Available online at www.sciencedirect.com
ScienceDirect
Neural Networks 21 (2008) 311–321
Neural Networks
www.elsevier.com/locate/neunet

2008 Special Issue

Some neural networks compute, others don't☆

Gualtiero Piccinini*

Department of Philosophy, University of Missouri – St. Louis, St. Louis, MO 63121-4400, USA

Received 6 August 2007; received in revised form 28 November 2007; accepted 11 December 2007

Abstract

I address whether neural networks perform computations in the sense of computability theory and computer science. I explicate and defend the following theses. (1) Many neural networks *compute*—they perform computations. (2) Some neural networks compute *in a classical way*. Ordinary digital computers, which are very large networks of logic gates, belong in this class of neural networks. (3) Other neural networks compute *in a non-classical way*. (4) Yet other neural networks *do not perform computations*. Brains may well fall into this last class.

Keywords: Connectionism; Neural network; Computation; Mechanism; Cognition; Brain

1. Introduction

In this paper, I draw and apply two distinctions: (i) between classical and non-classical computation and (ii) between connectionist computation and other connectionist processes.

Connectionist systems are sets of connected signal-processing units. Typically, they have units that receive inputs from the environment (input units), units that yield outputs to the environment (output units), and units that communicate only with other units in the system (hidden units). Each unit receives input signals and delivers output signals as a function of its input and current state. As a result of their units' activities and organization, connectionist systems turn the input received by their input units into the output produced by their output units.

A connectionist system may be either a concrete physical system or an abstract mathematical system. An *abstract* connectionist system may be used to model another system to some degree of approximation. The modeled system may be either a *concrete* connectionist system or something else; e.g., an industrial process.

Psychologists and neuroscientists use abstract connectionist systems to model cognitive and neural systems. They often propose their theories as alternatives to classical, or "symbolic", computational theories of cognition. According to classical theories, the brain is analogous to a digital computer (Fodor & Pylyshyn, 1988; Gallistel & Gibbon, 2002; Newell & Simon, 1976; Pinker, 1997; Rey, 1997). According to connectionist theories, the brain is a (collection of) connectionist system(s).

But given the standard way connectionist systems are defined, classical and connectionist theories are not necessarily in conflict. Nothing in the definition of 'connectionist system' prevents the brain, a concrete connectionist system par excellence, from being a (classical) digital computer.

'Connectionist system' is more or less synonymous with 'neural network'. Brains, of course, are neural networks. More precisely, there is overwhelming evidence that nervous systems carry out their information processing, cognitive, and control functions primarily in virtue of the activities of the neural networks they contain. Thus, it should be uncontroversial that brains are concrete connectionist systems and cognition is explained by connectionist processes. Both connectionists and classicists should agree on this much.

To bring out the contrast between the two theories, we need a more qualified statement: according to *paradigmatic* connectionist theories, the brain is a *non-classical* (collection of) connectionist system(s). This statement is informative only insofar as there is a nontrivial distinction between classical and non-classical systems. This is not the same as the distinction between systems that compute and systems that do not. We

☆ An abbreviated version of some portions of this article appeared in Piccinini (2007c) as part of the IJCNN 2007 Conference Proceedings, published under IEE copyright.

* Tel.: +1 314 516 6160; fax: +1 314 516 4610.
E-mail address: piccininig@umsl.edu.

0893-6080/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.neunet.2007.12.010

should be able to ask whether any, some, or all non-classical connectionist systems are computational. A clear answer to this question is needed to resolve a dispute about the nature of cognition.

Many mainstream connectionist theorists agree with classicists that brains perform computations and neural computations explain cognition (Bechtel & Abrahamsen, 2002; Churchland, 1989; Churchland & Sejnowski, 1992; Cummins & Schwarz, 1991; Eliasmith, 2003; Feldman & Ballard, 1982; Hopfield, 1982; Koch, 1999; Marr & Poggio, 1976; O'Brien & Opie, 2006; Roth, 2005; Rumelhart & McClelland, 1986; Schwartz, 1988; Shagrir, 2006; Smolensky & Legendre, 2006). The claim that distinguishes such connectionist computationalists from classicists is that according to connectionist computationalism, non-classical connectionist (computing) systems are a better model of the brain than classical computing systems. In reply, some classicists argue that (non-classical) connectionist systems do not perform computations at all (Fodor, 1975; Gallistel & Gibbon, 2002; Pylyshyn, 1984). According to such classicists, only classical systems perform genuine computations. This does not bother a different group of connectionist theorists, who reject or downplay the claim that brains compute (Edelman, 1992; Freeman, 2001; Globus, 1992; Horgan & Tienson, 1996; Perkel, 1990).

Who is right? Do brains compute? Do (non-classical) connectionist systems compute? Which kind of system – classical or non-classical, computational or non-computational – is the best model for the brain?

Making progress on these debates – between classicists and connectionists as well as between computationalists and anti-computationalists – requires independently motivated distinctions between, on the one hand, systems that compute vs. systems that do not, and on the other hand, classical vs. non-classical systems. By applying such distinctions to connectionist systems, we can find out which connectionist systems, if any, do or do not perform computations, and which, if any, are classical or non-classical. Yet it has proven difficult to draw such distinctions in a satisfactory way.

The same problem may be framed in terms of theories of cognition. Is cognition explained by *non-classical, connectionist computations*? The answer depends on both what the brain does and where we draw two lines: (i) the line between connectionist computation and other kinds of connectionist processing and (ii) the line between classical computation and non-classical computation. Drawing these lines in a satisfactory way is a contribution to several projects: a satisfactory account of computation, a correct understanding of the relationship between classical and connectionist theories of cognition, and an improved understanding of cognition and the brain.

Unfortunately, different participants in these debates use 'computation' in different ways. I will base my discussion on computation as the subject matter of computability theory (aka recursion theory) and computer science. This is the sense of 'computation' that is characterized by certain powerful mathematical theorems, comes with a special explanatory style (computational explanation), and inspired computational theories of cognition.

2. Do connectionist systems compute?

Accounts of the nature of computation have been hindered by the widespread view that computing requires executing programs. Several authors embrace such a view (Fodor, 1975; Pylyshyn, 1984).[1] Some authors endorse something even stronger: "To compute function g is to execute a program that gives o as its output on input i just in case $g(i) = o$. Computing reduces to program execution" (Cummins, 1989, p. 91) (Roth, 2005). The weaker view – namely, that program execution is a necessary condition for genuine computing – is strong enough for our purposes. Such a view is plausible when we restrict our attention to at least *some* classical systems. The same view gives rise to paradoxical results when we consider connectionist systems.

The view that computation requires program execution leads to a dilemma: either connectionist systems execute programs or they do not compute. Different people have embraced different horns of this dilemma.

A computationalist who is opposed to (paradigmatic) connectionist theories might wish to deny that connectionist systems – or at least, paradigmatic examples of connectionist systems – perform computations. Here is something close to an outright denial: "so long as we view cognition as computing in any sense, we must view it as *computing over symbols*. No connectionist device, however complex, will do" (Pylyshyn, 1984, p. 74, italics original). A denial that connectionist systems compute is also behind the view that connectionism is not a truly computationalist framework, but rather, say, an associationist framework, as if the two were mutually exclusive (Gallistel & Gibbon, 2002).[2]

In light of the thesis that computing requires executing programs, rejecting the idea that connectionist systems perform computations may sound like a reasonable position. Unfortunately, this position does not fit with the observation that the input–output mappings produced by many paradigmatic connectionist systems may be characterized by the same formalisms employed by computability theorists to characterize classical computing systems.

It is difficult to deny that many paradigmatic examples of connectionist systems perform computations in the same sense in which Turing machines and digital computers do. The first neural network theorist to call his theory 'connectionist'

[1] Another view that has hindered our understanding of computation is that computation requires representation (Cummins, 1983; Fodor, 1975; O'Brien & Opie, 2006; Pylyshyn, 1984; Shagrir, 2006). I have argued against it elsewhere (Piccinini, 2004a, 2007a, 2008a), so I will set it aside.

[2] A related red herring, coming from an anti-computationalist perspective, is the claim that certain systems (cognitive systems, *some* connectionist systems) do not compute because they are *dynamical* (Port & van Gelder, 1995; van Gelder, 1995, 1997, 1998). As we shall see, Van Gelder is right that *some* connectionist systems do not compute. But not because they are dynamical—computing systems are dynamical too! The relevant question is, how should we draw the line between those dynamical systems that compute and those that do not? That is one topic of this paper.