take in some coordinates. In typical simulations in climate science or in cosmology, tons of characteristics need to be specified in the premises. A similar point holds about the conclusion of a CS: It is typically very specific too. In some cases scientists may use scaling arguments or other arguments to generalize the conclusions of a simulation, but it is often more appropriate to say that such a generalization is not proper part of the simulation itself. Typical thought experiments, by contrast, are much less specific. Some of them such as Newton's bucket thought experiment can be described in purely qualitative terms. The essential point of Galileo's thought experiment with the falling bodies can be made on the basis of statements to the effect that some body carries more heaviness than another one (although it is true that the thought experiment is presented using more specific statements in the original text). As a consequence, the conclusions from TEs are commonly less specific.[50]

Let us now turn to the *form* of TEs and CSs. From a historical point of view, TEs are not often first presented as deductively valid arguments; they are sometimes based upon hidden premises (cf. Norton 2004b, p. 50). But later on, some discussion about the TE may arise, the hidden premises are discovered and explicitly stated such that a deductively valid argument arises. It is sometimes quite difficult to discover the dependency of a thought

---

[50] Recall Humphreys's point that CSs are the modern successors of TEs (see my page 2 above and Humphreys 2004, p. 115).

experiment on hidden premises.[51] A good example for a TE that has been subject to intense logical scrutiny is the EPR argument first presented by Einstein et al. (1935).[52] Nevertheless, the reconstructions of TEs in terms of arguments by Norton (1996) mostly provide deductively valid arguments, the only exception being the bucket experiment, which runs an inference to the best explanation (Norton 1996, pp. 347–349). It is not very different with CSs. If we abstract from roundoff and truncation errors and the arguments used in validation, CSs are deductive arguments, and no inference to the best explanation enters. A difference to TEs though is that computer simulations involve approximations. Regarding hidden premises, there is no principled problem to read off the premises from a computer simulation. However, working scientists who use a computer simulation may not be aware of some such premises.

Many CSs take the form of calculations. This need not be so for TEs, which may be purely qualitative. As a further difference, typical CSs provide much longer and more complicated arguments than do TEs. It would be very difficult for a human being to run through the argument of a computer simulation, which is precisely the reason why computers are used.

There is an interesting parallel here between CSs and the use of the computers in mathematical proof. Tymoczko (1979) has analyzed the computer-aided proof of the four-color theorem and argued that this proof is quite different from other proofs since it is formalizable, but not surveyable any more. Here a proof counts as surveyable if it can be

> looked over, reviewed, verified by a rational agent (Tymoczko 1979, 59).

Arguments that reconstruct computer simulations are not surveyable in this sense. This is important for the question of whether computer simulations may be explanatory. I will leave this question at one side because my interest here is in knowledge.

Let us finally consider the *use* that scientists make of TEs and CSs. Insofar as both are arguments, they can in principle be used in many ways and for many purposes. CSs are indeed put to many uses; they can figure in prediction, tests and in the design of experiments. The typical use of TEs seems more limited. One could of course cast predictions using thought experiments, but this is not done in practice for good reasons. Thought experiments are only feasible for simple scenarios that are easy to describe. Typically, the number of relevant causes and forces is stipulated to be small. As a consequence, the scenarios are too idealized to be realistic. The main use of TEs seems to be confined to textbooks, where TEs are used to expound theories and to trace basic consequences of certain hypotheses (e.g. to derive some basic effects that

---

[51] Cf. the discussion about Galileo's TE with the falling bodies; see e.g. Atkinson and Peijnenburg (2004).

[52] See Fine (2009) and references therein for philosophical discussion about the EPR argument.

a theory predicts). TEs are very well suited to this aim since they are built upon concrete scenarios that can be intuitively grasped. Also, the underlying argument is fairly simple and straight-forward. In this way, they provide an intuitive understanding to the student. Computer simulations are also often used for pedagogical purposes, and they also provide some understanding, but in a different way. In education, they can replace experiments of well-understood phenomena: The student is allowed to vary the details of the setting and the initial conditions and can therefore explore the effects of some target system. But in these cases, the argument that underlies the simulations is typically intransparent for the student, and she does not understand how a few hypotheses entail the predicted effects.

A couple of famous TEs have figured centrally in controversial discussions of theories and are thus linked to theory change. Examples are provided by Galileo's experiment with the falling bodies, Einstein's train TE (Einstein 1920, pp. 11–27) and the EPR experiment Einstein et al. (1935). The potential of TEs to initiate conceptual change was stressed by Kuhn (1964). Common computer simulations, by contrast, seem too complicated and too intransparent as to be able to motivate conceptual change.

Typical CSs take into account a plurality of causes, forces etc. For this reason, holism is much more an issue about computer simulations than about TEs: When a computer simulation ends up with an outcome that seems wrong or is unwelcome for other reasons, there are often too many details that may be changed in order to avoid this outcome.[53] Also, as emphasized before, computer simulations start often with very specific assumptions, and for this reason they do not typically tell us general lessons.

There are other respects in which TEs and CSs can be compared, for instance, the way the argument is executed. As already mentioned, the conclusions of TEs often come to mind spontaneously without any conscious effort. This is different from the way CSs arrive at a conclusion. But this difference does not seem to be important for the purposes of this paper unless it indicates that the argument view is inappropriate for TEs, which would be a point stretching beyond the scope of this paper.

The result of our comparison is that there are a number of differences between typical CSs and typical TEs even if they are both thought to be arguments. As a consequence, the argument views about CSs and TEs do not imply that CSs and TEs are exactly the same thing; they only imply that they draw on the same epistemic sources.

7 Conclusions

Scientists often gain new knowledge about a target system by simulating this very system using a computer, or so it is often claimed. How is this possible and

---

[53] Cf. Lenhard and Winsberg (2010).

how can we explain this? The argument view answers this question by saying that computer simulations are arguments. They generate new knowledge in the same way as arguments do more generally. In more detail, computer simulations can be reconstructed as arguments, where the arguments fully capture the epistemic power of the simulations. Furthermore, when a computer simulation is run, an argument is executed by a coupled system formed by the simulation scientist and the computer, or so I have argued. The argument view is thus not only about reconstruction, but also about what actually happens when a simulation is run.

The focus was on one run of a simulation program. I have argued that this run can be reconstructed as an argument. To some part, the premises of this argument specify the initial state of the target system. Otherwise, the premises state that the system behaves (in some respects) as if the algorithm of the simulation were applied to the empirical characteristics of the system. Often, the algorithm can be unpacked in a number of premises that state how certain characteristics of the system evolve in time.

The argument view about simulations has a number of merits. For one thing, it provides a simple explanation of how one can obtain new knowledge using computer simulations. The explanation is not problematic in any way because it is uncontroversial that we can gain new knowledge by going through arguments.

Second, the argument view about computer simulations seems also promising if we are to understand other epistemic work that computer simulations do more generally. This is clearly so when CSs complete their tasks in virtue of producing knowledge, as was suggested in the introduction to this paper. But the argument view is also fruitful if this condition is not fulfilled. Consider explanation as an example. It is arguable that some scientific computer simulations provide explanations. If computer simulations are arguments and if explanations are arguments (or are at least built upon arguments), it is obvious how computer simulations can figure in explanation (cf. Norton 1996, fn. 11 on p. 339).[54] An open question though is what kind of understanding we gain by running computer simulations.[55] Computer simulations are arguments that are not consciously followed and that are not often surveyable. The question is how much and what kind of understanding we can obtain in this way.[56] In any case, there is reasonable hope that the argument view can accommodate the various uses of computer simulations.

---

[54] Explanations are arguments according to the DN-model of explanation (Hempel and Oppenheim 1948). Even though many objections have been levelled against the DN-model (see Salmon 1989, particularly pp. 46–50, for an overview), most of them do not cast doubts on the idea that explanations are arguments. That explanations provide arguments is also important for e.g. the unificationist view of explanation (see Friedman 1974 and Kitcher 1981 for classic references).

[55] See Friedman (1974) for the connection between explanation and understanding.

[56] For an interesting assessment of the explanatory power of artificial society simulations see Grüne-Yanoff (2009). See also Weber (1999) for simulations and explanation.

As a third advantage, the argument view furnishes a formal account of computer simulations. Such a formal account can be useful to clarify the content of a simulation. We have above suggested a standard way to state the result of a computer simulation.

Fourth and finally, the argument view provides a common floor for comparing thought experiments and computer simulations, at least if the argument view about thought experiments has it right (Norton 1996).

Using the argument view, we can also explain how computer simulations provide representations of their targets, as e.g. Winsberg, Lenhard and Küppers emphasize. Computer simulations represent because they successively derive statements either about the target or about a model thereof. In the latter case, statements about the model are translated to the target to learn about it. In either case, working scientists obtain information about the target via the statements. That this is crucial for successful representation by models has been suggested by e.g. Bailer-Jones (2003).[57]

If the reconstructing argument is naturally thought to refer to a model of the target, then there is another relationship to models. Winsberg's point that computer simulations provide direct representations of the target can be understood as saying that the arguments of a computer-simulation do not often trace a prior model of the target, but rather define a new one.

The most elaborate account given so far of how computer simulations produce knowledge is due to Barberousse et al. (2009). The authors propose a multi-layer analysis of computer simulations qua processes. According to this analysis, the computer undergoes a sequence of states that can be interpreted as computational states; at a different level of description, a sequence of representations of the target arises. This analysis is compatible with the argument view defended in my paper because computations can be reconstructed in terms of arguments executed by the coupled system that is formed by the working scientist and the computer. Further, the argument refers to the target system or a model thereof and thus provides representations of the target. But I think that the argument view yields a more detailed description of how computer simulations produce knowledge.[58] Further, it provides an interesting parallel to the argument view about thought experiments.

In this paper, I have restricted myself to deterministic computer simulations. It remains to be shown whether Monte Carlo simulations can be accommodated in the argument view.[59] I think that they can, but I leave this for future work.

Philosophical work about computer simulations is quite generally faced with a difficult challenge. The simulations carried out in the sciences turn out to be diverse and to serve many purposes. It is desirable to provide a general,

---

[57] Cf. Suárez (2004)

[58] The paper by Barberousse et al. (2009) has a lot to say about the process of running simulations, but it does not end up with a clear statement about how this process produces knowledge.

[59] See Gillespie (1976) and Gillespie (1977) for famous examples of Monte Carlo simulations.

though informative account of computer simulations that faithfully represents the work that computer simulations do in science. I believe that the argument view is a promising start.

Acknowledgements An earlier version of this paper was presented at the workshop "Thought experiments and Computer Simulations" at the IHPST, Paris in March 2010. I'm grateful to the organizers and the other participants. Special thanks to John Norton for discussion and encouragement.

References

Atkinson, D., & Peijnenburg, J. (2004). Galileo and prior philosophy. *Studies in History and Philosophy of Science, 35*, 115–136.

Bailer-Jones, D. M. (2003). When scientic models represent. *International Studies in the Philosophy of Science, 17*, 59–75.

Balzer, W. (2009). *Die Wissenschaft und ihre Methoden* (2nd ed.). Karl Alber, Freiburg und München.

Barberousse, A., Franceschelli, S., & Imbert, C. (2009). Computer simulations as experiments. *Synthese, 169*, 557–574.

Bartuccelli, M. V., Gentile, G., & Georgiou, K. (2001). On the dynamics of a vertically-driven damped planar pendulum. *Proceedings of the Royal Society of London, Series A, 457*, 1–16.

Bertschinger, E. (1998). Simulations of structure formation in the universe. *Annual Review of Astronomy and Astrophysics, 36*, 599–654.

Bishop, M. A. (1999). Why thought experiments are not arguments. *Philosophy of Science, 66*, 543–541.

Brown, J. R., & Fehige, Y. (2010). Thought experiments. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (winter 2010 ed.).

Brown, J. R. (1991). *The laboratory of the mind: Thought experiments in the natural sciences*. Routledge, London.

Brown, J. R. (2004). Peeking into Plato's haeven. *Philosophy of Science, 71*, 1126–1138.

Clark, A. (2007). Curing cognitive hiccups: A defense of the extended mind. *Journal of Philosophy, 104*, 163–192.

Clark, A., & Chalmers, D. J. (1998). The extended mind. *Analysis, 58*(1), 7–19.

Cohnitz, D. (2006). *Gedankenexperimente in der Philosophie*. Mentis, Paderborn.

Cooper, R. (2005). Thought experiments. *Metaphilosophy, 3*, 328–347.

Dolag, K., Borgani, S., Schindler, S., Diaferio, A., & Bykov, A. M. (2008). Simulation techniques for cosmological simulations. *Space Science Reviews, 134*, 229–268. Preprint under 0801.1023v1.

Efstathiou, G., Davis, M., White, S. D. M., & Frenk, C. S. (1985). Numerical techniques for large cosmological N-body simulations. *Astrophysical Journal Supplement Series, 57*, 241–260.

Einstein, A., Podolsky, B., & Rosen, N. (1935). Can quantum-mechanical description of physical reality be considered complete? *Physical Review, 47*, 777–780.

Einstein, A. (1961). *Relativity, the special and the general theory. A popular exposition*. Methuen, London, 1920, here quoted after edition published by Crown, New York.

Fine, A. (2009). The Einstein-Podolsky-Rosen argument in quantum theory. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Fall 2009 ed.). http://plato.stanford.edu/archives/fall2009/entries/qt-epr/.

Frankfurt, H. G. (1978). The problem of action. *American Philosophical Quarterly, 15*, 157–62.

Friedman, M. (1974). Explanation and scientific understanding. *Journal of Philosophy, 71*, 5–19.

Frigg, R. P., & Reiss, J. (2009). The philosophy of simulation: Hot new issues or same old stew? *Synthese, 169*, 593–613.

Frigg, R. P., Hartmann, S., & Imbert, C. (Eds.) (2009). *Models and simulations. Special Issue. Synthese* (Vol. 169, pp. 425–626).

Frigg, R. P., Hartmann, S., & Imbert, C. (Eds.) (2011). *Models and simulations 2. Special Issue. Synthese* (Vol. 180, pp. 1–77).

Galilei, G. (1933). *Le Opere di Galileo Galilei*. Florence: G. Barbèra.

Galilei, G. (1974). *Two new sciences*. Translation by S. Drake. Madison (WI): University of Wisconsin Press.

Galison, P. (1996). *Computer simulations and the trading zone*. In P. Galison & D. J. Stump (Eds.), *The disunity of science. Boundaries, contexts, and power* (pp. 118–157). Stanford: Stanford University Press.

Galison, P. (1997). *Image and logic. A materical culture of microphysics*. Chicago: University of Chicago Press.

Gendler, T. S. (1998). *Galileo and the indispensability of scientific thought experiment*. *British Journal for the Philosophy of Science, 49*, 397–424.

Gendler, T. S. (2000). *Thought experiment. On the powers and limits of imaginary cases*. New York: Garland Publishing.

Gendler, T. S. (2004). Thought experiments rethought and reperceived. *Philosophy of Science, 71*, 1152–1163.

Giere, R. N. (2004). How models are used to represent. *Philosophy of Science, 71*, 742–752.

Giere, R. N. (2009). Is computer simulation changing the face of experimentation? *Philosophical Studies, 143*(1), 59–62.

Gillespie, D. T. (1976). A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *Journal of Computational Physics, 22*, 403–434.

Gillespie, D. T. (1977). Exact stochastic simulation of coupled chemical reactions. *The Journal of Physical Chemistry, 81*, 2340–2361.

Gramelsberger, G. (2010). *Computerexperimente. Zum Wandel der Wissenschaft im Zeitalter des Computers*. Transcript, Bielefeld.

Gramelsberger, G. (2011a). Generation of evidence in simulation runs: Interlinking with models for predicting weather and climate change. *Simulation & Gaming, 42*(2), 212–224.

Gramelsberger, G. (2011b). What do numerical (climate) models really represent?. *Studies in History and Philosophy of Science, Part A 42*, 296–302. Model-Based Representation in Scientific Practice.

Grice, P. (1989). Logic and conversation. In H. P. Grice (Ed.), *Studies in the ways of words* (pp. 1–143). Cambridge, MA: Harvard University Press.

Grice, P. (2001). *Aspects of reason*. Oxford: Oxford University Press.

Grüne-Yanoff, T. (2009). The explanatory potential of artificial societies. *Synthese, 169*, 539–555.

Hartmann, S. (1996). The world as a process: Simulations in the natural and social sciences. In R. Hegselmann, K. G. Troitzsch, & U. Mueller (Eds.), *Modelling and simulation in the social sciences from the philosophy of science point of view* (pp. 77–100). Dordrecht: Kluwer. Quoted from the revised version at http://philsci-archive.pitt.edu/archive/00002412/.

Hempel, C. G., & Oppenheim, P. (1948). Studies in the logic of explanation. *Philosophy of Science, 15*, 135–175.

Hockney, R. W., & Eastwood, J. W. (1988). *Computer simulation using particles* (special student ed.). ed., Adam Hilger, Bristol etc.

Humphreys, P. (1990). Computer simulations. In *PSA: Proceedings of the biennial meeting of the philosophy of science association 1990* (pp. 497–506) (English).

Humphreys, P. (2004). *Extending ourselves: Computational science, empiricism, and scientific method*. New York: Oxford University Press.

Humphreys, P. (2009). The philosophical novelty of computer simulation methods, *Synthese, 169*, 615–626.

Irvine, A. D. (1991). Thought experiments in scientific reasoning. In: T. Horowitz & G. J. Massey (Eds.), *Thought experiments in science and philosophy* (pp. 149–165). Savage, MD: Rowman and Littlefield.

Kitcher, P. (1981). Explanatory unification. *Philosophy of Science, 48*, 507–531.

Klypin, A. (2000). Numerical simulations in cosmology I: Methods. astro-ph/0005502.

Kronsjö, L. (1979). *Algorithms: Their complexity and efficiency*. Chichester: Wiley.

Kuhn, T. S. (1964). A function for thought experiments. In *L'Aventure de la Science, Mélanges Alexandre Koyré, Hermann, Paris* (Vol. 2). (Reprinted in Kuhn, T. S., The Essential Tension:

Selected Studies in Scientific Tradition and Change. Chicago: University of Chicago Press, 1977, pp. 240–265, 307–334).

Kühne, U. (2005). *Die Methode des Gedankenexperiments*. Suhrkamp, Frankfurt am Main.

Küppers, G., & Lenhard, J. (2005a). Computersimulationen: Modellierungen 2. Ordnung. *Journal for General Philosophy of Science, 36*(2), 305–329.

Küppers, G., & Lenhard, J. (2005b). Validation of simulation: Patterns in the social and natural sciences. *Journal of Artificial Societies and Social Simulation, 8*(4), 3.

Lenhard, J. (2007). Computer simulation: The cooperation between experimenting and modeling. *Philosophy of Science, 74*, 176–194.

Lenhard, J., & Winsberg, E. (2010). Holism, entrenchment, and the future of climate model pluralism. *Studies in History and Philosophy of Science Part B, 41*(3), 253–262.

Mach, E. (1926). *Erkenntnis und Irrtum. Skizzen zur Psychologie der Forschung* (5th ed.). Johann Ambrosius Barth, Leipzig (coincides with the 4th ed.).

McAllister, J. (2004). Thought experiments and the belief in phenomena. *Philosophy of Science, 71*, 1164–1175.

Menary, R. (Ed.) (2010). *The extended mind*. Cambridge, MA: MIT Press.

Morrison, M. (2009). Models, measurement and computer simulation: The changing face of experimentation. *Philosophical Studies, 143*, 33–57.

Moue, A., Masavetas, K. A., & Karyianni, H. (2006). Tracing the development of thought experiments in the philosophy of natural sciences. *Journal for General Philosophy of Science, 37*, 61–75.

Norton, S. D., & Suppe, F. (2001). Why atmospheric modeling is good science. In P. Edwards & C. Miller (Eds.), *Changing the atmosphere* (pp. 67–106). Cambridge, MA: MIT Press.

Norton, J. D. (1991). Thought experiments in Einstein's work. In T. Horowitz & G. J. Massey (Eds.), *Thought experiments in science and philosophy* (pp. 129–144). Savage, MD: Rowman and Littlefield.

Norton, J. D. (1996). Are thought experiments just what you thought? *Canadian Journal of Philosophy, 26*, 333–366.

Norton, J. D. (2004a). On thought experiments: Is there more to the argument? *Proceedings of the 2002 Biennial Meeting of the Philosophy of Science Association, Philosophy of Science, 71*, 1139–1151.

Norton, J. D. (2004b). Why thought experiments do not transcend empiricism. In C. Hitchcock (Ed.), *Contemporary debates in the philosophy of science* (pp. 44–66). Oxford: Blackwell.

Oberkampf, W. L., & Roy, C. J. (2010). *Verification and validation in scientific computing*. Cambridge: Cambridge University Press.

Pang, T. (2006). *An introduction to computational physics* (2nd ed.). Cambridge, MA: Cambridge University Press.

Parker, W. S. (2009). Does matter really matter? Computer simulations, experiments, and materiality. *Synthese, 169*(3), 483–496.

Peebles, P. J. E. (1980). *The large scale structure of the universe*. Princeton, NJ: Princeton University Press.

Perret-Gallix, D. (2002). Simulation and event generation in high-energy physics. *Computer Physics Communications, 147*(1–2), 488–493.

Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. (2007). *Numerical recipes. The art of scientific computing* (3rd ed.), New York: Cambridge University Press.

Russell, B. (1905). On denoting. *Mind, 14*, 479–493. (Reprinted in Russell, Bertrand, Essays in Analysis, London: Allen & Unwin, pp. 103–119 (1973))

Salmon, W. (1989). *Four decades of scientific explanation*. Minneapolis: University of Minnesota Press.

Soter, S. (2007). Are planetary systems filled to capacity? *Scientific American, 95*(5), 424.

Stöckler, M. (2000). On modeling and simulations as instruments for the study of complex systems. In M. Carrier, G. J. Massey, & L. Ruetsche (Eds.), *Science at the century's end: Philosophical questions on the progress and limits of science* (pp. 355–373). Pittsburgh, PA: University of Pittsburgh Press.

Strawson, P. F. (1950). On referring. *Mind, 59*, 320–344.

Suárez, M. (2003). Scientific representation: Against similarity and isomorphism. *International Studies in the Philosophy of Science, 17*, 225–244.

Suárez, M. (2004). An inferential conception of scientific representation. *Philosophy of Science, 71*, 767–779.

Tymoczko, T. (1979). The four-color problem and its philosophical significance. *Journal of Philosophy, 76*, 57–83.

Weber, K. (1999). *Simulation und Erklärung*. Waxmann, Münster.

Wedgwood, R. (2006). The normative force of reasoning. *Noũs, 40*, 660–686.

Weisberg, M. (2007). Who is a modeler? *British Journal for Philosophy of Science, 58*, 207–233.

Winsberg, E. (1999). Sanctioning models. The Epistemology of Simulation. *Science in Context, 12*, 275–292.

Winsberg, E. (2001). Simulations, models, and theories: Complex physical systems and their representations. *Philosophy of Science (Proceedings), 68*, 442–454.

Winsberg, E. (2010). *Science in the age of computer simulations*. Chicago: University of Chicago Press.