

Information Processing, Computation, and Cognition¹

Draft - 8/20/2009 – Please do not quote without permission

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
Neuroscience Institute
Georgia State University
P.O. Box 4089
Atlanta, GA 30302-4089 USA
Email: ascarantino@gsu.edu

Abstract:

Computation and information processing are perhaps the most fundamental notions in cognitive science. They are also among the most imprecisely discussed. Many cognitive scientists take it for granted that cognition involves computation, information processing, or both. Many write as if computation and information processing were the same thing. In addition, different cognitive scientists mean different things by ‘computation’ and ‘information processing’. To help improve on this situation, which leads to a great amount of talking at cross purposes, we outline a taxonomy of notions of computation and information processing and analyze how different notions of computation relate to different notions of information processing. We argue that cognition involves information processing in several senses of the term, and as a consequence, that cognition is computation at least in a weak sense that we spell out. We hope this will bring greater clarity and precision to the debate over the nature of cognition.

Keywords: computation, information processing, computationalism, computational theory of mind, cognitivism, Shannon information, Grice, natural meaning, non-natural meaning.

1. Introduction

Computation and information processing are perhaps the most fundamental notions in cognitive science. They are also among the most imprecisely discussed. Many cognitive

¹ Thanks to our audience at the 2009 meeting of the Southern Society for Philosophy and Psychology and to Ken Aizawa, Sonya Bahar, Bob Gordon, and Oron Shagrir for comments and encouragement. Special thanks to Neal Anderson for his extensive and insightful comments.

scientists take it for granted that cognition involves computation, information processing, or both. Many write as if computation and information processing were the same thing. In addition, different cognitive scientists mean different things by ‘computation’ and ‘information processing’.

In this paper, we take steps towards a systematic investigation of the notions of computation and information processing as well as the roles they can play in a theory of cognition. Our strategy is based on the following tenets:

- There is little clarity on how notions of computation and information processing ought to be characterized and applied in a theory of cognition. Different authors employ the terms ‘computation’ and ‘information’ differently, often without clearly explaining what they mean. Some regimentation is needed.
- There are a variety of legitimate notions of computation and information. Investigators ought to be explicit as to which notion(s) they are using, in order to make their commitments explicit as well as avoid talking at cross purposes.
- Whether computation is information processing, and what kind of information processing it is, depends on which notions of computation and information are being used. In the senses that are most important for a rigorous theory of cognition, computation and information processing are independent notions: computation is not the same as information processing.
- Nevertheless, both computation and information processing may play important roles in a theory of cognition.

In what follows, we will outline a taxonomy of notions of computation and information processing and analyze how different notions of computation relate to different notions of information processing. This will allow us to shed new light on two central debates in cognitive science: the one between computationalists and anti-computationalists, and the one between classicists and connectionists. We will point out some mistakes that are made when the distinctions between varieties of computation and varieties of information are not taken into account. For instance, it is a mistake to infer that cognition involves (digital) computation simply from the premise that cognition involves information processing. We will also argue that cognition involves information processing in several senses of the term. From this, it follows that that cognition is computation at least in a weak sense that we will spell out. We hope our conclusions will serve as a clear and cogent formulation of the information processing and computational approach to cognition as well as bring greater clarity and precision to the debate over the nature of cognition.

2. Computation

Our first objective is to clarify what is at stake in two central debates in cognitive science. The first is the debate between computationalists and anti-computationalists. The second is the

debate between classicists and connectionists. These two debates are often conflated, but they are not the same. Computationalists argue that cognition is computation; anti-computationalists deny it. Classicists argue that cognition operates over language-like structures; connectionists suggest that cognition is implemented by connectionist networks instead. In both cases, the opposing camps confusedly mix terminological disputes with substantive disagreements over the nature of cognition. This is why we begin by introducing a taxonomy of notions of computation.

As we will argue in Section 2.5, depending on what one means by ‘computation’, computationalism can range from being trivially true to being an explanatorily powerful, though fallible, thesis about cognition. Similarly, we will show the opposition between classicists and connectionists to be either spurious or substantive depending on what is meant by ‘connectionism’.

Historically, the notion of computation has been characterized in many ways. For our taxonomic purposes, different notions of computation vary along two important dimensions. The first is how encompassing the notion is, that is, how many processes it includes as computational. The second dimension has to do with whether being the vehicle of a computation requires possessing meaning, or semantic properties. We will look at the first dimension first.

2.1 Digital computation

We use ‘digital computation’ for the notion implicitly defined by the classical mathematical theory of computation, and ‘digital computationalism’ for the thesis that cognition is digital computation. Rigorous mathematical work on computation began with Alan Turing and other logicians in the 1930s (Church 1936, Gödel 1934, Post 1936, Turing 1936-7), and it is now a well-established branch of mathematics (Davis et al. 1994).

A few years after Turing and others formalized the notion of digital computation, Warren McCulloch and Walter Pitts (1943) used digital computation to characterize the activities of the brain. McCulloch and Pitts were impressed that the main vehicles of neural processes appear to be trains of all-or-none spikes, which are discontinuous events (Dayan and Abbott 2001). This led them to conclude, rightly or wrongly, that neural processes are digital computations. They then used their theory that brains are digital computers to explain intelligent behavior – a phenomenon that until then had resisted mechanistic explanation.

McCulloch and Pitts’s theory is the first theory of cognition to employ Turing’s notion of digital computation. Since McCulloch and Pitts’s theory had a major influence on subsequent computational theories, we stress that digital computation is the principal notion that inspired modern computational theories of cognition (Piccinini 2004a). This makes the clarification of digital computation especially salient.

Digital computation may be defined both abstractly and concretely. Roughly speaking, abstract digital computation is the manipulation of strings of discrete elements, that is, strings of letters from a finite alphabet. Here we are interested primarily in concrete computation, or physically implemented computation. Letters from a finite alphabet may be implemented by what we call ‘digits’. To a first approximation, concrete digital computation is the processing of sequences of digits according to general rules defined over the digits (Piccinini 2007a). Let us briefly consider the main ingredients of digital computation.

The atomic vehicles of digital computation are digits, where a digit is simply a macroscopic state (of a component of the system) whose type can be reliably and unambiguously distinguished from other macroscopic types by the system. To each (macroscopic) digit type, there correspond a large number of possible microscopic states. Digital systems are engineered so as to treat all those microscopic states in the same way – the one way that corresponds to their (macroscopic) digit type. To ensure reliable manipulation of digits based on their type, a physical system must manipulate at most a finite number of digit types. For instance, ordinary computers contain only two types of digit, usually referred to as ‘0’ and ‘1’.² Digits need not mean or represent anything, but they can; numerals represent numbers, while other digits (e.g., ‘|’, ‘/’, ‘\’) do not represent anything in particular.

Digits can be concatenated (i.e., ordered) to form sequences or strings. Strings of digits are the vehicles of computations. A digital computation consists in the processing of strings of digits according to rules. A rule in the present sense is simply a map from input strings of digits, plus possibly internal states, to output strings of digits. Examples of rules that may figure in a digital computation include addition, multiplication, identity, and sorting.³

When we define concrete computations and the vehicles – such as digits – that they manipulate, we need not consider all of their specific physical properties. We may consider only the properties that are relevant to the computation, according to the rules that define the computation. Since concrete computations and their vehicles can be defined independently of the physical media that implement them, we shall call them ‘medium-independent’.

In other words, a vehicle is medium-independent just in case the rules (i.e., the input-output maps) that define a computation are sensitive only to differences between portions of the vehicles along specific dimensions of variation – they are insensitive to any other physical properties of the vehicles. Put yet another way, the rules are functions of state variables associated with a set of "information-bearing" degrees of freedom (more on this below), which can be defined differently in different physical media. Thus, a given computation can be implemented in multiple physical media (e.g., mechanical, electro-mechanical, electronic, magnetic, etc.), provided that the media possess a sufficient number of dimensions of variation (or degrees of freedom) that can be appropriately accessed and manipulated.

In the case of digits, their defining characteristic is that they are unambiguously distinguishable by the processing mechanism under normal operating conditions. Strings of digits are ordered sets of digits; i.e., digits such that the system distinguishes different members of the set depending on where they lie along the string. The rules defining digital computations are, in turn, defined in terms of strings of digits and internal states of the system, which in turn are simply states that the system can distinguish from one another. No further physical properties of a physical medium are relevant to whether they implement digital computations. Thus, digital computations can be implemented by any physical medium with the right degrees of freedom.

² The term ‘digit’ is used in two ways. It may be used for the discrete variables that can take different values. For instance, binary cells are often called bits, which can take either 0 or 1 as values. Alternatively, the term ‘digit’ may be used for the values themselves. In this second sense, it is the 0’s and 1’s that are the bits. We use ‘digit’ in the latter sense.

³ Addition and multiplication are usually defined as functions over numbers. To maintain consistency in the present context, they ought to be understood as functions over strings of digits.

To summarize, a physical system is a digital computing system just in case it is a system that manipulates input strings of digits, depending on the digits' type and their location on the string, in accordance with a rule defined over the strings (and possibly the internal states of the system).

The notion of digital computation here defined is quite general. It should not be confused with three other commonly invoked but more restrictive notions of computation: classical computation, algorithmic computation, and computation of Turing-computable functions (Figure 1).

Let us begin with the most restrictive notion of the three: classical computation. A classical computation is a digital computation that has two additional features. First, it manipulates a special kind of digital vehicle: sentence-like symbolic structures. Second, it is algorithmic: it proceeds in accordance with algorithms – i.e., effective, step-by-step procedures that manipulate strings of digits and produce a result within finitely many steps. Thus, a classical computation is a digital, algorithmic computation whose algorithms are sensitive to the combinatorial syntax of the symbols (Fodor and Pylyshyn 1988).⁴

A classical computation is algorithmic, but the notion of algorithmic computation – i.e., digital computation that follows an algorithm – is more inclusive, because it does not require that the vehicles being manipulated be sentence-like.

An algorithmic computation, in turn, is computable by Turing machines (Turing-computable). This is known as the Church-Turing thesis, a thesis for which there is compelling evidence (Kleene 1952). But the computation of Turing-computable functions need not be carried out by following an algorithm. For instance, many neural networks compute Turing-computable functions, but such networks have no level of functional organization at which they follow the steps of an algorithm for computing their functions; no functional level at which their states and state transitions are discrete.⁵

Finally, the computation of a Turing-computable function is a digital computation, because Turing-computable functions are by definition functions of a denumerable domain – a domain whose elements may be counted – and the arguments and values of such functions are, or may be represented by, strings of digits. But it is equally possible to define functions of strings of digits that are not Turing-computable, and to mathematically define processes that compute such functions. Some authors have speculated that some functions that are not Turing-computable may be computable by some physical systems (Penrose 1994, Copeland 2000). According to our usage, any such computations still count as digital computations. Of course, it may well be that only the Turing-computable functions are computable by physical

⁴ Fodor and Pylyshyn (1988) restrict their notion of classical computation to processes defined over *representations*. In other words, the sentence-like symbolic structures manipulated by Fodor and Pylyshyn's classical computations must have semantic content. Thus, Fodor and Pylyshyn's classical computations are a kind of 'semantic computation'. We will discuss the distinction between semantic and non-semantic computation in Section 2.1.

⁵ Some authors use the term 'algorithm' for the computations performed by such networks too. In this broader sense of 'algorithm', any processing of a signal follows an algorithm, regardless of whether the process is defined in terms of discrete manipulations of strings of digits. Since this is a more encompassing notion of algorithm than the one employed in computer science – indeed, it is even broader than the notion of computing Turing-computable functions – our point stands. The point is that the computation of Turing-computable functions is a more inclusive notion than that of algorithmic computation in the standard sense (Piccinini 2008a).

systems; whether this is the case is an empirical question that does not affect our discussion. Furthermore, the computation of Turing-*un*computable functions is unlikely to be relevant to the study of cognition. Be that as it may – we will continue to talk about digital computation in general.

Many other distinctions may be drawn within digital computation, such as hardwired vs. programmable, special purpose vs. general purpose, and serial vs. parallel computation (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, such as the computations performed by standard digital computers, as well as many types of neural network computations. Since digital computation is the notion that inspired the computational theory of cognition, it is the most relevant notion for present purposes.

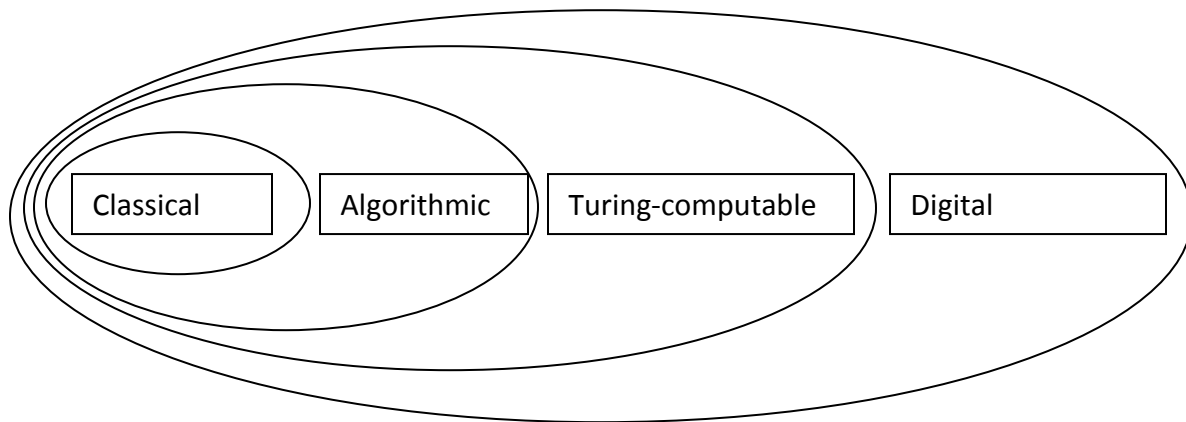


Figure 1. Types of digital computation and their relations of class inclusion.

2.2 Generic Computation

Even though the notion of digital computation is broad, we need an even broader notion of computation. In order to capture all relevant uses of ‘computation’ in cognitive science, we introduce the notion of ‘generic computation’. The banner of ‘generic computation’ will subsume digital computation as defined in the previous section, analog computation, and neural computation.

We use ‘generic computation’ to designate the processing of vehicles according to rules that are sensitive to certain vehicle properties, and specifically, to differences between different portions of the vehicles. This definition generalizes the definition of digital computation by allowing for a broader range of vehicles (e.g., continuous variables as well as discrete digits). We use ‘generic computationalism’ to designate the thesis that cognition is computation in the generic sense.

Since the definition of generic computation makes no reference to specific media, generic computation is medium-independent – it applies to all media. That is, the differences between portions of vehicles that define a generic computation do not depend on specific physical properties of the medium but only on the presence of relevant degrees of freedom.

Shortly after McCulloch and Pitts argued that brains perform digital computations, others countered that neural processes may be more similar to *analog* computations (Gerard 1951, see also Rubel 1985). The evidence for the analog computational theory included neurotransmitters and hormones, which are released by neurons in degrees rather than in all-or-none fashion.

Analog computation is often contrasted with digital computation, but analog computation is a vague and slippery concept. The clearest notion of analog computation is that of Pour-El (1974). Roughly, abstract analog computers are systems that manipulate continuous variables to solve certain systems of differential equations. Continuous variables are variables that can vary continuously over time and take any real values within certain intervals.

Analog computers can be physically implemented, and physically implemented continuous variables are different kinds of vehicles than strings of digits. While a digital computing system can always unambiguously distinguish digits and their types from one another, a concrete analog computing system cannot do the same thing with the exact values of continuous variables. This is because the values of continuous variables can only be measured within a margin of error. Primarily due of this, analog computations (in the present, strict sense) are a different kind of process than digital computations.⁶

The claim that the brain is an analog computer is ambiguous between two interpretations. On the literal interpretation, ‘analog computer’ is given Pour-El’s (1974) precise meaning. The theory that the brain is literally an analog computer in this literal sense was never very popular. The primary reason is that although spike trains are a continuous function of time, they are also sequences of all-or-none signals (Dayan and Abbott 2001).

On a looser interpretation, ‘analog computer’ refers to a broader class of computing systems. For instance, Churchland and Sejnowski use the term ‘analog’ so that containing continuous variables is sufficient for a system to count as analog: “The input to a neuron is analog (continuous values between 0 and 1)” (1992, p. 51). Under such a usage, even a slide rule counts as an analog computer. Sometimes, the notion of analog computation is simply left undefined, with the result that by ‘analog computer’ one refers to some kind of presumably non-digital but otherwise unspecified computing system.

The looser interpretation of ‘analog computer’ – of which Churchland and Sejnowski’s usage is one example – is not uncommon, but we find it misleading because prone to being confused with the literal interpretation of analog computation *sensu* Pour-El’s (1974). Let us now turn to the notion of ‘neural computation’.

In recent decades, the analogy between brains and computers has taken hold in neuroscience. Many neuroscientists have started using the term ‘computation’ for the processing of neuronal spike trains, that is, sequences of spikes produced by neurons in real time. The processing of neuronal spike trains by neural systems is often called ‘neural computation’. Whether neural computation is best regarded as a form of digital computation, analog computation, or something else is a difficult question, which we cannot settle here. Instead, we will subsume digital computation, analog computation, and neural computation under the banner of ‘generic computation’ (Figure 2).

⁶ For more details on the contrast between digital and analog computers, see Piccinini 2008b, Section 3.5.

While digits are unambiguously distinguishable vehicles, other vehicles are not so. For instance, concrete analog computers cannot unambiguously distinguish between any two portions of the continuous variables they manipulate. Since the variables can take any real values but there is a lower bound to the sensitivity of any system, it is always possible that the difference between two portions of a continuous variable is small enough to go undetected by the system. From this it follows that the vehicles of analog computation are not strings of digits. Nevertheless, analog computations are only sensitive to the differences between portions of the variables being manipulated, to the degree that they can be distinguished by the system. Any further physical properties of the media implementing the variables are irrelevant to the computation. Like digital computers, therefore, analog computers operate on medium-independent vehicles.

Finally, current evidence suggests that the vehicles of neural processes are neuronal spikes and that the functionally relevant aspects of neural processes are medium-independent aspects of the spikes—primarily, spike rates. That is, the functionally relevant aspects of spikes may be implemented either by neural tissue or by some other physical medium, such as a silicon-based circuit. Thus, spike trains appear to be another case of medium-independent vehicle, in which case they qualify as proper vehicles for generic computations. Assuming that brains process spike trains and that spikes are medium-independent vehicles, it follows by definition that brains perform computations in the generic sense.

In conclusion, generic computation includes digital computation, analog computation, neural computation (which may or may not correspond closely to digital or analog computation), and more.

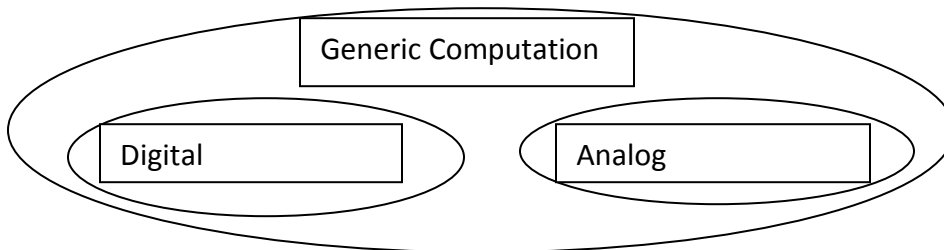


Figure 2. Types of generic computation and their relations of class inclusion. Neural computation is not represented because where it belongs is controversial.

2.3 Pancomputationalism

There are other notions of computation, even broader than generic computation. According to some authors, every physical process is a computation. This may be because at the most fundamental physical level, everything is a collection of discrete states, and every physical process is a digital computation defined over such discrete states (e.g., Wolfram 2002). The latter is a working hypothesis or article of faith for those interested in seeing how familiar physical laws might emerge from a discrete substrate. It is not a widely accepted notion, and there is no direct evidence for it.

This notion of computation is not directly relevant to theories of cognition, because theories of cognition attempt 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 a notion of computation that excludes at least some other processes as non-computational (cf. Piccinini 2007b).

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

We agree that, *if* brains perform computations, *then* the job of neuroscience and psychology is to discover which specific computations they 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 neuroscience and psychology as an empirical hypothesis, to explain intelligent behavior by analogy with digital computers. Most of the empirical import of the computational theory of cognition is in danger of being eliminated by stretching the notion of computation from digital to generic. Stretching the notion of computation even further, so as to embrace pancomputationalism, 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 debate. It is utterly implausible that critics of computationalism have simply failed to notice that every process is a computation. More likely, they object to what they perceive to be questionable empirical commitments of computationalism. For this reason, computation in the sense of pancomputationalism is a poor foundation for a theory of cognition. From now on, we will set pancomputationalism aside.

2.4 Semantic vs. Non-Semantic Computation

So far, we have taxonomized different notions of computation according to how broad they are, namely, how many processes they include as computational. 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.

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.

To maintain generality, we will consider both semantic and non-semantic notions of computation. In a nutshell, semantic notions of computation define computations as operating over representations. By contrast, non-semantic notions of computation define computations without requiring that the vehicles being manipulated be representations.

Let us take stock. We have introduced the following taxonomy of varieties of computation (Figure 3): digital computation (computation according to computability theory), generic computation (which includes digital computation as a special case), and computation in the sense of pancomputationalism (which is irrelevant to a theory of cognition). Within each category, we may distinguish a semantic notion of computation, which presupposes that the computational vehicles are representations, and a non-semantic one, which doesn't.⁷

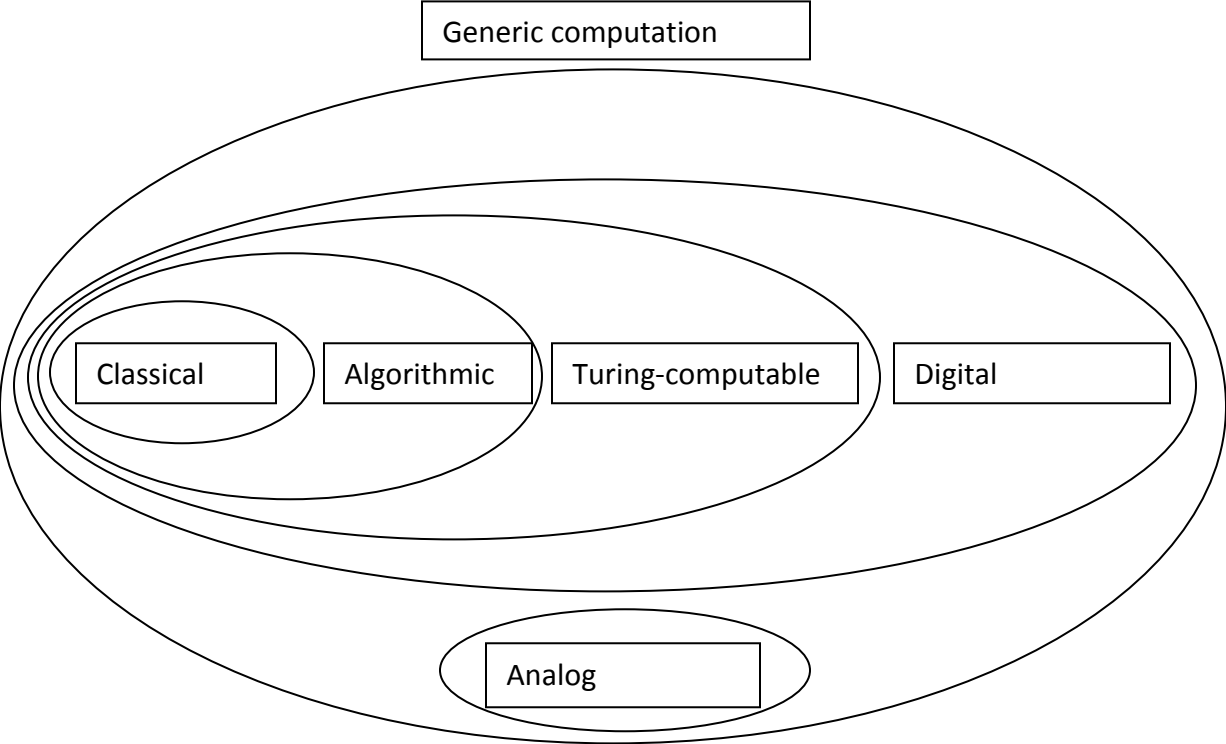


Figure 3. Types of computation and their relations of class inclusion. Once again, neural computation is not represented because where it belongs is controversial.

2.5 Computationalism, Classicism, and Connectionism Reinterpreted

The distinctions introduced so far allow us to shed new light on the longstanding debates on computationalism, classicism, and connectionism. Computationalism, we have said, is the view

⁷ One terminological caveat. Later on, we will also speak of *semantic* notions of information. In the case of non-natural semantic information, by ‘semantic’ we will mean the same as what we mean here, i.e., representational. A representation is something that can mis-represent, i.e., may be unsatisfied or false. In the case of natural information, by ‘semantic’ we will mean something weaker, which is not representational because it cannot mis-represent (see below).

that cognition is computation. Leaving aside pancomputationalism, we can now appreciate that this may mean at least two different things: that cognition is digital computation, and that cognition is computation in the generic sense. Let us consider them in reverse order.

If 'computation' means generic computation, i.e., the processing of medium-independent vehicles according to rules, then computationalism becomes the claim that cognition is the manipulation of medium-independent vehicles according to rules. This is not a trivial claim. Disputing 'generic computationalism' requires arguing that cognitive capacities depend in an essential way on some specific physical property of cognitive systems other than those required to distinguish computational vehicles from one another. For instance, one may argue that the processing of spike rates by neural systems is essentially dependent on the use of voltages instead of other physical properties to generate action potentials. Without those specific physical properties, one may propose, it would be impossible to exhibit certain cognitive capacities.

Something along these lines is sometimes maintained for phenomenal consciousness, due to the qualitative feel of conscious states (e.g., Block 1978). Nevertheless, many cognitive scientists maintain that consciousness is reducible to computation and information processing (cf. Baars, Banks, and Newman 2003). Since we lack the space to discuss consciousness, we will assume (along with mainstream cognitive scientists) that if and insofar as consciousness is not reducible to computation in the generic sense and information processing, cognition may be studied independently of consciousness. If this is right, then, given that cognitive processes appear to be implemented by spike trains and that spike trains appear to be medium-independent, we may conclude that cognition is computation in the generic sense.

If 'computation' means digital computation, i.e., the processing of strings of digits according to rules, then computationalism says that cognition is the manipulation of digits. As we saw above, this is what McCulloch and Pitts (1943) argued for. Importantly, it's also what both classical computational theories and many connectionist theories of cognition maintain. Finally, this view is what most critics of computationalism object to (Globus 1992, Perkel 1990). Although digital computationalism encompasses a broad family of theories, it is a powerful thesis with considerable explanatory scope. Its exact explanatory scope depends on the precise kind of digital computing systems that cognitive systems are postulated to be. If cognitive systems are digital computing systems, they may be able to compute a smaller or larger range of functions depending on their mechanistic properties.⁸

Let us now turn to another debate often confusedly mixed with the one just described, namely the debate between classicists and connectionists. By the 1970s, McCulloch and Pitts were mostly forgotten. The dominant paradigm in cognitive science was classical or symbolic AI, aimed at writing computer programs that simulate intelligent behavior without much concern for how brains work (Newell and Simon 1976). It was commonly assumed that digital computationalism is committed to classicism, that is, the idea that cognition is the manipulation of linguistic, or sentence-like, structures. On this view, cognition consists of performing

⁸ Digital computing systems belong in a hierarchy of systems that are computationally more or less powerful. The hierarchy is measured by the progressively larger classes of functions each class of system can compute. Classes include the functions computable by finite automata, pushdown automata, and Turing machines (Davis et al., 1994).

computations on sentences with a logico-syntactic structure akin to that of natural languages, but written in the language of thought (Harman 1973, Fodor 1975).

It was also assumed that, given digital computationalism, explaining cognition is an autonomous activity from explaining the brain, in the sense that figuring out how the brain works tells us little or nothing about how cognition works: neural descriptions and computational descriptions are at “different levels” (Fodor 1975, Marr 1982).

During the 1980s, connectionism re-emerged as an influential approach to psychology. Broadly understood, connectionism is the thesis that cognition is mediated by neural networks. Most connectionists deny that cognition is based on a language of thought, and affirm that a theory of cognition should be at least “inspired” by the way the brain works (Rumelhart and McClelland 1986).

The resulting debate between classicists and connectionists (e.g., Fodor and Pylyshyn 1988, Smolensky 1988) has been somewhat confusing. Part of what’s confusing about it is that different authors employ different notions of computation, which vary in both their degree of precision and their inclusiveness (see Piccinini 2008a for more details). But even after we factor out differences in notions of computation, further confusions lie in the wings.

Digital computationalism and connectionism are often described as being at odds with one another. This is because it is assumed that digital computationalism is committed to classical computation (the idea that the vehicles of digital computation are language-like structures) and to autonomy from neural processes, two theses flatly denied by many prominent connectionists. But many connectionists also maintain that neural networks perform computations defined over strings of digits and that such computations explain behavior, so they should be counted among the digital computationalists (Bechtel & Abrahamsen, 2002; Churchland & Sejnowski, 1992; Feldman & Ballard, 1982; Hopfield, 1982; Koch, 1999; Marr & Poggio, 1976; Rumelhart & McClelland, 1986; Smolensky & Legendre, 2006). To make matters worse, other connectionists reject digital computationalism—they maintain that their neural networks, while explaining behavior, do not perform digital computations (Edelman, 1992; Freeman, 2001; Globus, 1992; Horgan & Tienson, 1996; Perkel, 1990).

In addition, the very origin of digital computationalism calls into question the commitment to autonomy from neural processes. McCulloch and Pitts initially introduced digital computationalism as a theory of the brain, and some form of computationalism or other is now a working assumption of many neuroscientists. Furthermore, nothing in the notion of digital computation commits us to conceiving of its vehicles as language-like.

To clarify this debate, we need two separate distinctions. One is the distinction between digital computationalism (“cognition is digital computation”) and its denial (“cognition is something other than digital computation”). The other is the distinction between classicism (“cognition is digital computation over linguistic structures”) and connectionism (“cognition is computation – digital or not – by means of neural networks”).

We then have two versions of digital computationalism—the classical one (“cognition is digital computation over linguistic structures”) and the connectionist one (“cognition is digital computation by neural networks”)—standing opposite to the denial of digital computationalism. This includes the generic computationalist version of connectionism (“cognition is a kind of neural network processing that is not digital computation”). This may be

enough to accommodate most views in the current debate. But it still doesn't do justice to the relationship between classicism and connectionism.

A further wrinkle in this debate derives from the ambiguity of the term 'connectionism'. In its original sense, connectionism says that behavior is explained by the changing 'connections' between stimuli and responses, which are biologically mediated by changing connections between neurons (Thorndike 1932, Hebb 1949). This original connectionism is related to behaviorist associationism, according to which behavior is explained by the association between stimuli and responses. Associationist connectionism adds to associationism a biological mechanism to explain the associations: the mechanism of changing connections between neurons.

Contemporary connectionism is much more general than associationist connectionism. In its most general form, contemporary connectionism simply says that behavior is explained (at some level) by neural network activity. But this is a truism—or at least it should be. The brain is the organ of cognition, the cells that perform cognitive functions are (mostly) neurons, and neurons perform their cognitive labor by organizing themselves in networks. Modern connectionism in this sense is little more than a platitude.

The confusion arises, in part, because many contemporary connectionists are also connectionists in the original, associationist sense. So classicists object to connectionism in the associationist sense, while connectionists insist that the brain is made out of neural networks. Of course it is, but this does not refute classicism. What remains to be determined is which neural networks, organized in what way, actually explain cognition.

To sum up, everyone is (or should be) a connectionist in the most general contemporary sense, though not everyone is a connectionist in the associationist sense. Some people are classicists, believing that in order to explain cognition, neural networks must amount to manipulators of linguistic structures. Some people are non-classicist (but still computationalist) connectionists, believing that cognition is explained by non-classical neural network digital computation. Finally, some people are anti-computationalist connectionists, believing that cognition is explained by neural network processes, but these do not amount to digital computation (e.g., because they process the wrong kind of vehicles). To find out which view is correct, in the long run the only effective way is to study nervous systems at all levels of organization and find out how they produce behavior (Figure 4; see also Piccinini 2008c).

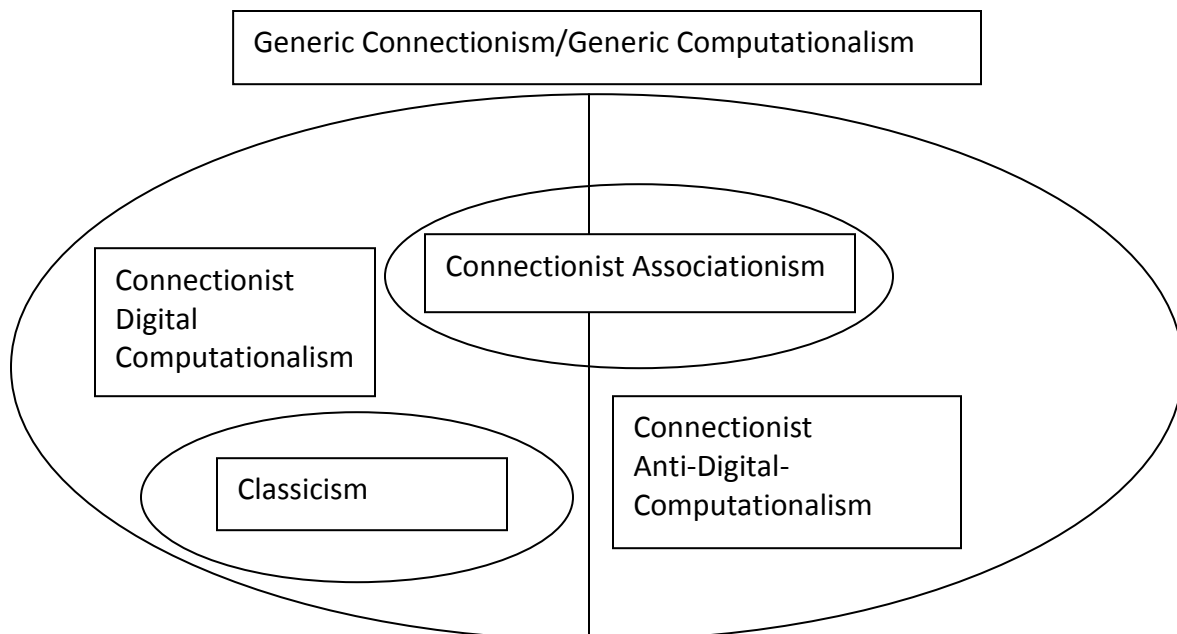


Figure 4. Some prominent forms of computationalism and their relations.

3. Information

We now turn our attention to the notion of information. Computation and information processing are commonly assimilated in the sciences of mind. As we have seen, computation is a mongrel concept that comprises importantly different notions. The same is true of information. Once the distinctions between notions of information are introduced, we will be in a position to consider whether the standard assimilation of information processing and computation is warranted and useful for a theory of cognition.

Information plays a central role in many disciplines. In the sciences of mind, information is invoked to explain cognition and behaviour (e.g. Miller 1951, Minsky 1968). In communication engineering, information is central to the design of efficient communication systems such as television, radio, telephone networks, and the internet (e.g. Shannon 1948, Fano 1961, Pierce 1980). A number of biologists have suggested explaining genetic inheritance in terms of the information carried by DNA sequences (e.g. Maynard Smith 2000; see also Godfrey-Smith 2000, Griffiths 2001). Animal communication theorists routinely characterize non-linguistic communication in terms of shared information (Bradbury and Vehrencamp 1998). Some philosophers maintain that information explains the intentionality of mental states, namely their being about states of affairs (Dretske 1981, Millikan 2004). Finally, information plays important roles in several other disciplines, such as computer science, physics, and statistics.

To account for the different roles information plays in all these fields, more than one notion of information is required. In this paper, we distinguish between three main notions of information.

3.1 Shannon information

We use 'Shannon information' to designate the notion of information initially formalized by Shannon (1948).⁹ We distinguish between information theory, which provides the formal definition of information, and communication theory, which applies the formal notion of information to engineering problems of communication. We begin by introducing two measures of information from Shannon's information theory. Later we will discuss how to apply them to the physical world.

Let X and Y be two discrete random variables taking values in $A_X = \{a_1, \dots, a_n\}$ and $A_Y = \{b_1, \dots, b_r\}$ with probabilities $p(a_1), \dots, p(a_n)$ and $p(b_1), \dots, p(b_r)$. We assume that $p(a_i) > 0$ for all $i=1, \dots, n$, $p(b_j) > 0$ for all $j=1, \dots, r$, $\sum_{i=1}^n p(a_i) = 1$ and $\sum_{j=1}^r p(b_j) = 1$.

Shannon's key measures of information are the following:

⁹ Shannon was building on important work by Boltzmann, Szilard, Hartley, and others (Pierce 1980).

$$H(X) = - \sum_{i=1}^n p(a_i) \log_2 p(a_i)^{10}$$

$$I(X, Y) = \sum_{i=1, j=1}^{n, r} p(a_i, b_j) \log_2 \frac{p(a_i | b_j)}{p(a_i)}$$

$H(X)$ is called ‘entropy’; it measures the *average information* produced by the selection of values in $A_X = \{a_1, \dots, a_n\}$. We can think of $H(X)$ as a weighted sum of n expressions of the form $I(a_i) = -\log_2 p(a_i)$. $I(a_i)$ is sometimes called the ‘self-information’ produced when X takes the value a_i . Shannon obtained the formula for *entropy* by setting a number of mathematical desiderata that any satisfactory measure of uncertainty should satisfy, and showing that the desiderata could only be satisfied by the formula given above.¹¹

Shannon information may be measured with different units. The most common unit is the ‘bit’.¹² A bit is the information generated by a variable, such as X , taking a value a_i that has a 50% probability of occurring. Any outcome a_i with less than 50% probability will generate *more* than 1 bit; any outcome a_i with more than 50% probability will generate *less* than 1 bit. The choice of unit corresponds to the base of the logarithm in the definition of entropy. Thus, the information generated by X taking value a_i is equal to 1 bit when $I(a_i) = -\log_2 p_i = 1$; that is, when $p_i = 0.5$.¹³

Entropy has a number of important properties. First, it is equal to zero when X takes some value a_i with probability 1. This tells us that information as entropy presupposes uncertainty: if there is no uncertainty as to which value a variable X takes, the selection of that value generates no Shannon information.

Second, the closer the probabilities p_1, \dots, p_n are to having the same value, the higher the entropy. The more uncertainty there is as to which value will be selected, the more the selection of a specific value generates information.

Third, the entropy is highest and equal to $\log_2 n$ when X takes every value with the same probability. This occurs when the “freedom of choice” of the selection process represented by X taking values in $\{a_1, \dots, a_n\}$ is maximal.

¹⁰ 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, 379) 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.

¹¹ The three mathematical desiderata are the following: (i) The entropy H should be continuous in the probabilities p_i , (ii) The entropy H should be a monotonic increasing function of n when $p_i = 1/n$, and (iii) If $n = b_1 + \dots + b_k$ with b_i

positive integer, then $H(1/n, \dots, 1/n) = H(b_1/n, \dots, b_k/n) + \sum_{i=1}^k b_i/n H(1/b_i, \dots, 1/b_i)$. Shannon further supported his

interpretation of H as the proper measure of information by demonstrating that the channel capacity required for most efficient coding is determined by the entropy (see Theorem 9, in Section 9 of Shannon 1948).

¹² Shannon credited John Tukey, a computer scientist at Bell Telephone Laboratories, with introducing the term in a 1947 working paper.

¹³ $-\log_2 0.5 = 1$.

$I(X,Y)$ is called ‘mutual information’; it is 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 $p(a_i,b_j)=p(a_i)p(b_j)$ for all i and j . This is to say that the transfer of mutual information between two sets of uncertain outcomes $A_X=\{a_1,\dots,a_n\}$ and $A_Y=\{b_1,\dots,b_r\}$ amounts to the *statistical dependency* between the occurrence of outcomes in A_X and A_Y . Information is ‘mutual’ because statistical dependencies are symmetrical.

Shannon’s measures of information have many applications. The best-known application is Shannon’s own: communication theory. Shannon was looking for an optimal solution to what he called the ‘fundamental problem of communication’ (Shannon 1948, 379), that is, the reproduction of messages from an information source to a destination. In Shannon’s sense, any device that produces messages in a stochastic manner can count as an information source or destination.

Shannon distinguished three phases in the communication of messages: (a) the *encoding* of the messages produced by an *information source* into channel inputs, (b) the *transmission* of signals across a channel with random disturbances generated by a *noise source*, (c) the *decoding* of channel outputs back into the original messages at the *destination*.

The process is summarized by the following picture of a *communication system*:

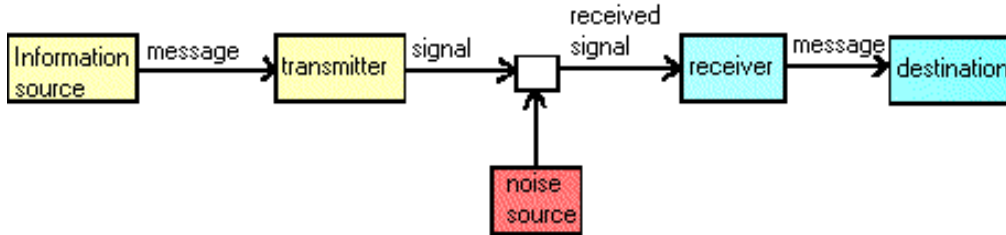


Figure 5: Shannon’s (1948) depiction of a *communication system*

Having distinguished the operations of *coding*, *transmission* and *decoding*, Shannon developed mathematical techniques to study how to perform them efficiently. Shannon’s primary tool was the concept of *information*, as measured by entropy and mutual information.

A common misunderstanding of Shannon information is that it is a semantic notion. This misunderstanding is promoted by the fact that Shannon liberally referred to *messages* as carriers of Shannon information. But Shannon’s notion of message is not the usual one. In the standard sense, a message has semantic content or meaning – there is something it stands for. By contrast, Shannon’s messages need not have semantic content at all – they need not stand for anything.¹⁴

We will continue to discuss Shannon’s theory using the term ‘message’, but the reader should keep in mind that the usual commitments associated with the notion of message do not apply. The identity of a communication-theoretic message is fully described by two features: it is a physical structure distinguishable from a set of alternative physical structures, and it

¹⁴ Furthermore, Shannon’s messages don’t even have to be strings of digits of finitely many types; on the contrary, they may be continuous variables. We gave the definitions of entropy and mutual information for discrete variables, but the definitions can be modified to suit continuous variables (see Shannon 1948, Part III).

belongs to an exhaustive set of mutually exclusive physical structures selectable with well-defined probabilities.

Under these premises, to *communicate* a message produced at a source amounts to generating a second message at a destination that replicates the original message so as to satisfy a number of desiderata (accuracy, speed, cost, etc.). The sense in which the semantic aspects are irrelevant is precisely that messages carry information merely *qua* selectable, physically distinguishable structures associated with given probabilities. A seemingly paradoxical corollary of this notion of communication is that a nonsense message such as ‘2#r %h@’ could in principle generate *more* Shannon information than a meaningful message such as ‘avocado’.

To understand why, consider an experiment described by a random variable X taking values a_1 and a_2 with 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.

Shannon’s insight was that the occurrence of a_2 generates *more information* than the occurrence of a_1 , because it is *less expectable*, or *more surprising*, in light of the prior probability distribution. This is reflected by the measure of self-information we introduced earlier: $I(a_i)=-\log_2 p(a_i)$. Given this measure, the higher (lower) is $p(a_i)$, the lower (higher) is $I(a_i)=-\log_2 p(a_i)$. Therefore, if ‘2#r %h@’ is a more surprising message than ‘avocado’, it will carry more information, despite its meaninglessness.

Armed with his measures of non-semantic information, Shannon was able to prove a number of seminal theorems, which had a profound impact on the field of communication engineering. In the “fundamental theorem for a noisy channel”, Shannon proved that it was theoretically possible to make the error rate in the transmission of information across a randomly disturbed channel as low as desired up until the point in which the source information rate in bits per unit of time becomes larger than the channel capacity, which is defined as the mutual information maximized over all input source probability distributions.¹⁵ Notably, this is true regardless of the physical nature of the channel (e.g., true for electrical channels, wireless channels, water channels, etc.).

Communication theory has found many applications, most prominently in communication engineering, computer science, neuroscience, and psychology. For instance, Shannon information is commonly used by neuroscientists to measure the quantity of information carried by neural signals about a stimulus and estimate the efficiency of coding (what forms of neural responses are optimal for carrying information about stimuli) (Dayan and Abbott 2001, chap. 4). Now, we move on from Shannon information, which is non-semantic, to semantic information.

3.2 Semantic Information

Suppose that a certain process leads to the selection of signals. Think, for instance, of how you produce letters of the alphabet when speaking. There are two dimensions of this process that

¹⁵ Shannon defined the channel capacity C as follows: $Max_{P(a)} I(X,Y) = Max_{P(a)} [H(X) - H(Y|X)]$.

are relevant for the purposes of information transmission. One is the uncertainty that characterizes the speaker's letter selection process as a whole – the probability that each letter be selected. This is the non-semantic dimension that would interest a communication engineer and it is captured by Shannon's theory. A second dimension of the letter selection process concerns what the selection of each letter 'means' (in the context of a word). This would be the dimension of interest, for example, to a linguist trying to decipher a tribal language.

Broadly understood, semantic notions of information aim to determine what the selection of a *particular signal* means (or stands for). To address the semantics of the signal, it is no longer enough to know which other signals might have been selected instead and with what probabilities. Whereas the selection of, say, any of forty equiprobable letters will generate the same non-semantic information, the selection of each individual letter will generate different semantic information depending on what that particular letter stands for.

Semantic and non-semantic notions of information are both connected with the reduction of uncertainty. In the case of non-semantic information, the uncertainty has to do with which among many signals will be selected. In the case of semantic information, the uncertainty has to do with what a particular signal means or stands for.

One of the foundational problems for the sciences of mind is that no uncontroversial theory of semantic information – comparable in rigor and scope to Shannon's theory – has emerged. The notion of semantic information is still to a large extent in search of a theory. This explains the temptation felt in many quarters to appropriate Shannon's theory of non-semantic information and squeeze into it a semantic story. The temptation is in effect to understand information transmission in terms of reduction of uncertainty, use Shannon's measures to quantify the information transmitted, and draw conclusions about the meanings of the transmitted signals.

This will not do. As we have stressed, Shannon information does not capture, nor is it intended to capture, the semantic content, or meaning, of signals. From the fact that Shannon information has been transmitted no conclusions follow concerning what semantic information, if any, has been transmitted.

To tackle the 'meaning' relation, we begin by distinguishing between two kinds of meaning, namely, natural and non-natural meaning (Grice 1957). Natural meaning is exemplified by a sentence such as 'those spots mean measles', which is true – Grice claimed – just in case the patient has measles. Non-natural meaning is exemplified by a sentence such as 'those three rings on the bell (of the bus) mean that the bus is full' (ibid., 85), which is true even if the bus is not full.

We will extend the distinction between Grice's two types of *meaning* to a distinction between two types of *semantic information*: *natural information* and *non-natural information*.¹⁶ Spots carry natural information about (or mean, or stand for) measles by virtue of a reliable physical correlation between measles and spots. By contrast, the three rings on

¹⁶ By calling this kind of information non-natural, we are not taking a stance on whether it can be naturalized, that is, reduced to some more fundamental physical process. We are simply using Grice's terminology to distinguish between two importantly different, and often conflated, notions of semantic information. We will briefly discuss the naturalization of non-natural information in Section 4.3.

the bell of the bus carry non-natural information about (or mean, or stand for) the bus being full by virtue of a convention.¹⁷

We will now consider each notion of semantic information in more detail.

3.2.1 Natural (Semantic) Information

When smoke carries information about fire, the basis for this informational link is the physical relation between fire and smoke. By the same token, when spots carry information about measles, the basis for this informational link is the physical relation between measles and spots. Both are examples of natural semantic information. The first task of an account of natural semantic information is to specify the physical relation that has to obtain between a source and a signal for the signal to carry natural information about the source.

Following Dretske (1981), we discuss natural information in terms of correlations between events, understood as particulars having properties (at a time). Roughly speaking, a signal carries natural information about anything it reliably correlates with. Reliable correlations are the sorts of correlations information users can count on to hold in some range of future and counterfactual circumstances. For instance, smoke reliably correlates with fire, spots reliably correlate with measles, and ringing doorbells reliably correlate with people at the door. It is by virtue of these correlations that one can dependably infer fire from smoke, measles from spots, and visitors from ringing doorbells.

Yet correlations are rarely perfect. Smoke is occasionally produced by smoke machines, spots are occasionally produced by mumps, and doorbell rings are occasionally produced by naughty kids who immediately run away. The natural information carried by signals in all such cases is *probabilistic*: a signal may carry natural information to the effect that *o* is *probably* G, whether or not *o* is actually G.¹⁸

This probabilistic notion of natural information makes sense of the central role information plays in the descriptive and explanatory efforts of scientists of mind. Organisms survive and reproduce by tuning themselves to reliable but imperfect correlations between internal variables and environmental stimuli, and between environmental stimuli and threats and opportunities. In comes the sound of a predator, out comes running. In comes the redness of a ripe apple, out comes approaching.

Consider a paradigmatic example from ethology. Vervet monkeys' alarm calls appear to be qualitatively different for leopards, eagles, and snakes (Struhsaker 1967, Seyfarth & Cheney 1992). As a result, vervet calls respond to eagle calls by hiding behind bushes, to leopard calls by climbing trees, and to snake calls by standing tall.

A theory of natural information with explanatory power should be able to make sense of the connection between the reception of different calls and different behaviors. To make sense of this connection, we need a notion of probabilistic information according to which vervet

¹⁷ This is not to say that conventions are the only possible source of non-natural meaning. For further discussion, see Grice 1957.

¹⁸ For a more precise and detailed theory of probabilistic information, see Scarantino, ms. The probabilistic notion of information we employ includes all-or-nothing natural information – roughly, the natural information that *o* is G with certainty – as a special, limiting case. Different theories of all-or-nothing natural information have been offered by Dretske 1981 and Cohen and Meskin 2006. For a critique of approaches focusing solely on all-or-nothing natural information, see Scarantino (2008).

monkey calls carry natural information about predators insofar as call types reliably correlate with predator types. For example, if the probability that an eagle is present is high given an eagle alarm call, but low in the absence of an eagle alarm call, eagle alarm calls will carry a lot of natural probabilistic information about eagles (the same holds, *mutatis mutandis*, for snake and leopard calls). This allows us to give an information-based explanation of why vervets take refuge under bushes when they hear eagle calls. They do because they have received probabilistic information that an eagle is present.

More generally, the sorts of correlations on which behaviors are based commonly include some uncorrelated token events along the way – the more uncorrelated token events there are, the weaker the correlation. Organisms do make mistakes, after all. Some mistakes are due precisely to the reception of probabilistic information about events that fail to obtain. For instance, sometimes non-predators are mistaken for predators because they sound like predators; or predators are mistaken for non-predators because they look like non-predators.

The key point is that an event's failure to obtain is compatible with the reception of natural information about its obtaining, just like the claim that the probability that *o* is *G* is high is compatible with the claim that *o* is not *G*. No valid inference rules take us from claims about the transmission of probabilistic information to claims about how things turn out to be.¹⁹

3.2.2 Non-natural (Semantic) Information

Bearers of natural information, we have said, “stand for” what they are about simply in virtue of being physically connected to it. We have provided a working account of this physical connection: there must be a reliable correlation between a signal and its source. Whatever the right account may be, one thing is clear: in the absence of the right physical connection, no natural information is carried.

Bearers of non-natural information, by contrast, need not be physically connected to what they are about in any direct way. Thus, there must be an alternative process by which bearers of non-natural information come to bear the non-natural information they bear. A convention, as in the case of the three rings on the bell of the bus, is a clear example of what may establish a non-natural informational link.

The convention may either be explicitly stipulated, as in the rings case, or emerge spontaneously, as in the case of the non-natural information attached to words in natural languages. But we do not wish to assume that non-natural information is always based on convention (cf. Grice 1957). Perhaps there are other mechanisms, such as learning or biological evolution, through which non-natural informational links may be established. What matters for something to bear non-natural information is that, somehow, it becomes associated with what it “stands for”. Later we will discuss what this association amounts to.

Since a signal that carries non-natural information is arbitrary – it need not be produced by the event it “stands for” – the signal may be produced whether or not the event occurred.

¹⁹ In Scarantino and Piccinini (Forthcoming), we have argued that an important implication of this point is that the transmission of natural information entails nothing more than the truth of a probabilistic claim. It follows that our distinction between natural and non-natural meaning differs from Grice's in one important respect. On our view, there is nothing objectionable in holding that ‘those spots carry natural information about measles, but he doesn't have measles’, provided measles are more likely given those spots than in the absence of those spots.

In these situations, we commonly 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 no ace, you might say that you have received false information. There are two principal strategies to make sense of this way of talking.

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

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

We choose this second option, because we are interested in a taxonomy of actual uses of the information concept in the sciences of mind (Scarantino and Piccinini, forthcoming). Cognitive scientists routinely say that cognition involves the processing of information. Sometimes they mean that cognition involves the processing of natural information. At other times, they mean that cognition involves the processing of non-natural information, without any commitment as to whether the information is true or false.

The notion of non-natural information used in cognitive science is best understood as the notion of *representation*, where a representation is by definition something that can get things wrong. The statement ‘water is not H₂O’ gets things wrong with respect to the chemical composition of water, but it does not fail to represent that water is not H₂O. By the same token, the statement ‘water is not H₂O’ contains false non-natural information to the effect that water is not H₂O.

One of the central challenges for a theory of cognition and behavior is to explain how a system can acquire the ability to represent, or non-naturally mean, or carry non-natural information, that p whether or not p is true. There is a large literature devoted to explaining how this works.

Our present objective is not to explain how representation works, nor whether representation can be naturalized using the notion of natural information. Our objective is simply to capture existing usage and distinguish the commitments pertaining to different notions of information. Non-natural information as (often) used in cognitive science has the capacity to get things wrong, so it can be false.

We are now in a position to refine the sense in which Shannon information is non-semantic. Shannon information is non-semantic in the sense that Shannon’s measures of information are not measures of semantic content, whether natural or non-natural. Shannon information is also non-semantic in the sense that, as Shannon himself emphasized, the signals studied by communication theory need not carry any non-natural information. However, the signals studied by communication theory always carry natural information about their source when they carry Shannon information.

In such cases, there is by definition a reliable correlation between source events and receiver events. For any source event, there is a certain probability of the corresponding receiver event. In addition, in virtually all practical applications of communication theory (as opposed to information theory more generally), the probabilities connecting source and receiver

events are underwritten by a causal process. For instance, ordinary communication channels such as cable and satellite television rely on signals that causally propagate from the sources to the receivers. Because of this, the signals studied by communication theory, in addition to producing a certain amount of Shannon information based on the probability of their selection, also carry semantic natural information about their source (depending on how reliable their correlation with the source is). Of course, in many cases, signals also carry non-natural information about things other than their sources – but that is a contingent matter, to be determined case by case.

3.3 Other Notions of Information

There are other notions of information, such as Fischer information (Fisher 1935) and information in the sense of algorithmic information theory (Li and Vitányi 1997). There is even an all-encompassing notion of information, somewhat analogous to some all-encompassing notions of computation. Physical information may be generalized to the point where every state of a physical system is defined as an information-bearing state. Then everything physical is said to be made out of information. Given this all-encompassing notion, allegedly the physical world is, at its most fundamental level, constituted by information (Wolfram 2002). 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 (non-arbitrary, truth-entailing semantic information²⁰), and non-natural information (arbitrary, non-truth-entailing semantic information).

The differences between these three notions can be exemplified as follows. Consider an utterance of the sentence “I have a toothache”. It carries Shannon information just in case the production of the sequence *I-blank-h-a-v-e-blank-a-blank-t-o-o-t-h-a-c-h-e* is a stochastic process that generates letters with certain probabilities. The same utterance carries natural information about having a toothache just in case there is a reliable correlation between the production of the sequence *I-blank-h-a-v-e-blank-a-blank-t-o-o-t-h-a-c-h-e* and having a toothache. Carrying natural information about having a toothache entails probably having a toothache. Finally, the same utterance carries non-natural information just in case the sequence *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 probably having a toothache.

4. How Computation and Information Processing Fit Together

Do the vehicles of computations necessarily bear information? Is information processing necessarily carried out by means of computation? What notion of information is relevant to the claim that a cognitive system is an information processor? Answering these questions

²⁰ See Scarantino and Piccinini (forthcoming) for an explanation of the limited sense in which we take natural information to be truth-entailing.

requires combining computation and information in several of the senses we have distinguished.

We will argue that the vehicles over which computations are performed may or may not bear information. Yet, as a matter of contingent fact, the vehicles over which cognitive computations are performed generally bear information, in several senses of the term. We will also argue that information processing must be carried out by means of computation in the generic sense, although it need not be carried out by computation in a more specific sense.

Before we begin, we should emphasize that there is a trivial sense in which computation can entail the processing of Shannon information. We mention it now but will immediately set it aside, because it is not especially relevant to cognitive science.

Consider an electronic digital computer. Each of its memory cells is designed to stabilize on one of two thermodynamically macroscopic states, called '0' and '1'. Suppose that, as soon as the computer is switched on, each memory cell randomly goes into one of those states. A nonzero Shannon entropy is associated with this state initialization process, since it is a random process, so manipulation of randomly generated initial states by the computer amounts to the processing of Shannon information in this trivial sense.

Note, however, that the digits created by random initialization of the memory cells carry no non-natural semantic information, since the digits in question do not have any conventional semantic content. They carry no natural semantic information either, since they do not reliably correlate with any source other than the thermal component of the electrons' dynamics in the memory cell at power up. Random noise is not the kind of source of natural information that the states of computers – or of cognitive systems, for that matter – have the function of correlating with. Thus, the digits do not carry any natural semantic information, except in the very weak sense that such states correlate with what caused them – whatever that may be.

Since a computer manipulates its digits and, in the above sense, a nonzero Shannon information is associated with randomly generated digits, it follows that a computer processing randomly generated digits is processing information-bearing states. In this trivial sense, such a computer is necessarily a processor of Shannon information. But here we are interested in something more. We are interested not just in whether computation can be regarded as information processing in this basic, physical sense. We are interested in how computation fits with the processing of Shannon information, natural semantic information, and non-natural semantic information about sources in the system's distal environment – the sources the system has the function of responding to. We will take up each of these in turn.

4.1 Computation and the Processing of Shannon Information

Cognitive scientists are generally interested in computational vehicles that bear semantic information; they rarely worry whether computation is the processing of Shannon information, which is not semantic. Nevertheless, it is useful to briefly examine the relationship between computation and Shannon information.

We emphasize at the outset that the notion of processing Shannon information is not entirely clear. Shannon information is not a property of individual signals, which can be manipulated as the signal is manipulated. Instead, Shannon entropy and mutual information quantify statistical properties of a selection process and a communication channel as a whole.

The first thing to note is that, strictly speaking, computational vehicles need not carry Shannon information at all. This is because Shannon information requires random variables – deterministic variables carry no Shannon information whatsoever – whereas computation is entirely compatible with deterministic variables. For instance, consider again a digital computer. For its inputs to be regarded as carriers of Shannon information, they have to be selected in a stochastic way. But the computer works just the same whether its inputs are selected deterministically or probabilistically. The same point can be made about other computing devices, such as analog computers and neural networks. Thus, computation need not entail the processing of Shannon information.

Another relevant observation is that computation need not create Shannon information. This is because computation may be either probabilistic or deterministic – in most real-world applications, computation is deterministic. Deterministic computation cannot create Shannon information – it can only conserve or destroy it. More precisely, Shannon information is conserved in logically reversible computations (which are necessarily deterministic), is reduced in logically irreversible deterministic computations, and may be created, conserved, or destroyed in noisy computations.

These points should not obscure the fundamental importance of communication theory to the understanding of computing systems. Even though, strictly speaking, computing systems need not process Shannon information, it is useful to assume that they do. That is, it is useful to characterize the inputs, internal states, and outputs of computing systems and their components stochastically and analyze system performance and resource requirements using information-theoretic measures and techniques. On this view, the components of computing systems may be treated as sources and receivers in the sense of communication theory (Winograd and Cowan 1963). This allows engineers to design efficient computer codes and effective communication channels within computing systems. The same techniques may be used to analyze neural codes and signal transmission within the nervous system.

So far, we have discussed whether computation is the processing of Shannon information. What about the converse? Processing Shannon information may or may not be done by means of digital computation. First, for Shannon information processing to be done by digital computing, the 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. Just to give an example, a digital-to-analog converter transforms digits into analog signals.

Shannon information processing is, however, a form of computation in the generic sense. As we defined it, generic computation is the functional manipulation of any medium-independent vehicle. Shannon information is a medium-independent notion, in the sense that whether Shannon information can be associated with a given vehicle does not depend on its specific physical properties, but rather on its probability of occurrence relative to its alternatives. Thus, generic computation is broad enough to encompass the processing of

Shannon information. In other words, if a vehicle carries Shannon information, its processing is a computation in the generic sense.²¹

The bottom line is this: although, strictly speaking, computation need not process vehicles that carry Shannon information, it usually does, and at any rate, it is useful to assume that it does. Conversely, the processing of Shannon information amounts to some form of generic computation. Communication theory places constraints on any kind of signal transmission (and hence on the processing of any information carried by the transmitted signals) within computing systems, including cognitive ones. This does not yet tell us what we are most interested in: how does computation relate to semantic information?

4.2 Computation and the Processing of Natural Information

Many people use ‘computation’ and ‘information processing’ interchangeably. What they often mean by ‘information processing’, we submit, is the processing of natural (semantic) information carried by the computation vehicles. This use of ‘information processing’ is especially common in the behavioral and brain sciences. It is thus important to examine whether and in what sense computation is the processing of natural information.

The notion of digital computation does not require that the computation vehicles carry natural information—or more precisely, computational vehicles are not required to carry natural information about the computing system’s *distal* environment.²² In this section, we will focus on whether digital computation entails the processing of natural information about the system’s distal environment.

Granted, natural information is virtually ubiquitous—it is easy enough to find reliable correlations between physical variables. This being the case, the digits present within a digital computer may well carry natural information about cognitively relevant 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 natural information processing.

But digits—the vehicles of digital computation—need not be reliably correlated with anything (in the distal environment) in order for a digital computation to be performed over them. Thus, a digital computation may or may not constitute the processing of natural information. If the computation vehicles *do* carry natural information, this may be quite important. In our example, that certain digits carry natural information about the state of the car explains why the car’s computer can regulate, say, fuel injection successfully. Our point is simply that digital computation does not entail the processing of natural information.

This point is largely independent of the distinction between semantic and non-semantic digital computation. A semantic digital computation is generally defined over the *non*-natural

²¹ In this section, we have not discussed the distinction between semantic and non-semantic notions of computation, as it makes no difference to our present concerns.

²² At least for deterministic computation, there is always a reliable causal correlation between later states and earlier states of the system as well as between initial states and whatever caused those initial states – even if what caused them are just the thermal properties of the immediate surroundings. In this sense, (deterministic) digital computation entails natural information processing. But the natural information that is always “processed” is not about the distal environment, which is what theorists of cognition are interested in.

information carried by the digits—independently of whether they carry natural information (about the distal environment) or what specific natural information they carry. For instance, ordinary mathematical calculations carried out on a computer are defined over numbers, regardless of what the digits being manipulated by the computer reliably correlate with. 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 cannot misrepresent. Thus, even semantic digital computation does not entail the processing of natural information. (More on the relation between natural and non-natural information below.)

This point generalizes to generic computation, which includes, in addition to digital computation, analog computation and neural computation (i.e., the processing of neuronal spike trains). Nothing in the notions of analog computation or neural computation mandates that the vehicles being manipulated carry natural information (about the distal environment). More generally, nothing in the notion of manipulating medium-independent vehicles mandates that such vehicles carry natural information (about the distal environment). Thus, computation may or may not be the processing of natural information.

What about the converse claim: that the processing of natural information must be carried out by means of computation? The answer is that yes, it must, although the computation need not be digital or of any specific kind. Natural information may be encoded by continuous variables, which cannot be processed by digital computers. Or it may be encoded by strings of digits, 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.

But if natural information is processed, some kind or other of generic computation must be taking place. Natural information, like Shannon information, is a medium-independent notion. As long as a variable reliably correlates with another variable, it carries natural information about it. Any other physical properties of the variables are immaterial to whether they carry natural information. Since generic computation is just the processing of medium-independent vehicles, any processing of natural information amounts to a computation in the generic sense.

What about cognition? Does it involve the processing of natural information? Intuitively, what we perceive seems to reliably correlate with our environment. But we don't have to stop at appearances. Large portions of neurophysiology are devoted to the detection, description, and explanation of the detailed causal correlations that exist between neural responses and variables in the external (distal) environment of organisms (Dayan and Abbott 2001, O'Reilly and Munakata 2000). Thus, cognition involves the processing of natural information. This is another way to reach the conclusion that cognition is computation in the generic sense.

In sum, computation vehicles need not carry natural information, but natural information processing is necessarily some kind of generic computation – though not necessarily either digital or analog computation. Since cognition involves the processing of natural information, it amounts to computation in the generic sense, though it doesn't follow that cognition is some more specific kind of computation, such as digital or analog computation.

4.3 Computation and the Processing of 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, and what this has to do with theories of cognition.

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 difference. In the case of *semantic* digital computation – digital computation that has semantic content by definition – any reasonable notion of semantic content turns digital computation into non-natural information processing. As a matter of fact, almost 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*, digital computation is the processing of non-natural information. It *is* if we define a notion of digital computation that entails that the digits are representations. 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 in the sense that provides a mechanistic explanation for the behavior of digital computers—digital computation need not be the processing of non-natural information.

Conversely, the processing of non-natural information need not be carried out by means of digital computation. Whether it is 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. Even when encoded digitally, non-natural information may be manipulated by digital computation or by other processes, as in digital-to-analog conversion.

Generic computation may or may not constitute the processing of semantic information, for similar reasons. Nothing in the notions of analog computation or neural computation mandates that the vehicles being manipulated carry non-natural information. They could be entirely meaningless and the computations would proceed just the same. More generally, nothing in the notion of manipulating medium-independent vehicles mandates that such vehicles carry non-natural information. As in the case of natural information, the converse does not hold. Instead, non-natural semantic information processing is necessarily done by means of computation in the generic sense. For non-natural semantic information is medium-independent – it's defined in terms of the association between an information-carrying vehicle and what it stands for, without reference to any physical properties of the vehicles. Since generic computation is defined as the processing of medium-independent vehicles, the processing of non-natural semantic information (i.e., the processing of vehicles carrying non-natural semantic information) is computation in the generic sense.

The connection between non-natural information processing and generic computation provides yet another route for establishing a generic form of computationalism. All that is required is that cognition involves the processing of non-natural information. That seems plausible enough. Intuitively, our thoughts appear to have non-natural semantic contents. For

example, Violet's thought *that angels are unhelpful* has semantic content, but its content appears to constitute non-natural information. There is no immediate correlation between Violet's thought and anything in her environment, which would be required to turn the content of her thought into natural information. In addition, mainstream theories of language and cognition individuate concepts and cognitive processes at least in part by their non-natural semantic contents (cf. Margolis and Laurence 1999).

Nevertheless, the reality, or at least the scientific explanatory value, of non-natural information has been denied (Quine 1960, Stich 1983). Since we lack room to do justice to this issue, we will simply assume that at least some cognitive systems do process non-natural information at least some of the time. Our reason is simply that our current best theories of cognition are built on that assumption. As long as we lack a better alternative, we propose to explore the commitments and viability of our best theories.

Our work is not done yet. We defined non-natural information in terms of the association between bearers of non-natural information and what they stand for. This representation relation requires explication.

In recent decades, philosophers have devoted a large literature to representation. They have articulated several theoretical options. One option is anti-naturalism: representation cannot and need not be explained or reduced in naturalistic or mechanistic terms – the representational character of cognition simply shows that psychology and perhaps neuroscience are *sui generis* sciences. A second option is anti-realism: representation is not a real (i.e., causally efficacious and explanatory) feature of cognition, although ascriptions of non-natural information-bearing states may be heuristically useful (Egan forthcoming). Finally, naturalists look for explanations of representation in non-intentional terms (Rupert 2008).

A prominent family of naturalistic theories attempts to reduce non-natural information 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; Millikan 2004). According to these theories, non-natural information is, at least in part, natural information. But theories according to which non-natural information reduces to natural information are only one family of naturalistic theories among others. According to other naturalistic theories, non-natural information reduces to things other than natural information (e.g., Millikan 1984, Harman 1987).

Again, we lack the space to discuss accounts of non-natural information in detail. We do wish to point out, however, that a complete theory of cognition cannot simply postulate that cognition involves the processing of non-natural information and take the notion of non-natural information for granted. On the contrary, a complete theory needs to take a stance on non-natural information – what it amounts to, whether it is causally efficacious and explanatory, and if so, how it is naturalistically explained.

Summing up the contents of this section, we have reached three main conclusions. First, computation may or may not be the processing of non-natural information – with the exception of semantic computation, which is not the most theoretically fundamental notion of computation. Second, the processing of non-natural information is a kind of generic computation. Third, although our best theories of cognition postulate that cognition involves the processing of non-natural information, the notion of non-natural information may require a naturalistic explanation. So it is not a conceptual truth that cognitive computation involves the

processing of non-natural information, but it is (plausibly) a contingent empirical truth that holds at least in some cases.

5. *Why the Difference between Computation and Information Processing Matters to Cognitive Science*

As our analysis demonstrates, ‘information processing’ and ‘computation’ are Protean terms. As soon as we clarify their multiple meanings – as we should prior to employing these terms in a rigorous theory of cognition – we realize that the mutual relations between computation and information processing are far from simple. The alleged synonymy between ‘computation’ and ‘information processing’ turns out to rest either on unfruitful stipulation or on disregarding crucial distinctions between the two terms and their multifarious senses.

If this is right, why are the two notions used interchangeably so often, 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 did not clearly distinguish either between Shannon information and semantic information or between semantic and non-semantic computation. But at least initially, they were fairly clear that information and computation played distinct roles within their theory. Their idea was that organisms and automata contain control mechanisms: information is transmitted within the system and between system and environment, and control is exerted by means of digital computation (or perhaps by means of analog computation, another special case of generic computation).

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, many stopped paying attention to the differences between the two. 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. These reasons are not sufficient to actually establish that cognition is digital computation. Whether 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 external (distal) 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 reliably correlated with events occurring in the external environment. Finally, non-natural information can serve to characterize the conventional meaning of mental representations, 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 both instantiated 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 we consider 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 computation in the generic sense. This is a relatively weak conclusion, however, because we defined ‘generic computation’ to include any type of process that manipulates medium-independent vehicles. That cognition involves such vehicles is well established. 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).²³ Theories according to which cognition involves computation but not representation have occasionally been accused of incoherence, on the grounds that computation (allegedly)

²³ 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.

requires representation. On the contrary, it is a corollary of our analysis that these theories are perfectly coherent. Whether they are correct, of course, is another story.

It is independently plausible that cognition does involve the processing of information. After all, natural information is carried by neural and cognitive events by virtue of reliable correlations with environmental events, and such correlations are commonplace (O'Reilly and Munakata 2000, Dayan and Abbott 2001). 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. The notion of non-natural information may be used without realizing that it is different from natural information and may require naturalization. In addition, it will be difficult to integrate findings and explanations from different research areas, because authors working in different areas may use key terms in different senses. For instances, many psychologists use 'computation' to mean digital computation, while many neuroscientists use it to mean computation in the generic sense. Similarly, many scientists use 'information' to mean non-natural information, while many philosophers use 'information' to mean natural information. 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 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. Here is an example. Since many cognitive and neural events reliably 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 in characterizing neural computation. This line of reasoning can be seen at work in some 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 computations in the *generic* sense—once again, a relatively weak 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.

The above mistake is further compounded by the tendency, present in some of the same authors, to describe neural systems as ‘analog’ as opposed to ‘digital’ (e.g., Churchland and Sejnowski 1992, 51). Analog variables are generally assumed to take any real values (within a relevant interval), which entails that they can take uncountably many values. Since computability theory is defined in terms of variables that can take only countably many values, computability theory does not apply to analog variables in any direct way (cf. Piccinini 2008, Section 3.5).

At this point, some people might be tempted to short-cut this discussion by appealing 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.²⁴

So one important consequence of our analysis is that the common view that cognition involves computation simply because it processes information is liable to generating serious mistakes. Nevertheless, there are three important senses—corresponding to the notion of Shannon information and 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 many important cases of connectionist computation). All that follows is that cognition is computation in the weak sense that it involves the processing of medium-independent vehicles.

6. Conclusion

‘Computation’ and ‘information processing’ are commonly used interchangeably, which presupposes that they are roughly synonymous terms. If that were true, computation would entail information processing, and information processing would entail computation. Alas, things are not so simple. Each of the two assimilated notions has more than one meaning. Here is how the landscape looks at the end of our survey:

Question: Is digital computation the same as information processing?

Thesis 1: Digital Computation vs. Processing Shannon Information

- (a) Digital computation (semantic or non-semantic) does not entail the processing of Shannon information;

²⁴ For further discussion of some relevant issues, see Piccinini 2007b, c.

- (b) The processing of Shannon information does not entail digital computation (semantic or non-semantic).

Thesis 2: Digital Computation vs. Processing Natural Information

- (a) Digital computation (semantic or non-semantic) does not entail the processing of natural information (about the distal environment);
- (b) The processing of natural information does not entail digital computation (semantic or non-semantic).

Thesis 3: Digital Computation vs. Processing Non-natural Information

- (a) Non-semantic digital computation does not entail the processing of non-natural information;
- (b) Semantic digital computation entails the processing of non-natural information;
- (c) The processing of non-natural information does not entail digital computation (semantic or non-semantic).

Question: Is generic computation the same as information processing?

Thesis 4: Generic Computation vs. Processing Shannon Information

- (a) Generic computation (semantic or non-semantic) does not entail the processing of Shannon information;
- (b) The processing of Shannon information entails generic computation (semantic or non-semantic).

Thesis 5: Generic Computation vs. Processing Natural Information

- (a) Generic computation (semantic or non-semantic) does not entail the processing of natural information (about the distal environment);
- (b) The processing of natural information entails (semantic) generic computation.

Thesis 6: Generic Computation vs. Processing Non-natural Information

- (a) Non-semantic generic computation does not entail the processing of non-natural information;
- (b) Semantic generic computation entails the processing of non-natural information;
- (c) The processing of non-natural information entails (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 each notion is taken to mean.

Question: Does cognition involve computation, information processing, or both?

Thesis 7: Cognition, Computation, and Information Processing

- (a) Cognition involves the transmission of signals, which are subject to the constraints of communication theory (Shannon information);

- (b) Cognition involves the processing of natural information;
- (c) Cognition, at least in some important cases, involves the processing of non-natural information;
- (d) Cognition is a form of computation in the generic sense.

Thesis 7 summarizes what we take to be the clearest and most plausible formulation of the information processing and computational approach to cognition. Within such an approach, there is room for disagreement among theories of cognitive architecture.

Theories of cognitive architecture are traditionally divided into classical and connectionist. As we have argued, the contrast implied in that division is misleading. If 'connectionism' means associationist connectionism, then there may be a genuine contrast with classical theories. But if a connectionist theory is simply a theory committed to an explanation in terms of neural networks, as the most common use of 'connectionism' suggests, then all cognitive scientists – including classicists – ought to be connectionists.

Many questions about cognition, computation, and information processing remain to be investigated. One that seems especially pressing to us is whether neural computation is a form of digital computation, analog computation, or neither.

References

- Baars, B. J., W. P. Banks, and J. B. Newman, Eds. (2003). Essential Sources in the Scientific Study of Consciousness. Cambridge, MA, MIT Press.
- Barwise, J. and J. Seligman, 1997. *Information Flow: the logic of distributed systems*. Cambridge Tracts in Theoretical Computer Science, 44.
- Bechtel, W., & Abrahamsen, A. (2002). *Connectionism and the Mind: Parallel Processing, Dynamics, and Evolution in Networks*. Malden, MA: Blackwell.
- Block, N. (1978). Troubles with Functionalism. Perception and Cognition: Issues in the Foundations of Psychology. C. W. Savage. Minneapolis, University of Minnesota Press. **6**: 261-325.
- Bradbury J. W., Vehrencamp S.L. (2000), Economic models of animal communication. *Animal Behavior*, 59:259-268.
- 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.
- Copeland, B. J. (2000). "Narrow Versus Wide Mechanism: Including a Re-Examination of Turing's Views on the Mind-Machine Issue." The Journal of Philosophy **XCVI**(1): 5-32.
- 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.
- Edelman, G. M. (1992). Bright Air, Brilliant Fire: On the Matter of the Mind. New York: Basic Books.
- Egan, F. (forthcoming). "A Modest Role for Content." Studies in the History and Philosophy of Science.
- Fano, R. M. (1961). Transmission of information; a statistical theory of communications. New York, M.I.T. Press.
- Feigenbaum, E. A. and J. Feldman (1963). Computers and Thought. New York, McGraw-Hill.
- Feldman, J. A. & Ballard, D. H. (1982). Connectionist Models and their Properties. Cognitive Science, 6, 205-254. Reprinted in Anderson & Rosenfeld, 1988, pp. 484-508.
- Fisher, R. (1935). The Design of Experiments. Edinburgh, Oliver and Boyd.
- Floridi, L. (2005). "Is Semantic Information Meaningful Data?" Philosophy and Phenomenological Research 70(2): 351-370.
- Fodor, J. A. (1975). The Language of Thought. Cambridge, MA, Harvard University Press.
- Fodor, J. A. (1981). "The Mind-Body Problem." Scientific American 244(January 1981).
- Fodor, J. A. and Z. W. Pylyshyn (1988). "Connectionism and Cognitive Architecture." Cognition 28: 3-71.
- 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.
- Globus, G. G. (1992). Towards a Noncomputational Cognitive Neuroscience. Journal of Cognitive Neuroscience, 4(4), 299-310.
- 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.
- Griffiths, P.E (2001) Genetic Information: A Metaphor in Search of a Theory. Philosophy of Science 68(3): 394-412
- Harman, G. (1973). Thought. Princeton, Princeton University Press.
- 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.
- Hopfield, J. J. (1982). Neural Networks and Physical Systems with Emergent Collective Computational Abilities. Proceedings of the National Academy of Sciences, 79, 2554-2558. Reprinted in Anderson & Rosenfeld 1988, pp. 460-464.
- Horgan, T., & Tienson, J. (1996). Connectionism and the Philosophy of Psychology. Cambridge, MA: MIT Press.
- Margolis, E. and S. Laurence, Eds. (1999). Concepts: Core Readings. Cambridge, MA, MIT Press.
- Li, M. and P. Vitányi (1997). An Introduction to Kolmogorov Complexity and Its Applications, Second Edition. New York, Springer.
- 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.
- Marr, D. & Poggio, T. (1976). Cooperative Computation of Stereo Disparity. *Science*, 194, 283-287.
- Maynard Smith, J. (2000). "The concept of information in biology." *Philosophy of Science* **67**(2): 177-194.
- 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*, McGraw-Hill
- 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.
- Minsky. M. (1968), *Semantic Information Processing*, 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.
- Penrose, R. (1994). *Shadows of the Mind*. Oxford, Oxford University Press.
- Perkel, D. H. (1990). Computational Neuroscience: Scope and Structure. In E. L. Schwartz, *Computational Neuroscience* (pp. 38-45). Cambridge, MA: MIT Press.
- 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 Logic1: 103-105.
- Quine, W. V. O. (1960). Word and Object. Cambridge, MA, MIT Press.
- Ramsey, W. M. (2007). Representation Reconsidered. Cambridge, Cambridge University Press.
- Rieke, F. et al. (1997). Spikes: Exploring the Neural Code. Cambridge, MA, MIT Press.
- Rubel, L. A. (1985). "The Brain as an Analog Computer." Journal of Theoretical Neurobiology4: 73-81.
- Rumelhart, D. E., J. M. McClelland, et al. (1986). Parallel Distributed Processing: Explorations in the Microstructure of Cognition. Cambridge, MA, MIT Press.
- Rupert, R. D. (2008). "Causal Theories of Mental Content." Philosophy Compass 3(2): 353-380.
- Scarantino, A. (2008). Shell Games, Information, and Counterfactuals. *Australasian Journal of Philosophy* 86 (4): 629 – 634.
- Scarantino, A. and Piccinini, G. (forthcoming). "Information without Truth". *Metaphilosophy*.
- Scarantino, A. (ms.), A Theory of Probabilistic Information.
- Seyfarth, R.M. & Cheney, D.L. 1992. Meaning and mind in monkeys. *Scientific American* 267, 122-129.
- Shagrir, O. (2006). "Why We View the Brain as a Computer." Synthese153(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.
- Smolensky, P. (1988). "On the Proper Treatment of Connectionism." Behavioral and Brain Sciences 11: 1-23.
- Smolensky, P. & Legendre, G. 2006. The Harmonic Mind: From Neural Computation to Optimality-Theoretic Grammar. Vol. 1: Cognitive Architecture; vol. 2: Linguistic and Philosophical Implications. Cambridge, MA: MIT Press.
- 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.
- Struhsaker, T.T. (1967). Auditory Communication among Vervet Monkeys (*Cercopithecus aethiops*). In *Social Communication among Primates*, ed. S.A. Altmann. University of Chicago Press.
- Winograd, S. and J. D. Cowan (1963). Reliable Computation in the Presence of Noise. Cambridge, MA, MIT Press.
- Wolfram, S. (2002). A New Kind of Science. Champaign, IL, Wolfram Media.