Computation vs. Information Processing:
How They Are Different and Why It Matters

Draft - 6/26/2008

Gualtiero Piccinini
Department of Philosophy
University of Missouri – St. Louis
599 Lucas Hall (MC 73)
One University Blvd.
St. Louis, MO 63121-4499 USA
Email: piccininig@umsl.edu

Andrea Scarantino
Department of Philosophy
Brains & Behavior Program
Georgia State University
P.O. Box 4089
Atlanta, GA 30302-4089 USA
Email: ascarantino@gsu.edu

Since the cognitive revolution, it's become commonplace that cognition involves both *computation* and *information processing*. Is this one claim or two? Is computation the same as information processing? The two terms are often used interchangeably, but this usage masks important differences. In this paper, we will distinguish information processing from computation and examine some of their mutual relations, shedding light on the role each can play in a theory of cognition. We will conclude by recommending that theorists of cognition be explicit and careful about which notions of computation and information they employ, as well as how they connect them together. Much confusion can be avoided by doing so.

Of course, one might explicitly stipulate that 'computation' and 'information processing' designate the same thing. As Smith (2002) points out, computation is sometimes *construed* as information processing. This tells us nothing about the independently established meanings of the terms—the meanings that are historically most influential and theoretically most important.

So we will set aside any stipulation that identifies computation and information processing. To understand the interesting relations between them, we need to take into account why and how the terms were introduced in the sciences of mind and brain. Once we do that, we can see that computation and information processing are different things and why it matters to keep them distinct.

To begin, we will introduce the main historically relevant and independently established uses of ‘computation’ and ‘information processing’. Then, we will discuss how computation and information processing may or may not go together. One moral of our story will be that, contrary to a common view, the thesis that cognition involves computation doesn’t follow from the thesis that cognition involves information processing.

1. Historical Preliminaries

Before proceeding, a brief reminder will help put our topic in the right historical perspective. The notions of computation and information came into the sciences of mind and brain from different places through different - though partially overlapping - routes. And they came to play different roles. Lacking room for a detailed historical treatment, we will limit ourselves to the following brief points.

The notion of computation came into cognitive science from mathematics, where a computation, in its original sense, is an algorithmic process: a process that generates correct results by following an effective procedure. This notion was made formally precise by Alan Turing and other logicians in the 1930s (Church 1936, Gödel 1934, Post 1936, Turing 1936-7, etc.). A few years later, neurologists Warren

McCulloch and Walter Pitts (1943) used Turing's formal notion of computation to characterize the activities of the brain. McCulloch and Pitts's computational theory had a large influence on computer design and artificial intelligence, and eventually, on the sciences of mind and brain (Piccinini 2004a). Via McCulloch and Pitts, Turing's notion of computation became a way to recruit the tools of the logician (such as proof-theoretic methods) and those of the computer scientist (algorithms, computer programs, certain neural networks) in order to characterize the functionally relevant properties of psychological and neural processes.

By contrast, the notion of information came into cognitive science from control engineering (Rosenblueth, Wiener, and Bigelow 1943; Wiener 1948) and communication engineering (Shannon 1948). For engineering purposes, information is what is transmitted by messages carried either within a system for purposes of feedback or across a physical channel for purposes of reproduction at a destination. Informal notions of information had been invoked in neurobiology as early as Edgar Adrian (1928; cf. Garson 2003), but no systematic formal analysis of the sense in which signals carry information had been provided prior to the mid-century. Norbert Wiener and Claude Shannon shared the insight that the transmission of information is fundamentally related to the reduction of uncertainty or entropy, and that the latter can be formally quantified.

But whereas Wiener (1948) focused on the role of information in the control of both mechanical and biological systems, Shannon (1948) was interested in the narrower question of how to transmit information efficiently across communication channels through proper encoding of the messages. Shannon's work immediately influenced psychologists such as George Miller, who put Shannon's technical notion of information at the foundation of a new science of mind (Miller and Frick 1949, Miller 1951; cf. Garner 1988, Crowther-Heyck 1999). Shannon's information theory entered neuroscience around the same time (MacKay and McCulloch 1952). Measures of information inspired by Shannon have been used ever since to quantify various aspects of neural signals (Rieke et al. 1997, Dayan and Abbott 2001). Shannon's (1948) theory also inspired philosophical theories of information (Dretske 1981) and the broader *informational semantics* movement, which seeks to use information to naturalize mental content.

Right around the time they entered psychology and neuroscience, the notions of computation and information merged into an appealing synthesis. Roughly, the mind/brain was seen as a computer, whose function is to receive information from the environment, process it, and use it to control the body and perform intelligent actions (Wiener 1948; McCulloch 1949; Jeffress 1951; Ashby 1952; Bowden 1953; Shannon and McCarthy 1956; Miller 1956; von Neumann 1958; Miller, Galanter, and Pribram 1960; Feigenbaum and Feldman 1963). Since then, computation and information processing have become almost inseparable—and often indistinguishable—in much literature on the mind and brain.

As it turns out, both 'computation' and 'information' are polysemous concepts, namely, concepts with multiple, yet related meanings. Our first task is to distinguish some importantly different varieties of computation and information.

2. Computation

The notion of computation has been characterized in many ways. For present purposes, different notions of computation vary along two important dimensions. The first is how encompassing a notion is, that is, how many processes it includes as computational vs. how many it excludes. The second dimension has to do with whether being computational requires possessing semantic properties. Let's look at the first dimension first.

2.1 Digital Computation

We use 'digital computation' for whatever notion is implicitly defined by the classical mathematical theory of computation, most famously associated with Turing. By this, we do not mean to appeal to a specific formalism, such as Turing machines. We mean to appeal to the class of formalisms of which Turing machines are an example and the kind of mathematical theorems that may be proven about them—including, of course, theorems about functions that are not computable by Turing machines (for an introduction, see Davis et al., 1994). We also do not mean to restrict ourselves to processes that follow algorithms. We include any process whose input-output relations can be characterized by the kind of mathematics originally developed by Turing and other computability theorists. These include the processes performed by many popular kinds of neural networks.[1]

There have been different attempts to explicate how the notion of digital computation applies to concrete physical systems. We will rely on the account one of us has developed in recent years (Piccinini 2007a). Roughly, digital computation is the processing of strings of digits according to general rules defined over the digits. A digit in this sense need not represent a number—it is simply a particular or a state whose type can be reliably distinguished from other types by the computing system.

It is important not to confuse the present notion of digital computation with the notion of *classical* computation sometimes employed in cognitive science debates.

---

[1] Note for participants in the conference on Computation in Cognitive Science: These qualifications should make clear that our notion of digital computation appears to be broader than Aizawa's notion of "Turing equivalent computation" (Aizawa, "Computation in Cognitive Science: It is not all about Turing equivalent computation," ms.).

To a first approximation, classical computation is digital computation performed in accordance with a step-by-step algorithm, perhaps in response to an internal representation of the rule that defines the computation. Digital computation includes classical computation as a special case. Digital computation in the present sense requires neither that the rules that define the computation be represented in the computing system, nor that the rules constitute a classical algorithm followed step-by-step by the computing system. All the rules need to do is specify what relationship obtains between the strings of digits constituting the input and the strings of digits constituting the output.

Many important distinctions may be drawn within digital computation, such as hardwired vs. programmable, special purpose vs. general purpose, and serial vs. parallel (cf. Piccinini 2008a, b). Nevertheless, digital computation is the most restrictive notion of computation that we will consider here. It includes processes that follow ordinary algorithms and effective procedures, such as the computations performed by standard digital computers, as well as many types of connectionist processes, including those performed by standard backpropagation networks. Since digital computation is the notion that inspired the computational theory of cognition—in both its classical and connectionist variants—it is the most relevant notion for present purposes.[2]

---

[2] In contrasting the notion of classical computation with that of connectionist process, we are following common usage (Fodor and Pylyshyn 1988). (Incidentally, notice that Fodor and Pylyshyn's notion of classical system is even more restrictive than the one we employ here, since for them, classical computations are defined over interpreted symbols and are sensitive to the combinatorial syntax of the symbols.) But that contrast is a false one. Not only are many paradigmatic types of connectionist computation digital in the present sense, some of them—such as those of McCulloch-Pitts nets (McCulloch and Pitts 1943)—are classical too. Speaking more carefully, we may use the notion of digital computation to taxonomize theories of cognition into classical theories (which may or may not be connectionist), non-classical computational (connectionist) theories, and non-computational (connectionist) theories (Piccinini 2008a).

2.2 Generic Computation

Digital computation is traditionally contrasted with analog computation (for a detailed account of this contrast, see Piccinini 2008b, Section 3.5). Analog computation, in turn, is a vague and slippery concept. The most clearly defined notion of analog computation is that of computation by an analog computer (Pour-El 1974). Roughly, an analog computer is a device whose function is to process continuous variables so as to solve systems of differential equations. It is worth remembering that shortly after McCulloch and Pitts argued that the brain performs digital computations, others offered the alternative theory that the brain performs analog computations, meaning, roughly, processes like those performed by analog computers (Gerard 1951, see also Rubel 1985).

The theory that the brain is literally an analog computer was never very popular, perhaps for the good reason that the main vehicles of neural processes appear to be trains of neuronal spikes, which do not look much like continuous variables.[3] In fact, they are made of all-or-none spikes, which are quite discontinuous events; this was the main reason for McCulloch and Pitts to conclude that neural processes are digital computations.

In recent decades, many neuroscientists started using the term 'computation' for the processing of neuronal spike trains (i.e., sequences of spikes produced by neurons in real time). We will call the processing of neuronal spike trains by neural systems 'neural computation'. Whether neural computation is a form of digital computation, analog computation, or a distinct type of process is a difficult question,

---

[3] The claim that the brain is *literally* an analog computer (in the sense of Pour-El 1974) should not be confused with the similar-sounding but much fuzzier claim that the brain is an analog computer, where 'analog computer' is used to denote a not-well-specified class of non-classical, and perhaps non-digital, systems. This use of the term 'analog computer' is not uncommon, but we find it misleading.

which we cannot settle here. Instead, we will subsume digital computation, analog computation, and neural computation under the banner of 'generic computation'.

We use 'generic computation' for any *manipulation of medium-independent vehicles*, where a medium-independent vehicle is such that all that matters for its processing are the differences between the values of different portions of the vehicle along a relevant dimension (as opposed to more specific physical properties, such as the vehicle's material composition). Since it is medium independent, a generic computation can be instantiated by many different physical mechanisms, so long as they possess a medium with an appropriate dimension of variation.[4]

For instance, digits are a kind of medium-independent vehicles: they can be implemented by many kinds of physical media (mechanical, electro-mechanical, electronic, etc.), but all that matters for the computations defined over them is that they belong to types that are unambiguously distinguishable by the processing mechanism. In other words, digital computers that manipulate different media may well perform the same computation.

By contrast, analog computers cannot always distinguish between any two portions of the continuous variables they manipulate—there is always a margin of measurement error. So, continuous variables are not digits or strings of digits. Nevertheless, all that matters for analog computations is the differences between the different portions of the variables being manipulated, to the degree that they can be distinguished by the system. Any further physical properties of the media that implement the variables are irrelevant to the computation. Like digital computations, analog computations operate on medium-independent vehicles.

---

[4] This notion of medium-independence was inspired in part by Garson 2003.

Finally, current evidence suggests that the vehicles of neural processes are neuronal spikes, and the functionally relevant aspects of neural processes are medium-independent aspects of the spikes—primarily, spike rates (as opposed to any more concrete properties of the spikes). Thus, spike trains are another case of medium-independent vehicle. In conclusion, generic computation includes digital computation, analog computation, neural computation (which may or may not reduce to digital or analog computation), and perhaps more. Assuming that brains process spikes trains and spikes are medium-independent vehicles, it follows by definition that brains perform generic computations.

2.3 Other Notions of Computation

There are many other notions of computation. For instance, some authors maintain that "the 'Theory of Computation' … is neither more nor less than *a mathematical theory of the flow of causality*" (Smith 2002, p. 43, emphasis original). Other authors maintain that every process is, at its most fundamental physical level, a (digital) computation (e.g., Wolfram 2002). Under either of these views, every (causal) process is a computation. These all-encompassing notions of computation are not directly relevant to the theory of cognition, because the theory of cognition attempts to find out what *distinguishes* cognition from other processes—not what it shares with everything else. Insofar as the theory of cognition uses computation to distinguish cognition from other processes, it needs to use a notion of computation that excludes at least some other processes (cf. Piccinini 2007b).

Someone may object as follows. Even if everything is computational, it doesn't follow that the claim that cognition involves computation is vacuous. A theory of cognition still has to specify which specific computations are performed by the brain. The job of neuroscience and psychology is to discover the specific computations that distinguish brains from other entities (cf. Shagrir 2006).

We agree that, *if* brains perform computations, the job of neuroscience and psychology is to discover the specific computations brains perform. But the *if* is important. The job of psychology and neuroscience is to find out how brains work, regardless of whether they compute. The claim that brains compute was introduced in neurosciences and psychology as an empirical hypothesis, to explain some recalcitrant phenomena by an analogy with digital computers. Most of the empirical import of the computational theory of cognition is already eliminated by stretching the notion of computation from digital to generic. Stretching the notion of computation further, from generic to all-encompassing, erases all empirical import from the claim that brains compute.

Here is another way to describe the problem. The view that cognition involves computation has been fiercely contested. Many psychologists and neuroscientists reject it. If we adopt an all-encompassing notion of computation, we have no way to make sense of this fact. It is implausible that critics of computationalism have simply failed to notice that everything is computational. More likely, they object to what they perceive to be questionable empirical commitments of computationalism. For this reason, all-encompassing computation is a poor foundation for a theory of cognition. Hence, we will set aside these all-encompassing notions of computation.

2.4 Semantic vs. Non-Semantic Computation

So far, we have taxonomized different notions of computation according to how encompassing they are. Now we will consider a second dimension along which notions of computation differ. Consider digital computation. The digits are often taken to be representations, because it is assumed that computation requires representation (Fodor 1981). A similarly semantic view may be taken with respect to generic computation and even all-encompassing computation.

(Well, with all-encompassing computation, it seems strange to say that computation requires representation, because computation encompasses more or less everything. Is everything representational? Some people seem prepared to bite this bullet (Shagrir 2006, Smith 2002).)

One of us has argued at length that computation per se, in the sense implicitly defined by the practices of computability theory and computer science, does not require representation, and that any semantic notion of computation presupposes a non-semantic notion of computation (Piccinini 2004b, 2008c). Meaningful words such as 'avocado' are both strings of digits and representations, and computations may be defined over them; nonsense sequences such as '2#r %h@', which represent nothing, are strings of digits too, and computations may be defined over them just as well.

Although computation does not *require* representation, it certainly *allows* it. In fact, generally, computations *are* carried out over representations. For instance, almost without exception, the states manipulated by ordinary computers are representations. So, it will be convenient to consider both semantic and non-semantic notions of computation. Semantic notions of computation define computations as operating over representations. Non-semantic notions of computation define computations without requiring that the vehicles being manipulated be representations.

To summarize, we now have the following taxonomy of computation: digital computation (computation according to computability theory), generic computation (which includes digital computation as a special case), and all-encompassing computation (which is irrelevant to the theory of cognition); in all cases, we may define a semantic notion of computation (in which the computational vehicles are representations) or a non-semantic one.

3. Information Processing

In principle, instances of information processing can differ along two dimensions: the type of processing involved and the type of information involved. We wish to be tolerant on what counts as processing, putting no principled restriction of what types of processing there may be. To differentiate between different notions of information processing, the key term for our purposes is ‘information’.

In the broadest sense, information has to do with the reduction of uncertainty. Signals or messages are informative to the extent that they reduce uncertainty. We distinguish two ways of spelling out this basic notion, which lead respectively to what we call *non-semantic* and *semantic* theories of information. On one hand, the occurrence of a signal or message can reduce uncertainty by being more or less surprising given the prior probability of its occurrence. This notion of information is non-semantic because what signals stand for is irrelevant to how informative they are: all that matters are the prior probabilities that the signals be selected. On the other hand, the occurrence of a signal can reduce uncertainty about some other state of affairs. This notion of information is semantic, because what signals ‘stand for’ is crucial to determine how informative they are. In what follows, first we describe Shannon’s influential theory of non-semantic information, and then we distinguish between two different types of semantic information, natural and non-natural.

3.1 Shannon information

We use ‘Shannon information’ to designate the notion of non-semantic information formalized in Shannon’s (1948) theory of communication.[5] Shannon developed his technical notion of information with a specific objective in mind. He was interested in an optimal solution to what he called the ‘fundamental problem of communication’ (Shannon 1948, 1), that is, the reproduction of messages from an information source to a destination. In Shannon’s sense, any device that produces a succession of “symbols” in a probabilistic manner can count as an information source or destination.

Shannon’s notion of symbol is not the usual one. In the standard sense, a symbol necessarily has meaning or semantic content. But Shannon’s symbols may or may not mean anything. They don’t even have to fall into finitely many types (like digits do); on the contrary, they may be continuous variables. We will continue to discuss Shannon’s theory using the term ‘symbol’ for historical accuracy, but it’s important to remember that symbols in Shannon’s sense are nothing more than physical structures distinguishable by their probability of being selected by the information transmitter and receiver.

To communicate a sequence of symbols produced at a source simply amounts to generating a second sequence of symbols at a destination so as to replicate the original sequence in a way that is satisfactory given some desiderata (accuracy, speed, cost, etc.). The sense in which the semantic aspects are irrelevant is that the only thing that matters for engineering purposes is that the symbols be selected from a discrete set (discrete information source) or a continuous interval (continuous information source) with a certain probability distribution.

---

[5] For a history of Shannon’s theory of communication, see for example Pierce (1980).

This being the case, a nonsense sequence of symbols such '2#r %h@' can in principle generate more information than a meaningful sequence of symbols such as 'avocado'. To get an intuitive grip on this idea, consider an experiment involving a random variable taking values a_1 and a_2 with prior probabilities $p(a_1)=0.9999$ and $p(a_2)=0.0001$ respectively. Before the experiment takes place, outcome a_1 is almost certain to occur, and outcome a_2 almost certain not to occur. The occurrence of both outcomes generates information, in the sense that it resolves the uncertainty characterizing the situation before the experiment takes place (in the absence of uncertainty, e.g., when $p(a_1)=1$ and $p(a_2)=0$, no Shannon information would be produced).

Shannon's intuition was that the occurrence of a_2 generates *more information* than the occurrence of a_1, because it is *less expectable* in light of the prior probability distribution. Consequently, if a sequence of symbols such as '2#r %h@' is more surprising than a sequence of symbols such as 'avocado', it will carry more information.

A further distinction between information as ordinarily understood and Shannon information is that, whereas ordinary information is carried by individual signals, communication theoretic measures of information are average measures. To introduce them, we need a bit of formalism. Let X be an input *ensemble*, namely a random variable taking values in $A_X=\{a1,\ldots,an\}$ with probabilities $p(a_1),..,p(a_n)$ respectively, and let Y be an output *ensemble*, namely a random variable taking values in $A_Y=\{b1,\ldots,br\}$ with probabilities $p(b_1),..,p(b_r)$ respectively. We assume that $p(ai)>0$ for all $i=1,..,n$, $p(bj)>0$ for all $j=1,..,r$, $\sum_{i=1}^{n} p\,(a_i)=1$ and $\sum_{j=1}^{r} p\,(b_j)=1$.

Shannon's central measures of information are the following:

$$H(X)=-\sum_{i=1}^{n} p\,(a_i)\log_2 p(a_i)^{6}$$

$$I(X,Y)=\sum_{i=1}^{r} p(ai,bj)\log\frac{p(ai/bj)}{p(ai)}$$

H(X) is called “entropy”, and it measures the *average information* produced by the selection of values in A_X={a1,…,an}. The higher the uncertainty as to which symbol will be selected, the higher will be the entropy (the entropy will be maximal when all values have an equal probability of being selected). I(X,Y) is called “mutual information”, and it is defined as the difference between the entropy characterizing X on average *before* and *after* Y takes values in A_Y. Shannon proved that X carries mutual information about Y whenever X and Y are statistically dependent, i.e., whenever it is not the case that p(ai,bj)=p(ai)p(bj) for all i and j. This is to say that the transfer of mutual information between two sets of uncertain outcomes A_X={a1,…,an} and A_Y={b1,…,br} amounts to the *statistical dependency* between the occurrence of outcomes in A_X and A_Y (information is ‘mutual’ because statistical dependencies are symmetrical).

The communication-theoretic measures of entropy and mutual information served well Shannon’s purpose of finding optimal engineering solutions to the ‘fundamental problem of communication’. Soon after its appearance, however, the theory started being used in many fields for which it had not been designed, including neuroscience and psychology, economics, cryptography, and quantum physics. For

---

[6] The logarithm to the base b of a variable x – expressed as $\log_b x$ - is defined as the power to which b must be raised to get x. In other words, $\log_b x=y$ if and only if $b^y=x$. We stipulate that the expression 0 $\log_b 0$ in case any of the addenda of H(X) is equal to 0. Shannon (1948, 1) pointed out that in choosing a logarithmic function he was following Hartley (1928), and added that logarithmic functions have nice mathematical properties, are more useful practically because a number of engineering parameters “vary linearly with the logarithm of the number of possibilities”, and are “nearer to our intuitive feeling as to the proper measure” of information.

instance, Shannon information is commonly used by neuroscientists to measure the quantity of information carried by neural signals about a stimulus and estimate efficiency of coding (what forms of neural responses are optimal for carrying information about stimuli) (Dayan and Abbott 2001, chap. 4).

What worried Shannon was mostly the association of semantic aspects to the programmatically non-semantic notion of information he had formulated. "Workers in other fields", Shannon emphasized, "should realize that the basic results of the subject are aimed in a very specific direction . . . the hard core of information theory is, essentially, a branch of mathematics, a strictly deductive system" (Shannon in Sloane (1992, 462)). In what follows, we will keep this lesson in mind, and carefully distinguish between Shannon information and semantic information, which is a theoretical refinement of the ordinary notion of information.

3.2 Semantic Information

We call 'semantic information' the information a signal carries by reducing uncertainty *about* some state of affairs. In this case, semantic aspects are crucial: what the signal "stands for" is key to determining what information it carries. Moreover, the theorist is no longer primarily interested in the average information carried by *sets* of signals, but rather in the information carried by *individual* signals.

We distinguish two forms of semantic information, which depend on what grounds the "standing for" relation: natural (semantic) information and non-natural (semantic) information (cf. Grice 1957). Roughly speaking, signals carry *natural* information by standing in an appropriate physical relation with what they are about. This is the sense in which smoke produced by fire carries information about fire. By contrast, signals carry *non-natural* information by being arbitrarily associated with what they are about by means of a convention. This is the sense in which the word ‘smoke’ carries non-natural information about smoke. We will now explore the distinction between theories of semantic information in a bit more detail.

3.2.3 Natural (Semantic) Information

When smoke carries natural information about fire, a vervet monkey call carries natural information about a flying predator, or a bee dance carries natural information about the location of nectar, what is the basis for this informational link? The idea at the core of theories of natural information is that the connection between informative signals and what they are about is physical.

In the contemporary literature, we find three main accounts of the physical connection that distinguishes informationally related from informationally unrelated events: *probabilistic* (Dretske 1981, Millikan 2004), *nomic* (Dretske 1981, Fodor 1990), and *counterfactual* (Loewer 1983, Cohen and Meskin 2006). These contributions find their primary source of inspiration in Dretske’s (1981) attempt to develop a theory of semantic information on the basis of Shannon’s (1948) communication theory. (Interesting remarks on natural information can also be found in Peirce 1931 and Grice 1957, see below.)

What got Dretske interested in Shannon’s (1948) theory was primarily that Shannon had treated information as an objective, physical commodity. Dretske wanted information to be objective in order to use it for the naturalization of intentional content (Stampe (1975, 1977) is an earlier attempt in this direction). But he also wanted information to capture the semantic aspects of communication Shannon had explicitly excluded. Accounting for these semantic aspects, Dretske suggested, requires constructing a theory of information around the idea that “information is that commodity capable of yielding knowledge” (Dretske 1981, 44).

An important consequence follows from this approach: since knowledge entails truth, so does natural information. The intuition here is that since one cannot come to know that p from a signal unless p, the falsity of p makes it impossible for a signal to carry the natural semantic information that p.

Dretske developed a number of communication-theoretic constraints for the transmission of semantic natural information, and defended as uniquely capable of satisfying them the following probabilistic definition: "A signal r carries the information that s is F = The conditional probability of s's being F, given r (and k), is 1 (but given k alone, less than 1)" (Dretske 1981, 65).

This account of information was soon criticized for including reference to k, which designates the knowledge state of the receiver about what possibilities exist at the source.[7] Dretske's critics thought that defining information in terms of knowledge, a state already imbued with intentional content, prevented the use of information for the naturalization of intentional content.

Two main strategies have been put in place to circumvent this problem. The first is to identify the physical connection between informative signals and what they are about with a nomic connection (Dretske 1981, Fodor 1990). Under this view, smoke carries information about fire just in case there is a nomic correlation between smoke and fire.

The second strategy is to identify the physical connection between informative signals and the states of affairs they are about with a counterfactual connection (Cohen and Meskin 2006). Under this view, smoke carries information about fire just in case smoke would not have occurred had fire not occurred.

---

[7] Another prominent criticism was that Dretske's (1981) theory sets too high a bar for the transmission of information, because signals in natural environments rarely raise the probability of what they are about to 1.

In the rest of our paper, we focus on nomic theories of natural information, which are currently the received view in informational semantics (for a discussion of the relation between nomic and counterfactual theories of information, see Scarantino, forthcoming).

3.2.4 Non-natural (Semantic) Information

In ordinary language, we talk about false information or misinformation. If someone were to whisper in your ear, "Hey, your opponent has an ace," when, in fact, your opponent has a queen, you might say that you received false information. There are in principle two strategies to make sense of this way of talking.

On one hand, we could say that you have received what *purported to be* the information that your opponent has an ace, but deny that any actual information was received. Under this view, information entails truth: if your opponent does not have an ace, you cannot possibly receive the information that he does. This is the road taken by most philosophers of semantic information, including Dretske (1981), Millikan (2004), and Floridi (2005).

On the other hand, we could say that you have actually received the information that your opponent has an ace even though, as a matter of fact, he does not. This interpretation presupposes the existence of a kind of semantic information that, unlike natural information, does not entail truth.

We choose the second option, because we are interested in a taxonomy of actual usages of the information concept, and the idea that one can receive information that happens to be false is an important part of ordinary information talk. Moreover, notions of information that do not entail truth are often presupposed by psychologists and computer scientists when they speak of information processing. We call the kind of semantic information that can be false 'non-natural information'.

The natural vs. non-natural distinction imports in the domain of information a distinction Grice (1957) drew between two concepts of meaning. The first kind of meaning is exemplified by a sentence such as "those spots mean measles", which is true, Grice claimed, just in case the patient has measles. The second kind of meaning is exemplified by a sentence such as "those three rings on the bell (of the bus) mean that the bus is full" (1957, 85), which is true even if the bus in not full. Similarly, we stipulate that "x carries natural information that p" entails p, whereas "x carries non-natural information that p" does not entail p.

The distinction between natural and non-natural information maps onto another historically important distinction between kinds of signs introduced by C. S. Peirce. Peirce defined a sign as "something which stands to somebody for something in some respect or capacity" (Peirce 1931). But the "standing for" relation can differ among signs.

Peirce distinguished three main varieties of signs: icons, indices, and symbols. Icons stand for something by virtue of resembling it. This is the sense in which the portrait known as "La Gioconda" stands for Lisa del Giocondo, born Gherardini. Indices stand for something by virtue of being physically connected to it. This is the sense in which a knocking on the door stands for someone's presence at the door. Symbols stand for something by virtue of a convention. The meaning-grounding convention can either be the effect of an explicit stipulation, as in the case of artificial languages, or the effect of spontaneous emergence, as in the case of natural languages. This is the sense in which we say that the word 'smoke' stands for smoke.

Our taxonomy maps onto the Peircean taxonomy as follows: Peircean indices are bearers of natural information, and Peircean symbols are bearers of non-natural information. Examples of bearers of non-natural semantic information can be found in languages (words, lexical concepts, etc.), and more generally, whenever a non-natural meaning is attached to a signal by virtue of a convention (the "three rings means full bus" case).

What is the relation between natural and non-natural information? As mentioned above, one of the main philosophical motivations for theorizing about natural information is the project of using it to naturalize intentional content. We have now introduced a notion of information that, by corresponding to non-natural meaning, is already imbued with intentional content. We emphasize that by calling such information non-natural we are not taking a stance on whether it can be naturalized. We are simply employing Grice's terminology to distinguish between two semantic notions of information. We wish to remain neutral on the prospects of reducing one to the other.

3.4 Other Notions of Information

There are other technical notions of information, such as Fischer information and information in the sense of algorithmic information theory. They play a valuable role within their respective fields. In some cases, they are related in interesting ways to the notion of computation. There is even an all-encompassing notion of information, somewhat analogous to some all-encompassing notions of computation. Given this all-encompassing notion, allegedly the physical world is, at its most fundamental level, constituted by information (Wolfram 2002). Perhaps, then, all-encompassing computation is the same as all-encompassing information processing, but this thesis is no help to a theory of cognition. Since these other notions of information are not directly relevant to our present concerns, we set them aside.

In what follows, we focus on the relations between computation and the following three notions of information: Shannon information (information according to Shannon's communication theory), natural information (truth-entailing semantic information), and non-natural information (non-truth-entailing semantic information).

The three notions are instantiated in different circumstances. Consider an utterance of the sentence 'I have a toothache'. It carries Shannon information just in case the production of the sequence of symbols *I-blank-h-a-v-e-blank-a-blank-t-o-o-t-h-a-c-h-e* can be modeled by a stochastic process. The same utterance carries natural information about having a toothache just in case the production of the sequence of symbols *I-blank-h-a-v-e-blank-a-blank-t-o-o-t-h-a-c-h-e* is nomically correlated to having a toothache. Carrying natural information about having a toothache entails having a headache. Finally, the same utterance carries non-natural information just in case the sequence of symbols *I-blank-h-a-v-e-blank-a-blank-t-o-o-t-h-a-c-h-e* has a non-natural meaning in a natural or artificial language. Carrying non-natural information about having a toothache does not entail having a toothache.

4. Computation May or May Not Be Information Processing

Are the vehicles over which computations are performed necessarily bearers of information? Is information processing necessarily carried out by means of computation? Answering these questions requires combining several of the senses of 'computation' and 'information' we have distinguished.

We will argue that the vehicles over which computations are performed may or may not bear information, and that information processing may or may not be carried out by means of computation. Yet, as a matter of contingent fact, the vehicles over which computations are performed generally bear information.

4.1 Computation vs. Processing Shannon Information

Few people worry about whether computation is the processing of Shannon information. This is because cognitive scientists are generally interested in semantic information. Nevertheless, for completeness' sake, let's briefly look at why computation need not be the processing of Shannon information, and why the processing of Shannon information need not be computation.

The notion of digital computation does not require that the computation vehicles bear or generate Shannon information. The vehicles of digital computation are strings of digits. Digits *carry* Shannon information *if and only if* they are statistically correlated to the selection of other symbols; they *generate* Shannon information *if and only if* they have a probability between 0 and 1 of being selected (but not 0 or 1). By these criteria, digits need not carry or generate Shannon information.

On one hand, digits need not be statistically correlated with anything else in order to undergo a digital computation. Consider a computer performing multiplication. Before it starts computing, it has input data and an initial internal state, both of which may be characterized as strings of digits. Input data and initial internal states may correlate in some interesting way with some other variable, but they don't have to. In the typical case, they are simply inserted in the computer by its user. More importantly, what matters for a computation to be well defined and carried out successfully is that the digits be there, not that they correlate with anything else. Thus, a computation's digits need not carry Shannon information.

On the other hand, a digit may well have probability 1 of being selected. In fact, that is the typical case. Input digits may well be selected with probability 1—nothing in the notion of digital computation prevents it. In addition, computation may be deterministic or indeterministic, but most computational steps are deterministic. Deterministic computations generate new digits with probability 1. Therefore, most computationally produced digits generate no Shannon information. Thus, a computation's digits need not generate Shannon information.

In summary, since computation is a process defined over digits and digits need not carry or generate Shannon information, digital computing does not entail the processing of Shannon information. (To be sure, however, Shannon's information theory can be usefully applied to studying the efficiency of computer codes (e.g., for the purpose of data compression, eliminating redundancies in a code, or devising the most efficient code for maximizing the average rate of information communicated) as well as the transmission of signals within computers (noise minimization).)

Conversely, processing Shannon information may or may not be done by means of digital computation. First, for it to be done by digital computing, Shannon information must be produced and carried by strings of digits. But Shannon information can be produced and carried by continuous signals, which may be processed by analog means. Second, even when the bearers of Shannon information are digits, there exist forms of processing of such digits other than digital computation. For instance, a digital-to-analog converter transforms digits into analog signals.

In conclusion, digital computation may or may not constitute the processing of Shannon information, and Shannon information may or may not be processed by means of digital computation.

Analogous considerations apply to generic computation. Generic computation may or may not constitute the processing of Shannon information, and vice versa. This case is similar to the previous one. We already established it in the case of digital computation, which is a special case of generic computation. It is just as easy to establish it for other kinds of generic computation, such as analog computation. Nothing in the notion of analog computation mandates that the continuous signals being manipulated produce or carry Shannon information.

The main difference from the previous case is that now, Shannon information processing is necessarily a form of generic computation. As we defined it, generic computation is the manipulation of any medium-independent vehicle. Information is a medium-independent notion, in the sense that whether a vehicle carries information does not depend on its specific physical properties. Thus, generic computation is broad enough to encompass any form of information processing. If a vehicle carries Shannon information, its processing is a generic computation.

4.2 Computation vs. Processing Natural Information

Many neuroscientists use ‘computation’ and ‘information processing’ interchangeably. What they generally mean by ‘information processing' is the processing of natural information carried by neural responses. It is thus important to examine whether and in what sense computation is the processing of natural information.

4.2.1 Digital Computation

The notion of digital computation does not require that the computation vehicles bear natural information. The vehicles of digital computation are strings of digits, which may or may not carry natural information.

Granted, natural information is virtually ubiquitous—it is easy enough to find nomically underwritten correlations between physical variables. This being the case, the digits present within a digital computer may well carry natural information about cognitively interesting sources, such as environmental stimuli that the computer responds to. Consider a car's computer, which receives and responds to natural information about the state of the car. The computer uses feedback from the car to regulate fuel injection, ignition timing, speed, etc. In such a case, a digital computation will be a case of processing natural information.

But the carrying and processing of natural information is optional: digits need not be nomically correlated with any state of affairs in order for a digital computation to be performed over them. Thus, digital computation may or may not constitute the processing of natural information.

This point is largely independent of the distinction between semantic and non-semantic digital computation. Consider a paradigmatic example of semantic digital computation: an arithmetical operation defined over natural numbers. The strings of digits represent numbers, but numbers are not a source of natural information—numbers do not nomically correlate with anything (in the relevant sense). Furthermore, it is quite possible for a digital computation to yield an incorrect output, which misrepresents the outcome of the operation performed. As we have seen, natural information is incompatible with the notion of misrepresentation. Thus, even semantic digital computation does not entail the processing of natural information. This is because semantic digital computation is generally defined over the *non*-natural information carried by the digits—independently of whether they carry natural information or what specific natural information they may carry.

At this stage, someone might reply that non-natural information reduces to natural information—plus, perhaps, some other naturalistic ingredients. Several efforts have been made to this effect (Drestke 1981, 1988; Barwise and Seligman 1997; Fodor 1990). According to these theories, non-natural information is, at least in part, natural information. If this is right, then semantic digital computation entails the processing of natural information.

But we take the possibility that the antecedent is false seriously. Theories according to which non-natural information reduces to natural information are only one family of theories among others. There are other theories according to which either non-natural information is irreducible, or it reduces to things other than natural information (e.g., Millikan 1984, Harman 1987). Since the project of reducing non-natural information to natural information may or may not go through, semantic digital computation may or may not be the processing of natural information.

What about the converse claim: that the processing of natural information must carried out by means of digital computation? Again, it may or may not be. It all depends on whether the natural information is digitally encoded and how it is processed. Natural information may be encoded by continuous variables, which cannot be processed by digital computers. Or it may be encoded by continuous variables, whose processing consists in converting them to analog signals. In both of these cases, non-natural information is processed, but not by means of digital computation.

In sum, digital computation vehicles need not carry natural information, and natural information processing need not be performed by digital computation.

4.2.2 Generic Computation

Generic computation may or may not constitute the processing of natural information. This case is analogous to the previous one. We already established it in the case of digital computation, which is a special case of generic computation. It is just as easy to establish it for other kinds of generic computation, such as analog computation. Nothing in the notion of analog computation mandates that the continuous variables being manipulated carry natural information. More generally, nothing in the notion of manipulating medium-independent vehicles mandates that such vehicles carry natural information.

The main difference from the previous case is that the converse thesis does not hold. Now, natural information processing is necessarily a form of generic computation, because, again, the notion of generic computation was intentionally left broad enough to encompass all the relevant processes.

4.3 Computation vs. Processing Non-natural Information

Non-natural information is a central notion in our discourse about minds and computers. We attribute conventional semantic contents to each other's minds. We do the same with the digits manipulated by computers. We will now examine whether and in what sense computation is the processing of non-natural information.

4.3.1 Digital Computation

Digital computation may or may not be the processing of non-natural information, and vice versa. This case is analogous to that of natural information, with one major difference. In the case of *semantic* digital computation, any reasonable notion of semantic content turns digital computation into non-natural information processing.

As a matter of fact, virtually all (digital) computing conducted in artifacts *is* information processing, at least in the ordinary sense of non-natural information.

But the main question we are trying to answer is whether necessarily, computation is semantic information processing. It *is* if we define a notion of digital computation that entails that the digits have semantic content. But as pointed out in Section 2.4, nothing in the notion of digital computation used by computer scientists and computability theorists mandates such a definition. So, in the theoretically most important sense of digital computation—digital computation as implicitly defined by the practices of computability theory and computer science—digital computation need not be semantic information processing.

Conversely, the processing of non-natural information need not be carried out by means of digital computation. It all depends on whether the non-natural information is digitally encoded and how it is processed. It may be encoded digitally, or by continuous variables, or by other vehicles. And even when encoded digitally, it may be manipulated by digital computation or by other processes, such as a digital-to-analog conversion.

In conclusion, digital computation vehicles need not carry non-natural information (though they do if the notion of digital computation is semantic), and the processing of non-natural information need not be a digital computation.

4.3.2 Generic Computation

Generic computation may or may not constitute the processing of semantic information, for similar reasons. Again, the converse does not hold. Semantic information processing must be done by means of generic computation, for this notion of computation is broad enough to encompass the relevant processes.

4.4 What We Have Accomplished So Far

So, is computation information processing? We are now in a position to summarize the theses we defended:

Question: Is digital computation information processing?

Thesis 1: Digital computation (semantic or non-semantic) may or may not be the processing of Shannon information & The processing of Shannon information may or may not be a digital computation (semantic or non-semantic).

Thesis 2: Digital computation (semantic or non-semantic) may or may not be the processing of natural information & The processing of natural information may or may not be a digital computation (semantic or non-semantic).

Thesis 3: Non-semantic digital computation may or may not be the processing of non-natural information, Semantic digital computation must be the processing of non-natural information, & The processing of non-natural information may or may not be a digital computation (semantic or non-semantic).

Question: Is generic computation information processing?

Thesis 4: Generic computation (semantic or non-semantic) may or may not be the processing of Shannon information & The processing of Shannon information must be a generic computation (semantic or non-semantic).

Thesis 5: Generic computation (semantic or non-semantic) may or may not be the processing of natural information & The processing of natural information must be a (semantic) generic computation.

Thesis 6: Non-semantic generic computation may or may not be the processing of non-natural information, Semantic generic computation must be the processing of non-natural information, & The processing of non-natural information must be a (semantic) generic computation.

These six theses summarize what we consider to be the relations of conceptual entailment, or lack thereof, between 'computation' and 'information processing', depending on what one means by each notion. It remains to be seen whether any of this matters.

5. Why the Difference between Computation and Information Processing Matters

'Information processing' and 'computation' are both Protean terms, used in all kinds of ways. If they are used vaguely enough, or if their mutual assimilation is stipulated, then the identification of computation with information processing is innocuous enough, but not particularly enlightening. When we move from generic umbrella terms to the more precise notions we have distinguished, however, the mutual relations between computation and information processing become far from obvious, as the six theses we listed testify.

So how is it that the two notions are so often used interchangeably, without a second thought? We suspect the historical reason for this conflation goes back to the cybernetic movement's effort to blend Shannon's information theory with Turing's computability theory (as well as control theory). Cyberneticians were initially clear on the difference between the two. Their idea was that organisms and automata are control mechanisms, information is transmitted within the system and between system and environment for purposes of feedback, and control is exerted by means of digital computation (or perhaps by means of analog computation, another special case of generic computation). The two concepts were still understood independently.

Then the waters got muddier. When the cybernetic movement became influential in psychology, AI, and neuroscience, ‘computation’ and ‘information’ became ubiquitous buzzwords. Many people accepted that computation and information belong together in a theory of cognition. After that, they stopped paying attention to the differences between the two, with the unfortunate consequence that confusions on computation started combining with confusions on information. To set the record straight and make some progress, we must get clearer on the independent roles computation and information can fulfill in a theory of cognition.

The notion of digital computation was imported from computability theory into neuroscience and psychology primarily for two reasons: first, it seemed to provide the right mathematics for modeling neural activity; second, it inherited mathematical tools (algorithms, computer program, formal languages, logical formalisms, and their derivatives, including many types of neural networks) that appeared to capture some aspects of cognition. Whether these reasons are enough to establish that cognition is digital computation is a difficult question, which lies outside the scope of this essay.

The theory that cognition is computation became so popular that it progressively led to a stretching of the operative notion of computation. In many quarters, especially neuroscientific ones, the term ‘computation’ is used, more or less, for whatever internal processes explain cognition. We have included this notion under the rubric of ‘generic computation’. Unlike ‘digital computation,’ which stands for a mathematical apparatus in search of applications, ‘neural computation’ is a label in search of a theory. Of course, the theory is quite well developed by now, as witnessed by the explosion of work in theoretical and computation neuroscience over the last decades (O'Reilly and Munakata 2000, Dayan and Abbott 2001). The point is that such a theory need not rely on a previously existing and independently defined notion of computation, such as 'digital computation' or even 'analog computation' in its most straightforward sense.

By contrast, the various notions of information have entirely distinct roles to play. By and large, they serve to make sense of how organisms keep track of their environments. Shannon information can serve to address quantitative problems of efficiency of communication in the presence of noise, including communication between the environment and the nervous system. Natural information can serve to give specific semantic content to particular states or events. This may include cognitive or neural events, which are often nomically correlated with events occurring in the environment. Finally, non-natural information can serve to characterize the conventional meaning of concepts, words, and the thoughts and sentences they constitute.

Whether cognitive or neural events fulfill all, or any, of these job descriptions is in part an empirical question and in part a conceptual one. It's a conceptual question insofar as we can mean different things by 'information' and 'computation', and insofar as there are conceptual relations between the various notions. It's an empirical question insofar as, once we fix the meanings of 'computation' and 'information', the extent to which computation and the processing of information are coinstantiated in the brain depends on the empirical facts of the matter.

In this essay, we have charted the main notions of computation and information that are relevant in cognitive science as well as their conceptual relations (or lack thereof). To wit, we have found that computational vehicles may or may not carry information; in turn, information processing may or may not be done by means of (digital) computing.

To be sure, cognition involves generic computation. This is a trivial conclusion, however, because we defined 'generic computation' to include any type of process that manipulates medium-independent vehicles. That cognition involves such vehicles is uncontroversial. It would be more interesting if cognition turned out to involve digital computation, which is defined independently of the theory of cognition and is the notion that initially inspired the computational theory of cognition. But the question of whether cognition involves digital computation, or some other nontrivial notion of computation, has yet to be settled.

Regardless of what kind of computation cognition may involve, it doesn't follow that cognition involves the processing of information. In fact, there are theories according to which cognition involves computation (or at least *may* involve computation), but it does not involve the processing of information (Stich 1983, Ramsey 2007). More precisely, what these authors reject is the role of representation in a theory of mind. They don't explicitly reject the notion of information. But notice that the notion of non-natural information is fully representational (it allows for representational error), and even the notion of natural information is typically used by scientists and philosophers to underwrite a notion of representation (Dretske 1981, 1988).[8] Theories according to which cognition involves computation but not representation have occasionally been accused of incoherence, on the grounds that computation (allegedly) requires representation. On the contrary, it is a corollary of

---

[8] Ramsey argues that this is not a genuine notion of representation. We are not persuaded, but we lack the space to do justice to this topic.

our analysis that these theories are perfectly coherent. Whether they are correct, of course, is another story.

It is independently plausible that, as a matter of contingent fact, cognition does involve the processing of information. After all, natural information is carried by neural and cognitive events by virtue of nomic correlations with environmental events, and such correlations are commonplace. Moreover, some aspects of such correlations may be measured by means of Shannon information. Finally, the concepts possessed and the words spoken by human cognizers are generally individuated in part by the non-natural information they carry. That human cognition involves the processing of non-natural information is not only part of commonsense; it is presupposed by many mainstream psychological theories. Whether non-natural information reduces (among other things) to natural information is an important and difficult question, which we cannot take up here.

Ok, but do these distinctions really matter? Why should a cognitive theorist care about the differences between computation and information processing? The main theoretical advantage of keeping them separate is to appreciate the independent contributions they can make to a theory of cognition. Conversely, the main cost of conflating computation and information processing is that the resulting mongrel concept may be too vague to do all the jobs that are required of it.

If it's unclear which notions (if any) of information and computation different theorists employ and how they are related, cross-purpose talk lurks at every corner. It will be difficult to compare and contrast theories; non-equivalent theories may be taken to be equivalent and equivalent theories may be taken to be non-equivalent. Making progress in our understanding of cognition requires clarity and precision on what kind of computation and information processing are involved in cognition, and on whether and how the two are related. If they are mutually related, it is also important to understand whether they are related conceptually, because of how we use the notions of computation and information processing, or whether they are related in ways that constitute an empirical discovery about cognition.

The conflation of computation with information processing leads to fallacious arguments. For instance, we have become accustomed to arguments to the effect that cognition does not involve computation because cognition does not involve representations (e.g., van Gelder 1995). Aside from the implausibility of the premise, the conclusion simply doesn't follow. Computationalism can survive the rejection of representationalism, and vice versa.

Here is another example. Since many cognitive and neural events nomically correlate with environmental events, and since such correlations seem to have much to do with the production of behavior, we may conclude that the brain processes natural information. Since information processing is carried out by means of computations, in turn, we may further conclude that the brain is a computing system. So far, so good—except that we didn't specify which notion of computation we were using. Since we know there is a mathematical theory of computation that establishes many of the properties and limitations of computation, we may be tempted to appeal to such a theory to characterize neural computation. This line of reasoning can be seen at work in much contemporary neuroscience literature (e.g., Churchland and Sejnowski 1992; Koch 1999; Churchland 2007). But the last step is fallacious: at best, we are entitled to the conclusion that the brain performs *generic* computation—once again, a fairly trivial claim. Whether and in what way neural computations are characterized by the results of computability theory—i.e., whether neural computations are a kind of digital computation—is a matter that needs to be established independently.

At this point, some people might be tempted to appeal to the Church-Turing thesis, which says that any computable function is computable by some Turing machine. Since Turing machines and other equivalent formalisms are the foundation of computability theory, doesn't it follow that all computations are covered by the results established by Turing and other computability theorists? Not at all. The Church-Turing thesis applies directly only to *algorithmic* digital computation. The relationship between algorithmic digital computation and digital computation simpliciter, let alone other kinds of computation, is quite complex, and the Church-Turing thesis does nothing to settle it.[9]

So one important consequence of our analysis is that the common view that cognition involves computation simply because it processes information is liable to generate serious mistakes. There are two important senses—corresponding to the two notions of natural and non-natural information—in which, plausibly, cognition involves information processing. Neither of these senses entails that cognition is computation in the historically and theoretically most significant sense of the term, i.e., digital computation (which includes connectionist computation in the most straightforward sense). Thus, cognition may well involve the processing of information. Whether this information is processed by means of (digital) computation remains an open question.[10]

References

Adrian, E. D. (1928). The Basis of Sensation: The Action of the Sense Organs. New York, Norton.
Ashby, W. R. (1952). Design for a Brain. London, Chapman and Hall.

---

[9] For further discussion of some relevant issues, see Piccinini 2007b, c.
[10] One of us believes the answer turns out to be negative (Piccinini 2007d).

Barwise, J. and J. Seligman, 1997. *Information Flow: the logic of distributed systems*. Cambridge Tracts in Theoretical Computer Science, 44.

Bowden, B. V. (1953). Faster than Thought. London, Sir Isaac Pitman.

Church, A. (1936). "An Unsolvable Problem in Elementary Number Theory." The American Journal of Mathematics **58**: 345-363.

Churchland, P. M. (2007). Neurophilosophy at Work. Cambridge, Cambridge University Press.

Churchland, P. S. and T. J. Sejnowski (1992). The Computational Brain. Cambridge, MA, MIT Press.

Cohen, J. and A. Meskin 2006. An Objective Counterfactual Theory of Information, *Australasian Journal of Philosophy* 84: 333–52.

Crowther-Heyck, H. (1999). "George E. Miller, Language, and the Computer Metaphor of Mind." History of Psychology **2**(1): 37-64.

Davis, M. D., R. Sigal, and E. J. Weyuker (1994). Computability, Complexity, and Languages. Boston, Academic.

Dayan, P. and L. F. Abbott (2001). Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems. Cambridge, MA, MIT Press.

Dretske, F. (1981). Knowledge and the Flow of Information. Oxford, Blackwells.

Dretske, F. (1988). Explaining Behaviour. Cambridge, MA, Bradford/MIT.

Feigenbaum, E. A. and J. Feldman (1963). Computers and Thought. New York, McGraw-Hill.

Floridi, L. (2005). "Is Semantic Information Meaningful Data?" Philosophy and Phenomenological Research 70(2): 351-370.

Fodor, J. A. (1981c). "The Mind-Body Problem." Scientific American **244**(January 1981).

Fodor, J. (1990). *A Theory of Content and Other Essays*. Cambridge, MA: MIT/Bradford.

Garner, W. R. (1988). The Contribution of Information Theory to Psychology. The Making of Cognitive Science: Essay in Honor of George A. Miller. W. Hirst. Cambridge, Cambridge University Press**:** 19-35.

Garson, J. (2003). "The Introduction of Information into Neurobiology." *Philosophy of Science* **70.5**: 926-936.

Gerard, R. W. (1951). Some of the Problems Concerning Digital Notions in the Central Nervous System. Cybernetics: Circular Causal and Feedback Mechanisms in Biological and Social Systems. Transactions of the Seventh Conference. H. v. Foerster, M. Mead and H. L. Teuber. New York, Macy Foundation**:** 11-57.

Gödel, K. (1934). On Undecidable Propositions of Formal Mathematical Systems. The Undecidable. M. Davis. Ewlett, NY, Raven**:** 41-71.

Grice, P. (1957), "Meaning," *The Philosophical Review*, 66: 377-88.

Harman, G. (1987). (Nonsolipsistic) Conceptual Role Semantics. New Directions in Semantics. E. LePore. London, Academic Press**:** 55-81.

Hartley, R.V. (1928), '*Transmission of Information*', Bell System Technical Journal, 7, 535-563.

Jeffress, L. A., Ed. (1951). Cerebral Mechanisms in Behavior. New York, Wiley.

Loewer, B. 1983. Information and Belief, *Behavioral and Brain Sciences* 6: 75–6.

Koch, C. (1999). Biophysics of Computation: Information Processing in Single Neurons. New York, Oxford University Press.

MacKay, D. and W. S. McCulloch (1952). "The Limiting Information Capacity of a Neuronal Link." Bulletin of Mathematical Biophysics **14**: 127-135.

McCulloch, W. S. (1949). "The Brain as a Computing Machine." Electrical Engineering **68**: 492-497.

McCulloch, W. S. and W. H. Pitts (1943). "A Logical Calculus of the Ideas Immanent in Nervous Activity." Bulletin of Mathematical Biophysics **7**: 115-133.

Miller, G. A. (1951). Language and Communication. New York, MgGraw-Hill.

Miller, G. A. (1956). "The Magical Number Seven, Plus or Minus a Two: Some Limits on our Capacity for Processing Information." Psychological Review **63**: 81-97.

Miller, G. A. and F. Frick (1949). "Statistical Behavioristics and Sequences of Responses." Psychological Review **56**: 311-325.

Miller, G. A., E. H. Galanter, and K. H. Pribram (1960). Plans and the Structure of Behavior. New York, Holt.

Millikan, R. G. (1984). Language, Thought, and Other Biological Categories. Cambridge, MA: MIT Press.

Millikan, R. G. (2004). Varieties of meaning : the 2002 Jean Nicod lectures. Cambridge, Mass., MIT Press.

O'Reilly, R. C. and Y. Munakata (2000). Computational Explorations in Cognitive Neuroscience: Understanding the Mind by Simulating the Brain. Cambridge, MA, MIT Press.

Peirce, C. S. (1931). *Collected Papers*. Harvard University Press, Cambridge, Massachusetts.

Piccinini, G. (2004a). “The First Computational Theory of Mind and Brain: A Close Look at McCulloch and Pitts’s ‘Logical Calculus of Ideas Immanent in Nervous Activity’.” Synthese **141**(2): 175-215.

Piccinini, G. (2004b). “Functionalism, Computationalism, and Mental Contents.” Canadian Journal of Philosophy **34**(3): 375-410.

Piccinini, G. (2007a). “Computing Mechanisms.” Philosophy of Science **74**(4): 501-526.

Piccinini, G. (2007b). “Computational Modeling vs. Computational Explanation: Is Everything a Turing Machine, and Does It Matter to the Philosophy of Mind?” Australasian Journal of Philosophy **85**(1): 93-115.

Piccinini, G. (2007c). “Computationalism, the Church-Turing Thesis, and the Church-Turing Fallacy.” Synthese **154**(1): 97-120.

Piccinini, G. (2007d). Computational Explanation and Mechanistic Explanation of Mind. Cartographies of the Mind: Philosophy and Psychology in Intersection. M. De Caro, F. Ferretti and M. Marraffa. Dordrecht, Springer**:** 23-36.

Piccinini, G. (2008a). “Some Neural Networks Compute, Others Don’t.” Neural Networks **21**(2-3): 311-321.

Piccinini, G. (2008b). “Computers.” Pacific Philosophical Quarterly **89**(1): 32-73.

Piccinini, G. (2008c). “Computation without Representation.” Philosophical Studies **137**(2): 205-241.

Pierce, J. R. (1980). An Introduction to Information Theory New York, Dover Publications.

Pour-El, M. B. (1974). "Abstract Computability and Its Relation to the General Purpose Analog Computer (Some Connections Between Logic, Differential Equations and Analog Computers)." Transactions of the American Mathematical Society **199**: 1-28.

Post, E. (1936). "Finite Combinatorial Processes. Formulation I." Journal of Symbolic Logic **1**: 103-105.

Ramsey, W. M. (2007). Representation Reconsidered. Cambridge, Cambridge University Press.
Rieke, F. e. a. (1997). Spikes: Exploring the Neural Code. Cambridge, MA, MIT Press.
Rubel, L. A. (1985). "The Brain as an Analog Computer." Journal of Theoretical Neurobiology **4**: 73-81.
Scarantino, A. (forthcoming). "Shell games, information and counterfactuals," *Australasian Journal of Philosophy*.
Shagrir, O. (2006). "Why We View the Brain as a Computer." Synthese **153**(3): 393-416.
Shannon, C. E. (1948). "A Mathematical Theory of Communication." Bell System Technical Journal: 379-423, 623-656.
Shannon, C. E. and J. McCarthy (1956). Automata Studies. Princeton, NJ, Princeton University Press.
Sloane, N. J. A., Wyner, A. D. , Ed. (1992). Claude Elwood Shannon: Collected Papers IEEE Press.
Smith, B. C. (2002). The Foundations of Computing. Computationalism: New Directions. M. Scheutz. Cambridge, MA, MIT Press**:** 23-58.
Stampe, D. (1975). Show and tell. In Freed, B., Marras, A., and Maynard, P., editors, *Forms of Representation: Proceedings of the 1972 Philosophy Colloquium of the University of Western Ontario*. North-Holland Publishing Company, Amsterdam.
Stampe, D. (1977). Toward a causal theory of linguistic representation. In French, P., Uehling, T., and Wettstein, H., editors, *Midwest Studies in Philosophy*, volume 2. University of Minnesota Press, Minneapolis.
Stich, S. (1983). From Folk Psychology to Cognitive Science. Cambridge, MA, MIT Press.
van Gelder, T. (1995). "What Might Cognition Be, if not Computation?" The Journal of Philosophy **XCII**(7): 345-381.
von Neumann, J. (1958). The Computer and the Brain. New Haven, Yale University Press.
Wiener, N., Bigelow J. and Rosenblueth, A. (1943), 'Behavior, Purpose and Teleology', Philosophy of Science, Vol. 10 (1943), 18-24.
Wiener, N. (1948). Cybernetics: Or Control and Communication in the Animal and the Machine. MA: MIT Press.
Wolfram, S. (2002). A New Kind of Science. Champaign, IL, Wolfram Media.