9-11-2011

The "Human Mental State Multiple Realizability in Silicon Thesis" is False: An Argument Against Computational Theories of Consciousness

Matthew Stuart Piper
University of Missouri-St. Louis, mpiper5@slu.edu

"The 'Human Mental State Multiple Realizability in Silicon Thesis' is False:

An Argument Against Computational Theories of Consciousness"

Matthew Stuart Piper
M.A., Interdisciplinary Studies, Sonoma State University, 2008
B.A., Honors Psychology, Saint Louis University, 1996

A Thesis

Submitted to The Graduate School of the

UNIVERSITY OF MISSOURI – ST. LOUIS
in partial fulfillment of the requirements for the degree

MASTER OF ARTS
in

PHILOSOPHY
with an emphasis in Philosophy of Mind

December 2011

Advisory Committee

Berit Brogaard, Ph.D.
Chairperson

Gualtiero Piccinini, Ph.D.

John Brunero, Ph.D.

The "Human Mental State Multiple Realizability in Silicon Thesis" is False: An Argument Against Computational Theories of Consciousness

Matthew Piper

1. Abstract

Two of the most important concepts in contemporary philosophy of mind are *computation* and *consciousness*. This paper explores whether there is a "strong" relationship between these concepts in the following sense: is a computational theory of consciousness possible? That is*, is the right kind of computation sufficient for the instantiation of consciousness*? In this paper, I argue that the abstract nature of computational processes precludes computations from instantiating the concrete properties constitutive of consciousness. If this is correct, then not only is there no viable computational theory of consciousness, the Human Mental State Multiple Realizability in Silicon Thesis is almost certainly false.

2. Introduction

The HMSMRST is the Human-Mental-State-Multiple-Realizability-in-Silicon Thesis, which holds that silicon-based systems having artificial experience[1] can be physically constructed (i.e., silicon-based machine consciousness is physically possible). Although computers need not be made from silicon, they are the archetypal, and most complex, silicon-based systems. So, on the most plausible reading, the HMSMRST is true if computers can be conscious. Given that common sense suggests that computers are not conscious, the burden for proponents of the HMSMRST, *given that the sine qua non of computers is computation*, is to provide a computational theory of consciousness.

In this paper, I argue that *all* computational theories of consciousness (CTCs) fail. CTCs hold that the right kind of computation is sufficient for the instantiation of consciousness. Given the widely-recognized importance of his work, I will use David J. Chalmers' Thesis of Computational Sufficiency as a paradigm case. I will argue that it fails for a reason that can be formalized as a general problem plaguing any CTC: the medium-independent properties (MIPs) constitutive of computational processes are insufficient to instantiate the medium-dependent properties (MDPs) constitutive of consciousness. MIPs, like graphemes, are properties whose causal role (e.g., symbolic meaning) does not depend upon the physical properties of the vehicles by which the relevant information is transferred (e.g., paper), while MDPs, like digestion, have causal roles (e.g., decomposition) that directly depend on physical properties of the relevant vehicles (e.g., enzymes). Since computations, as abstract descriptions, are MIPs, they must be *implemented* to generate MDPs. However, this makes potential implementation properties central to the feasibility of instantiating consciousness in artificial systems.

[1] "Artificial experience" denotes the instantiation of phenomenal consciousness in non-biological artifacts/systems.

The problems that arise for CTC advocates are two-fold, taxonomical and empirical. The taxonomic problem is that adverting to detailed physio-causal properties of implementation vehicles threatens to subvert the legitimacy of calling such theories "computational." The empirical problem is the following. Given the necessary role of implementation properties, and the fact that functions supervene on structures, it follows that physical differences can legislate mental differences. After distinguishing weaker and stronger varieties of implementation requirements, and showing why a plausible CTC requires a *thick* theory of implementation, I will examine the implementation requirements for human consciousness. Given empirical data suggesting that consciousness depends on very specific physical properties of the brain, for which there are *no known implementation surrogates*, I argue that CTCs will fail to generate the relevant MDPs. I will conclude by showing why this implies that the HMSMRST is almost certainly false.

The outline of this paper is as follows. In § 3, I will explain Chalmers' Thesis of Computational Sufficiency (TCS). In § 4, I will argue that Chalmers need for implementation properties either makes his account computationally vacuous or false, in the case of human consciousness. Either way, the TCS cannot carry the weight asked of it. In § 5, I will respond to objections. In § 6, I will show that Chalmers' TCS fails for reasons that are applicable to all CTCs, and I will formulate a general argument against CTCs based upon the difference between MIPs and MDPs. In § 7, I will consider objections to this general argument. In § 8, I spell out the implication of the foregoing for the HMSMRST, and then summarize the route taken in § 9.[2]

3. Chalmers' Thesis of Computational Sufficiency

In the essay, *A Computational Foundation for the Study of Cognition* (1994)[3], Chalmers argues for the Thesis of Computational Sufficiency, which is essentially the following:[4]

---

[2] This paper will assume that one set of properties can realize another set of properties (e.g., structural properties realizing functional properties). Such property sets will, in general, be related mereologically. In some cases, it is possible that what appears to be one set of properties (X) realizing another (Y) is actually a case of one set of properties (X) known under two different modes of presentation (X_1 and X_2) (cf. Shoemaker's subset account of properties (2007)).

[3] Despite the title of the paper in view, theories of cognition are not my target. I am not contesting Chalmers' claims in those regards, and am, in fact, sympathetic to many aspects of a computational approach to cognition. That having been said, Fodor (2008) presents some key obstacles any theory of cognition must overcome (esp. chapter 4). Thus, the CTC arguments presented here are tangential to (and this paper is agnostic regarding) the viability of *computational theories of cognition* in general (including Chalmers').

[4] This argument is also presented in chapters 7 and 9 of Chalmers (1996) book, *The Conscious Mind*. The unpublished (1994) paper will be the chief reference as it is a relatively concise statement, containing a minimum of extraneous material, whose import is generally recognized: notably, *only 3 months after the submission of this paper to the Journal of Consciousness Studies*, the Journal of Cognitive Science announced a Call for Papers *solely* on responses to Chalmers' unpublished (1994) paper!

(P1) "*A computation provides an abstract specification of the causal organization* of a system." (p. 11)

(P2) "Cognitive systems have their *mental properties in virtue of their [abstract] causal organization*." (p. 11)

(C1) "The right kind of *computational structure suffices for possession of a mind*." (p. 1)

Let us look at each step of the argument.[5]

(P1) Chalmers' argument rests upon a *causal theory of computation* (C_C). According to C_C, a system implements computation X if there is an isomorphism between the causal relations between the states of the physical system Y and the state transitions specified by the formal description of X. Thus, the key feature is "that the formal state-transitional structure of the computation mirrors the causal state-transitional structure of the physical system." (Chalmers, 1994; p. 4) Given C_C, P1 holds by definition: computations preserve abstract causal organization (ACO) since, syntactically speaking, they are nothing above and beyond ACOs.

(P2) The idea behind P2 is that, unlike most properties, mental properties depend only upon ACO. Properties depending only upon ACO are, in Chalmers' words, "organizationally invariant." By organizational invariance, Chalmers means that "[i]t does not matter how we stretch, move about, or replace small parts of a cognitive system: as long as we preserve its causal topology, we will preserve its mental properties." (ibid., p. 12) That is, Chalmers argues that mental properties depend only upon the *abstract* organization of a system's causal patterns and not the causal realizers as such. In other words, as long as the ACO between two systems is invariant, they will share the same mental properties, if any. In order to argue that computation (individuated causally) is sufficient for the possession of a (conscious) mind, Chalmers claims that mental properties are of two kinds, psychological and phenomenological, and then argues that each kind of property is organizationally invariant.

Psychological properties (Chalmers gives "belief and perception" as examples) are declared to be those "characterized by their causal role" within an overall causal system. Chalmers adverts to the work of Armstrong (1968) and Lewis (1972) in claiming that "[s]ystems with the same causal topology…will share their psychological properties" (Chalmers, 1994; p. 12). The basic idea is consonant with Functionalist theories of mind, in which mental properties are individuated by their

[5] Although the argument as stated here is formally invalid, it serves the purpose of presenting the basic ideas in the author's own words. Charitably, one valid form of the argument might be the following: (P1) Computations = ACOs. (P2) The implementation of *some* ACOs is sufficient for realizing mental properties. (C1) The implementation of *some* computations is sufficient for realizing mental properties.

functional roles within a system, as opposed to any physical properties *per se*. If this is right, then a C_C should instantiate any psychological properties a system would possess in virtue of preserving ACO.

Phenomenological properties, on the other hand, are not *prima facie* definable in terms of their causal roles, as Chalmers himself admits. Establishing that phenomenological properties can be accounted for via ACO therefore requires argument. Chalmers provides his "Dancing Qualia" Argument (DQA) – a *reductio* – for this purpose. Chalmers begins by assuming that agents with identical ACOs could have different experiences in virtue of having different material constitutions (silicon vs. neurons, e.g.), as type-physicalism might require. He then asks us to conceive of changing one agent into the other by the replacement of parts (neural parts replaced by silicon, say) while preserving its ACO.[6] *Ex hypothesi*, the experience of the agent under transformation would change (as the parts were replaced), but there would be no change in abstract causal organization and therefore no means whereby the agent could "notice," as he puts it, the shift in experience. To imagine, however, that it makes sense to say an agent could have qualitative changes in experience but be unable to notice those changes in experience seems incoherent. Given the absurdity of the conclusion, therefore, Chalmers' rejects the initial premise that agents with identical ACO can have different experiences. Thus, by his DQA, Chalmers defends his view that phenomenological properties are organizationally invariant. Having argued that both types of mental properties – psychological and phenomenological – are organizationally invariant, Chalmers has defended P2.

(C1) In arguing for the TCS, Chalmers advocates a C_C, in which a computation is simply an abstract specification of causal organization. Accordingly, if computations are ACOs, as P1 argues, and some ACOs are sufficient for mental properties, as P2 argues, then some computations are sufficient for mental properties, and C1, "the *right kind* of computational structure suffices for the possession of a mind," straightforwardly follows. Interestingly, as Chalmers admits, there are no positive proposals for what kind/type of computational structure is sufficient for consciousness. In any case, to a first approximation, such is Chalmers' argument for the TCS.

The core point for our purposes is that Chalmers presents the TCS as a CTC.[7] In other words, it holds that computation is sufficient for the instantiation of (both psychological and phenomenological properties of) consciousness. Let us turn to a critique of the TCS, arguably the most well-known and direct CTC on offer. After exposing a critical dilemma, I will eventually argue that it is symptomatic of a more general problem undermining all CTCs.

---

[6] Of course, this *assumes* that the relevant causal organization can be preserved when changing structures, which begs the question against type-physicalists. See below.

[7] Chalmers' TCS is arguably the clearest and most straightforward CTC. A small selection of other authors developing or working towards CTC accounts include Mathis and Mozer (1996), Cotterill (1998), Haikonen (2003), Holland (2003), Aleksander (2005), Sloman (2005), Cleermans (2005), Rolls (2007), Ballard (2010), Stuart (2010) and Chella and Manzotti (2011). Later, I will argue that the problems infecting TCS apply, in a more general form, to all CTCs.

4. Critique of Chalmers' Thesis of Computational Sufficiency

Chalmers' TCS is complex and subject to varied objections.[8] To keep discussion manageable, I will restrict my objections to those concerning a fundamental tension that arises out of the deployment of the concept ACO. As I'll explain, Chalmers' account faces the burden of either showing how *abstract* descriptions can generate *real* semantic properties or showing that silicon-based systems (SBSs) implementations can have all the kinds of *concrete* mechanisms whose activities are causally efficacious in the generation of human consciousness. As I'll show, given this dilemma, and the best scientific evidence about the realization conditions of consciousness, the TCS proves false, vacuous or question-begging.

Obviously, ACO depends upon the concepts of *abstraction*, *causation* and *organization*. Chalmers' basic idea is that there is a basal level of causation (in the brain) that generates consciousness. However, if mental properties are "organizationally invariant," as he argues, it is not necessary to generate this causal activity in the exact medium-dependent manner it occurs in us; rather, it is only necessary to instantiate a computation whose states and transitions have a structure that is isomorphic to the transitions of the actual causal activity. Before looking in more depth at the proposal, it will be helpful to consider the general idea in play. What is interesting about this thesis is that it might *seem* to assume that an abstract representation of causal organization can be productive.

But is this putative assumption warranted? It is unlikely. Take a blueprint of a watch, faucet or computer. What causal powers do such blueprints possess? Though they may be said to possess certain dispositional powers (if utilized by an engineer, perhaps), they are causally inert on their own. However, they are most definitely ACOs. Perhaps *abstract specifications of causal organizations are, as a genus, causally impotent*. Similarly, I can lecture (with diagrams, mathematics and movies, no less) all day about the causal forces in solids that permit them to suspend objects (a drink, say), but these have yet to keep my coffee from falling to the floor. This is of course because descriptions *qua* descriptions – even when accurate formalisms of real causal properties – never possess the actual causal powers of such properties. Why would we expect consciousness to be any different?

---

[8] There is not room to mention, let alone discuss, all the objections to aspects of the TCS. Certainly, David Chalmers has done the best job of cataloging replies (see his very helpful "Responses to articles on my work"). Here, I will make only a few points. Generally speaking, a number of writers have taken issue with Chalmers' ambition for an account of computation. Perhaps the most well-known objection comes from John Searle, who criticized computational accounts of the mind for trying to get semantic water from a syntactic well. Roughly speaking, Searle's "intrinsic intentionality" objection is that computations are individuated non-semantically (e.g., syntactically), and, given that human consciousness is self-evidently semantic, computations cannot instantiate human consciousness (Searle, 1980). This objection dovetails with discussion below. Other objections have been leveled against Chalmers' principle of organizational invariance (e.g., Seager, 1995; Pelczar, 2008), which we will cover below, and Chalmers' DQA and related thought experiments (e.g., Van Heuveln et al., 1998; Bostrom, 2006), which we will not discuss, since Chalmers himself admits that his thought experiments are inconclusive at best (Chalmers, 1999).

As we have already seen, one aspect of Chalmers' reply is his DQA thought experiment. Although ingenious, its legitimacy depends on the assumption that a silicon chip can instantiate the relevant causal powers of a neuron, which, however, just begs the question (against a type-physicalist characterization of consciousness, e.g.).

Tellingly, the difficulty reconciling the concepts "abstract" and "causal organization" can be clearly seen in Chalmers' own attempts to deploy the synonym concept "organizational invariance." The TCS depends on the premise that mental properties are organizationally invariant. As we saw above, organizational invariance obtains insofar as the physical deformation of a system, or system part, doesn't affect its functional role; that is, organizational invariance is a property of systems characterized by ACO. This is because as long as the abstract causal organization is unchanged (regardless of the concrete organization), the function of organizationally invariant systems will remain unchanged.

Chalmers illustrates the concept of organizational invariance by contrast with properties that are not organizationally invariant, such as flying and digesting. He explains that flying is not organizationally invariant because "we can move an airplane to the ground while preserving its causal topology, and it will no longer be flying" and digestion is not organizationally invariant because "if we gradually replace the parts involved in digestion with pieces of metal, while preserving causal patterns, after a while it will no longer be an instance of digestion…" (Chalmers, 1994; p. 9)

However, contrary to Chalmers' intent, the causal organizations that generate flying and digestion are exactly what are *not* conserved in these examples. Grounding the plane, in the first example, fails to replicate the requisite airflow patterns whose causal organization underpins the property of flying, while replacing digestive parts with metals in the second case fails to replicate the requisite enzyme patterns whose causal organization underpins the property of digestion. Since Chalmers' examples fail, it is unclear what the idea amounts to. Perhaps Chalmers' difficulties are to be expected given the nature of the complex concept he employs.

The key tension within ACO is that "abstraction" is the *glossing over of* "causal" roles. Straightforwardly, if anything is a concrete, medium-dependent property, *real causation* is; and if anything is a medium-independent property, *abstraction* is. Real causation involves the transformation of real forces, while abstractions do not. Thus, the causal properties of real forces are medium-dependent. Consider: the same force F (a bullet of mass M traveling at acceleration A, say) plays much different causal roles within different media (gas vs. water vs. glass vs. steel). Conversely, abstract properties are medium-independent because the vehicles utilized are irrelevant to their causal role. For example, whether a number is written, spoken or simply thought (or otherwise) does not affect its relevant (mathematical) role.

So, what are common examples of ACOs? Blueprints, models, maps, and simulations. Now, I am not arguing that various scientists and philosophers haven't made good use of blueprints, models, maps and simulations (and therefore the concept of ACO). Obviously they have. But the uses are *descriptive*. No one gets wet from a faucet blueprint, a model of a storm front, or a map of the Great Lakes. Because ACOs are syntactic in nature, the properties of ACOs (unlike the properties of the entities they represent) are medium-independent, instantiation-independent. ACOs lack the intrinsic powers of the properties they formalize.

Accordingly, it is inappropriate to apply ACO to the concrete phenomenon of consciousness because it makes no sense to hold something simultaneously medium-dependent (concrete) *and* medium-independent (abstract) *in respect of the same property*. While in the case of blueprints and maps, ACO may be an adequate characterization, in the case of real artifacts and terrain, real implementation details matter – abstract characterizations fail to instantiate the relevant features precisely *because* they are abstractions. In other words, *you can't eat causal cake with an abstract fork*.

Above, I claimed that Chalmers faces a dilemma. It is this. There are two aspects to Chalmers' account: the computational formalism, which is a syntactic description of causal state-transitions, on the one hand, and an implementation requirement, on the other. Chalmers accepts that a theory of mental properties must account for semantic properties. Thus, he must explain how *either* the computational formalism *or* the implementation properties *can account for such semantics*.

Chalmers clearly rejects the first option, stating "syntax may not be enough for semantics…" (ibid.; p. 13). He agrees that abstract entities (such as algorithms/programs) are insufficient to generate real causal powers, let alone cognition or semantics. In more detail he writes, "It is the program that is syntactic; *it is the implementation that has semantic content*….syntax may not be enough for semantics, but the right kind of causation is." (ibid. Italics added.) Thus, he accepts the second horn. In order to account for the semantics of consciousness, he posits that they are derived from "real physical causation" implementation properties (ibid.).

However, this invites two new problems. First, if the "real physical causation" implementation details are the crucial ones for cognition, semantics and consciousness (ibid.), then a type-physicalist account becomes, to some degree, preferable, because the essential causal work is being done by *physical* properties. To salvage a computational approach to consciousness, an advocate might hold that while physical tokening is necessary, it is the causal organization as given by the computational formalism that gives shape, so to speak, to the mental properties. However, it is unclear that this is different than (i.e., not extensionally equivalent to) garden-variety type-physicalism. Any type-physicalist approach will hold that it is both the physical properties and their organizational relations that matter (Scheutz, 2001). If this is right, calling such an approach "computational" begins to sound vacuous.[9]

The second problem is that locating the semantics in causally real implementation properties has the consequence that physical implementation asymmetries can generate semantic asymmetries, which can preclude multiple realizability. This follows from the simple reason that not any substrate can implement any function. No amount of CO_2 or toilet tissue can implement a submarine or a quantum computer, e.g. In general, the more complex the function in question, the more complex the substrate needed for its implementation. Arguably, humans are categorically more complex than any thermodynamically inert SBS can be, given the complexity of the hierarchical nesting of thermodynamic work cycles concurrently causally efficacious in humans (cf. Kauffman, 2002; Collier, 2008). Accordingly, *if the functions constitutive of consciousness are complex enough, then silicon-based systems will not possess enough structural degrees of freedom to implement the relevant properties*.

Below, I will argue that this is exactly the case. However, even if my demonstration were deemed inconclusive, Chalmers' argument remains critically question-begging. His assertion that mental properties can be captured by the implementation of their *abstract* causal properties simply begs the question against those who would assert that consciousness is a *medium-dependent* phenomenon, not realizable in any abstract process (e.g., Block, 1978; Searle, 2004; Hill, 2009; McLaughlin, 2010; Polger, 2011). Since it is quite unclear both what the causal organization of mental properties as instantiated in *humans* is, on the one hand, and whether other systems can implement *our* causal organization, on the other (cf. Kary and Mahner, 2002), this seems a bit gratuitous on his part.[10] More specifically, Chalmers would seem to beg the question against anyone who holds that "the causal structure of a physical system (i.e., its causal complexity) is determined by its various physical states *and* their causal relations among each other…[because] then computation, and as a consequence functionalist descriptions, will not be able to cover and capture *all* aspects of the causal structure." (Scheutz, 2001; p. 563)

As we have seen, the TCS faces multiple objections. First, if the TCS is taken to suggest that abstract formalisms can generate concrete phenomena, then it is simply false. To escape from this, Chalmers holds that implementation properties are necessary for realizing mental properties, as implementation properties, being physical properties themselves, generate real physical causality, and so can account for the semantics of mental properties. This, though, has two sequelae. One, because the critical role is played by physical properties, it appears the theoretical uses are

[9] I am indebted to an anonymous referee for suggesting the possibility of pressing Chalmers on this point.

[10] This is important because, *pace* Shoemaker (2007), *human* consciousness serves as the necessary indexicalization of *our* concepts of consciousness/qualia/phenomenology (McLaughlin, 2003). The key point is that if we do not semantically ground phenomenology and qualia in/by *human* consciousness, we have no way of recognizing the truth value of *our* concepts (i.e., properly employing them).

readily available to type-physicalist approaches. On this reading, the TCS is a somewhat vacuous account of computation, as all physicalist approaches would thereby become computational. The second result is that because implementation properties are physical properties, physical differences become critical in the determination of mental properties. The TCS, on this reading, is at least question-begging, and, given the evidence discussed in § 6, below, likely false.

5. Replies and Counters

How might would-be defenders of the TCS reply to the objections in § 4? There are multiple considerations.

First, it might be objected that my critique of ACO is misplaced given that because "causal relations" are what Functionalists take to be medium-independent (i.e., abstract), there is no tension within ACO. This is tangential to my point, however. Certainly, "causal relations," as typically adverted to in the literature, denote causal pattern-*descriptions* (without the instantiated forces), and hence are abstract. But causal relations, *qua* abstract, formal descriptions, have no instantiation-dependent properties whatsoever. To avoid the *fallacy of reification*, one must hold that causal relations need to be *implemented* to instantiate "real" properties. Thus, what I take to be in tension are the concepts "abstract," on the one hand, and, "causation," on the other, *within the context of implementation*. As they say, you can't have your cake and eat it too.

Second, a more substantive objection might be the following: It seems open to a defender of the TCS to say that, very obviously, the TCS can account for both the abstract (medium-independent) and concrete (medium-dependent) aspects of mental properties: the implementation generates the medium-dependent properties, while the ACO generates the medium-independent properties. Accordingly, both the syntax and semantics are explained, and all objections countered.

However initially attractive, adverting to the MDPs of an implementation as critical features of a TCS account invites the rejoinders encountered above. As we saw, not only is it vacuous, or at least trivial, to call an approach that relies *essentially* on the *physical* properties of systems "computational," but physical differences in implementation substrates, on this view, can be expected to generate semantic differences in certain contexts, or to preclude certain systems from implementing computations *tout court*. This is because *MDPs are not multiply realizable in a holistic sense*. So, different implementation substrates will have, and lack, different MDPs. Chalmers himself is very clear that "for any given complex [computation], very few physical systems will have the causal organization required to implement it" (1996; p. 319).   It is not at odds with this latter point, for instance, to claim that, for "mechanical" reasons, the set of non-human implementations of human consciousness is zero, as I will argue in § 6.

But just which MDPs are necessary to generate the requisite mental states? And what is the relationship between implementation properties and phenomenal properties?

One view – call it the *thin* theory of implementation – is that implementation details *as such* are tangential to phenomenology. On this view, what is important is that there is *a* physical substrate – providing *some* real causality – of a computational process. Other than the simple requirement of having enough parts to represent the informational states of the computation, what physical mechanisms operate between elements is tangential to phenomenology. Chalmers' position on the Chinese Room (1996; chapter 9) and elsewhere (e.g., 1994; p. 5) suggests this reading. The idea is that no matter what the actual causal commerce between informational elements/symbols (slips of paper, cans and string, etc.), the fact of causal interaction *simpliciter* provides all the semantic grounding necessary. On this view, implementations need not realize most of the MDPs of the original system; rather, the fact that there is an implementation substrate at all provides all the MDPs necessary.

The problem here is bald implausibility. Two illustrations should suffice to make the point. First, imagine a huge network of slips of paper tied to each other in a fashion isomorphic to the causal commerce of a neural network whose instantiation in us generates consciousness. Imagine that this network of string and paper slips is spread out over a plateau 1000 km^2. If a great wind blows over this plain, we can imagine that the slips of paper will change "states" depending on both the network of connections and the (input) energy from the wind (cf. Chalmers' (1996; Chapter 9) discussion of the Chinese Room thought experiment). Now, such a network is a concrete implementation of the ACO of a system whose activity generates consciousness. Just as neural states affect one another, the slips of paper are so connected to each other (by the strings) that the activity in each paper partially determines the activity of others. We can picture this grand network of strings and little slips of paper spread out over a massive plateau, rustling in the breeze. So…is the system of strings and paper slips conscious?

Chalmers very bravely claims "yes." To his way of thinking, the slips of paper purportedly capture the ACO of the real neural network that generates consciousness, and the strings and papers generate the real physical causality necessary for semantic content. But, given that his argument for why ACO should generate phenomenology (his DQA, above) begged the question, there doesn't seem to be any good reason to think a vast field of strings and paper slips should qualify as a conscious entity. And there seem to be a number of obvious reasons that it shouldn't. One, the strings and paper fail to instantiate a unified substrate; the material collective of strings and paper is quite *ad hoc* and fails to resemble a natural kind, which minds and consciousness reasonably are. Two, the system has no intrinsic dynamics. While it is true that the system implements a pattern of rule transitions, as governed by string arrangements between papers, it has no intrinsic movement of its own. It is essentially inert, and, as Searle argues, has no "intrinsic intentionality" (Searle, 2004). Three, admitting that such a "system" is conscious would imply that minds are either omnipresent in any causally dense region of the universe, if the states are observer-independent, or grossly underdetermined, as the ascription of causal relations and symbols is observer-dependent. But if any one posit is certain for each of us, it is that our being conscious is independent of any else's observation of us. This is certainly not a welcome result for those who take consciousness seriously. These are not reasonable conclusions.

Secondly, compare a huge abacus and a supercomputer. Both can serve to implement computations. If mental states can be realized by variously implemented computational formalisms, then each system, given enough time, should be able to instantiate the same mental state. Experience, though, is not atomic; experience is temporally structured (Blackmore, 2011). But in what sense could these two systems realize the same temporally-structured phenomenology if the experience were dependent on contingent and transiently-available environmental stimuli, as happens, say, when playing music in an ensemble, coordinating in-battle military maneuvers or talking one's way out of a speeding ticket?

Of course, a supercomputer will process a given computational sequence orders of magnitude faster than much simpler machines, such as abaci. Accordingly, the processes will quickly fall out of step relative to any varying environment/feedback process. How can they realize the *same* experiences when the temporal dynamics of their implementations (such as their temporal sensitivities to input) are so different? More generally, *given that temporal and information-density dynamics are defining features of, arguably all, mental states, in what sense could implementation substrates with vastly different processing capacities generate equivalent experiences?* Examples will sharpen the picture. Whereas the three musical notes A – C – E, if they are presented quickly enough, can generate the expectation of a return to A, or simply an appreciation of the A-minor chord, this will not happen at slower sampling/presentation rates. Similarly, the phenomenology of watching films is obliterated if the still images occur too slowly. The general reason is that, in these and other cases involving perception and perceptual illusion, there are temporal dynamics that must obtain in order to generate key gestalt effects. Without such temporal dynamics, the relevant phenomenology is absent.

Thus, the problem for a proponent of a thin theory of implementation is that "processing speed," "energy throughput," and "resource efficiency" are key constituents of causal individuation: *the informational, energetic, and resource efficiency capacities of a system critically inform its causal topology* (cf. Eliasmith, 2002). A thin theory of implementation cannot recognize these essential differences, and thus implies that adult human consciousness can be instantiated in a system of strings and paper slips. Accordingly, in the context of consciousness, a thin theory of implementation seems to have consequences far too unbelievable and unmotivated to allow us to take it seriously.

One remedy is to take a stronger view of implementation. According to a *thick* theory of implementation, implementation details are directly relevant to the realization of phenomenology. On this view, implementing structures must have as many functional degrees of freedom per unit time as the system being represented, which requires implementation substrates have a sufficient number of structural degrees of freedom. The following quote from Chalmers might be taken to suggest such a view: "For each neuron, there will be a memory location that represents the neuron, and each of these locations will be physically realized in a voltage at some physical location. *It is the causal patterns among these circuits,* just as it is the causal patterns among the neurons in the brain*, that are responsible for any conscious experience that arises*" (Chalmers, 1996; p. 321; Italics added). Among other things, this passage seems to imply that real physical implementation properties generate analogous phenomenological properties (cf. Chalmers' (1996) principles of coherence and double aspect theory), that potential implementation substrates have medium-dependent limitations, and that there are strong constraints on what materials can implement a given function, as determined by the design and functioning of the original system whose behaviors are being represented/mimicked. On this, much more plausible, view, the Chinese Room, the strings and papers mentioned above, and many other *ad hoc* arrangements will not be conscious in virtue of the fact that the implementation substrates will fail to generate the necessary MDPs. Such MDPs might include, but are not limited to, electrodynamical circuit properties, implementation substrate unity properties, or processing and resource capacity properties, e.g.

The problem for the CTC advocate is that this cuts against her position, too. It can motivate the conclusion that no silicon-based system can generate human-like experiential states, because, if we say that the *physical* differences between these systems' implementations (which subvene differences in functional degrees of freedom through differences in processing speed, power, and vehicles) can ground mental state differences, then we are forced to hold that physical differences between SBS and humans can ground essential mental differences. As we'll see below, there are excellent reasons to think that this is, in fact, the case. Perhaps the simplest argument is the following. Adult human systems are categorically more complex than SBSs. It is reasonable that our consciousness is generated by the most complex class of functions we instantiate. Accordingly, SBSs may be debarred from realizing *human* consciousness, for the same reason that an adding machine cannot reproduce the functions achievable on a supercomputer *in the same amount of time with the same amount of resources*.

The upshot of the present analysis is the following. On the one hand, since semantic and phenomenological properties depend on informational processing functions that require causal underpinning, it is implausible that systems with vastly different causal capacities per unit time, sharing no structural commonalities, can generate identical semantic and phenomenological properties. Hence, the thin theory of implementation appears infeasible. On the other hand, given a thick theory of implementation, it becomes questionable whether the specific causal underpinnings of human-like consciousness can be multiply realized. We will examine this directly in the next section. For now, the general problem for CTCs is that it is unclear how the real physical causation, which provides the essential semantics, can be assumed to be sufficient to multiply realize properties in systems that are quite obviously asymmetric in respect of not just specific MDPs, but also the number of degrees of freedom instantiated (also an MDP). Assuming that SBSs can generate the requisite causal, and hence semantic properties, is an article of faith that type-physicalists can justifiably argue either, at best, begs the question, or, given the empirical conditions under which awareness is generated (below), is false.

Third, it might be objected that my arguments are orthogonal to Chalmers' key commitments. He claims that implementation properties subvening computational formalisms (specifying causal organization changes) suffice for mentality. He does not say either 1.) how the TCS comports with physicalist approaches as such, or 2.) whether his account portends practical success in generating "artificial consciousness." I am sympathetic to these points, especially the second. Chalmers is a very astute and careful thinker, and has guarded himself well. Perhaps the letter of his message (notwithstanding the spirit) precludes the charge of false promises. Since there is no room to discuss the subtleties of his metaphysical position re: consciousness, it must to suffice to say that granting Chalmers the abovementioned points is tangential to my concerns in this paper. Either his account is a physicalist theory in disguise, in which case it is logically viable but empirically empty vis-à-vis the class of SBSs, or it is legitimately computational, in which case it is false.

Lastly, it might be argued that Chalmers' C_C is immune to the objections I have offered, either because (i) I have given an incorrect account of C_C, and thus my critique does not apply, or (ii) C_C, in virtue of stipulated causal sensitivities, can handle medium-dependent properties. Perhaps, e.g., "electrophysiological properties…are just the sorts of causal relations that, Chalmers argues, his conception of computation can handle…"[11] These objections can best be met by introducing and explaining the central thesis of this paper.

6. The General Critique of Computational Theories of Consciousness

The foregoing discussion of Chalmers' concept of ACO involves various facets of the tension involved in harmonizing the abstract with the concrete. This can be reformulated as a general problem applicable to all CTCs: the inherent gap between medium-independent properties and medium-dependent properties. The general argument against CTCs can be formulated straightforwardly.

(P1*) Computation is medium-independent.

(P2*) Consciousness is medium-dependent.

(P3*) Nothing that is medium-independent suffices for the instantiation of something that is medium-dependent.

[11] Thanks to an anonymous referee for requesting clarification here.

(C1*) No CTC can be true.

Let us look at each step of the argument.

(P1*) P1* asserts that computations are individuated independently of the physical media in which (a particular instance of) implementation occurs. Essentially computational entities, like algorithms and software programs, are medium-independent in that their essential specification is abstract/formal, and does not include specific implementation details.[12] In other words, since there are many physically asymmetrical ways to formally replicate algorithmic descriptions, there are many physically distinct ways to implement a given computation, hence the *de facto* multiple realizability of computation.

But what is it that is to be implemented in a medium-independent manner? The nuts and bolts of computation, so to speak, are *state-transitional sequences* that are *rule-following* (Egan, 1995). State-transitional structure is implied by the Church-Turing thesis, which holds that anything effectively computable can be computed by (and hence is formally equivalent to) a Turing machine (Copeland, 2002). Since Turing machines have state-transitional structures, it follows that computational systems will have state-transitional structures (or their formal equivalent) as well. And rule-following functionality is implied by recursion theory, and is required to permit input-output relation preservation. If the foregoing is correct, then computations are *fundamentally* individuated by three variables: *medium-independence*, *state-transitional structure*, and *rule-following functionality*. Call this the Minimal Theory of Computational Individuation.[13]

---

[12] More accurately, computational implementation involves only *one* medium-dependent (logical) constraint: the medium must have *enough degrees of freedom* to encode the relevant computational descriptions. This has been correctly noted by Piccinini and Scarantino (2011), and is implied in Chalmers (1996; Chapter 9). The importance of this, as we'll see below, follows from the fact that different media have different implementation limitations.

[13] An interesting ongoing debate is whether the rule-following function is necessarily semantic (i.e., whether computations have semantic individuation conditions). "Traditional" approaches have assumed so (e.g., Shagrir, 2006; Fodor, 2008; Sprevak, 2010): they turn on considerations such as, "Without some semantic constraint, how can we know if transitions are *rule*-following?" "In what sense can we meaningfully call something a "rule" unless we have a clear understanding of what makes it so?" Further support for this view comes from the fact that computation is useful only insofar as it is representational/semantic. Conversely, non-semantic accounts hold that representation is unnecessary to the concept of rule-following *as such* (cf. Egan, 1995; Piccinini, 2010b). A syntactic rule, it is argued, still suffices as a rule. These non-semantic accounts hold that the most general account of computational individuation should not advert to representational contents; the nature of the rules and the potential semantic values of states are further matters whose details individuate specific types and tokens of computation, but not computational processes *simpliciter*. One unsolved problem for non-semantic accounts is how to usefully unpack the notion of syntactic rules (ibid.). Is there a viable hybrid account? Obviously, there is no room for a discussion, but perhaps we can distinguish metaphysical and epistemic notions of computation. Metaphysically speaking, computations might be non-semantic, but, given the requirement of cognitive access, we must individuate usable computations semantically. The idea is that if there are non-semantic "syntactic" rules, they are epistemically beyond our ken, and cannot be *classified* as computations *by us*, in any case.

Whether the Minimal Theory is correct, however, is tangential to the truth of P1*. Computations, as abstractly formalized functional descriptions, are not medium-dependent entities. Neither Chalmers, nor any computationalist, denies that. As long as a state-transitional rule-following sequence (of formal description X (mirroring the causal organization of Y, e.g.)) is implemented in a medium-independent manner, a computation (X(Y)) has occurred. It should be clear that Chalmers' C_C, e.g., is a variant of this general account of computational individuation, thus answering objection (i) in § 5 above.[14] P1*, moreover, is consonant with Chalmers' claim that mental properties are organizationally invariant; thus, there is nothing contentious about P1* from a defender of TCS/CTC's point of view.

(P2*) P2* asserts that consciousness is a medium-dependent property. P2 obviously begs the question against defenders of the CTC, and therefore requires independent support. I will offer two arguments, one analytical in nature (A) and the other empirical (B).[15]

(A) The *analytical* argument for the medium-dependence of consciousness comes from an analysis of the nature of experience itself. The basic idea is that consciousness – conscious experience – displays a fundamental unity or holism (see Brook and Raymont (2010) for an overview), which requires us to treat it as fundamentally *concrete* in nature.

One way to argue is as follows. Conscious experience is a *self-evident* "given." Even, *a la* Descartes, if some representational contents of experience are illusory, one cannot doubt one's having an experience *simpliciter*. Experience, moreover, requires an "experiencer": a subject. Subjecthood, furthermore, is individuated by *phenomenal holism* (Bayne and Chalmers, 2003; Alter, 2010; Dainton, 2010), which is the instantiation of a metaphysically real "strong unity:" the individuating condition of "objecthood" (Strawson, 2010). Thus, subjects of experience – given the strong unity evident in phenomenal consciousness – are objects: "metaphysically real concretely existing entit[ies]" (ibid.; p. 88).

Thus, the first part of the analytic argument establishes the corollary theses that (1) consciousness is a metaphysically real, concrete phenomenon and that (2) conscious subjects, given phenomenal holism, are objects. But in virtue of what is something (1) real/concrete and (2) an object? One obvious answer is that something is "real" (i.e., an object) in virtue of its ability to have causal impact on the world (cf.

---

[14] Remember, for Chalmers, if "the formal state-transitional structure of the computation mirrors the causal state-transitional structure of the physical system," then a causally-individuated computation has occurred (Chalmers, 1994; p. 4).

[15] One noteworthy argument I will not *directly* deploy is how a computational theory could account for *emotional* experience, given that emotions are probably the most obvious candidates for MDPs. *Prima facie*, this is a considerable problem for computational theories, considering the qualitative concreteness of emotions and their neurochemical (i.e., plausibly type-physicalist) realization. However, there simply isn't space to open up such a complex and vexing topic. See Charland (1995), Sizer (2000), Nussbaum (2003) and Stuart (2010) for a balanced overview of the issue.

Collier, 2008). Since causal impact obtains in virtue of instantiation-dependent properties (of a system), objects are individuated by their instantiation-dependent properties (IDPs).

What, then, are IDPs? Plausibly, IDPs are constitutive physical properties of systems/objects/processes. Internalist/narrow-content IDPs might be contrasted with externalist/broad-content CDIPs (context-dependent instantiation properties), in that the latter are dispositional properties that depend on environmental contingencies for their realization. But what can we say of the class comprised by IDPs and CDIPs? Given that both classes of properties depend upon the exact nature of actual and potential physical interactions, it seems correct to say that IDPs and CDIPs are subsets of the larger class, MDPs. If so, then IDPs imply MDPs, and the analytic argument for P2* is thus: since concrete entities are concrete in virtue of their MDPs, and consciousness is self-evidently concrete, we can conclude that consciousness is a medium-dependent phenomenon.

Now, Strawson characterizes *functions* as "concretely existing unities" (ibid.; p. 89), and so, returning to our paradigm case, Chalmers' ACO would count as concrete by Strawson's own lights. Hence, Strawson's analysis provides no problem for the TCS or CTCs in general. But I think Strawson is wrong here. Specifically, I think Strawson makes the same mistake Chalmers does – a certain metaphysical gratuity with abstract "objects" (*a la* Plato). Functions may be concrete descriptions, but that is a far cry from amounting to concrete instantiations *in respect of the same properties*. Again, reification fallacies loom.

Nevertheless, the analytic argument given here does not seem to be able to deliver the goods. There are at least two central objections to the analytical argument as presented. One, it can be accused of begging the question in multiple respects; first, in assuming consciousness to be causally efficacious in order to derive that consciousness entails IDPs, and, second, then deducing that phenomenal holism *entails* medium-dependence. This leads to the second problem, which is that, even if the outlines of a valid argument are in view, it is presented far too quickly. The discussion skates over a number of contentious metaphysical and epistemological assumptions, whose utilization requires stronger defense. Given that the empirical argument (B) is much clearer, we will shortly turn to it in lieu of trying to navigate these murkier waters.

Before moving on, however, I think that the analytical argument does bring up an interesting challenge for CTCs. Any complete theory of the mental must be able to explain the phenomenological holism each of us, as subjects, experiences. If computations generate mental states, then coherence between mental states would seem to require coherence between computations. But by what mechanism will all these computations cohere? In other words, by what resources can a CTC account for the meta-coherence that must obtain among computations to instantiate phenomenal holism?

Say that computations are mutually coherent if it is possible to *simultaneously implement* them in the *same* spatiotemporal system. Two routes of response seem open to the defender of a CTC. One route is to argue that *there is only one computation with mental properties*. There would of course be no possible conflict between mental states in such a case, as there would only be one complex computation in question. However, there are no reasons to think such a case is possible, let alone plausible, and, given the self-evident diversity of mental contents within conscious experience, assuming that there is only a single computation might seem both unlikely and at odds with the practices of typical defenders of the computational theory of mind (cf. Fodor, 2008; and see Horst (2009) for an overview).

The other option for a defender of a CTC is to argue that the set of simultaneously implemented "mental" computations will be mutually physically coherent. The question is how to explain, in a non ad-hoc manner, how a theory of computation can generate the requisite *simultaneous implementability that must obtain between vast numbers of physically-realized computations*. In the brain, the nested, hierarchical structure – evolutionarily "designed" over millions of years – assures a tendency towards massive mutual coherence (Simon, 1973; Werner, 2010). In silicon-based systems (SBSs), such massive mutual coherence must be designed top-down. Given the complexity of human consciousness, it is a fair question whether SBSs displaying the requisite computational complexity to realize human-like consciousness can be built.

Of course, defenders of CTCs seem free to respond that the foregoing concerns are practical ones, and are therefore orthogonal to their theoretical claims. This is a fair reply. And, insofar as such considerations about physical implementation constraints are potentially decisive, they will reappear in the context of the empirical argument. Let us turn to it.

(B) The *empirical* argument for P2 involves the insight that, epistemically speaking, it is an *a posteriori* matter what physical details are functionally relevant to the generation of consciousness (Levin, 2008; Hill, 2009; McLaughlin, 2010; Weisberg, forthcoming). I want to examine two special structural properties whose instantiation arguably realizes/subvenes consciousness: our electrodynamical (B1) and our fractal (B2) nature.

(B1) In examining the correlates of consciousness, one fact is uncontested: consciousness varies in accordance with electroneurodynamical properties. But which electroneurodynamical aspects are irrelevant to the generation of consciousness is unknown. Historically speaking, part of the motivation for positing that the mind can be understood computationally is that neurons function, in part, in a discrete (all-or-none firing) fashion, analogous to the digital computation characteristic of typical computers (McCulloch and Pitts, 1943). But it is an open question whether the brain generates mental states by such discrete operations alone.

It has long been established that the electrical fields in the brain affect neural firing, which is unsurprising as electrical fields move electrically-sensitive ions (whose movement through neural membranes realizes both neural spikes and graded potentials) (Pockett, 2000). Since neural firing, by inference to the best explanation, realizes (or at least subvenes) consciousness in humans, it is natural to speculate that the electrodynamical field generated in the brain generates consciousness-changing causal effects (ibid.; John, 2001; McFadden, 2002; Frolich and McCormick, 2010).[16] Key evidence comes from the use of repetitive Transcranial Magnetic Stimulation, in which the application of electromagnetic fields has been shown to directly change conscious experience (Overgaard et al., 2004; Babiloni et al., 2007). If this is correct, then consciousness (at least partially) depends on the instantiated *electrodynamical field properties* (EFPs) of the brain (cf. Fingelkurts et al., 2010). It is in this sense, as I will argue, that consciousness is inherently medium-dependent. EFPs are crucially *inseparable from the media* they constitute because EFPs have holistic/global properties, such as nonlocality (Belot, 1998; Frisch, 2002) and unified physical continuity (Lipkind, 2005). Given this, EFPs are necessarily MDPs. Thus, if consciousness depends upon EFPs, as neuroscience suggests, then consciousness depends upon MDPs, and P2* is true.

What, though, are EFPs, and why should they be considered to be MDPs? In a review of field properties, Michael Lipkind notes the following individuating conditions of fields: physicality, causal nonlocality, continuity and sensitivity to internal parameters (ibid.). These properties fall out of the fact that EFPs are electrodynamic physical forces, containing energy and momentum, applicable to each intra-field "point." Electrodynamical fields thereby affect the behavior of electro-responsive "particles" within them, and are thus causally efficacious. Because fields affect all points by a given quantity (at a particular time), they possess holistic/global properties, like nonlocality and continuity (Belot, 1998; Frisch, 2002). Accordingly, accounting for EFPs requires instantiating holistic/global causal properties. Given that the individuating conditions for EFPs and computations are largely antithetical, the worry for CTCs is how EFPs are to be realized in computational processes.

There are two aspects of computation to consider: their abstract/formal characterization and their implementation conditions. The problem, as we've seen, is this: on the one hand, it is unclear how *real field causality* can be captured abstractly, in a medium-independent state transitional sequence. An abstract formalism is not the kind of entity that generates real causal properties. And, on the other hand, adverting to implementation properties doesn't seem to help either: If computational systems are required to implement EFP-subvening structures, then physical differences between SBSs and humans make such implementation highly implausible.

[16] Such a conclusion is unsurprising since living systems depend on electrodynamical activity in manifold ways and on multiple spatiotemporal scales (Adey, 1984; West, 2006). See, e.g., the abstracts from the latest (2011) Electrodynamic Activity of Living Cells conference (online at http://edalc11.ufe.cz/doku.php).

It is a question of mechanical dynamics: whether it is possible to build SBSs that can implement the relevant field properties required to realize the "real physical causality" that generates the necessary semantic properties (at least partially) constitutive of consciousness. The problem with implementing the kind of EFPs *generated in organic brains* – the *only* structures whose capacity to realize consciousness is undisputed – is that SBSs, as presently conceived, are designed in a manner *antithetical* to that of natural systems. Specifically, SBSs are designed to *minimize EFPs*, so as to maximize component independence and to minimize signal noise, while the components of organic systems are inherently interdependent, given that mutual functionality is a prerequisite for natural selection (cf. Simon, 1973; West, 2006; Werner, 2010). More specifically, electrodynamics in organic systems are generated in an aqueous environment that serves as a conduction-enhancing substrate, while SBSs have no analogous field-ground (Giertz, 2010). This is nowhere more true than it is in regards to the brain's intrinsic electrodynamical field and its local field potentials, which have been argued to be non-instantiable in SBSs as currently conceived (Pockett, 2000).

The basic argument of (B1) can be easily summarized. As revealed by empirical experiment, consciousness depends on EFPs. Field properties involve holistic physical forces. Such properties are medium-dependent in that they sensitively depend upon the exact media in which they occur. Given that computations are medium-independent, such field properties could only be realized by implementation properties. However, due to antithetical design differences, EFP-relevant physical differences preclude SBSs (as currently conceived) from realizing human-like EFPs. Yet the hopes for CTCs are actually even worse, since there is another consciousness-subvening feature that greatly deepens the difficulty here discussed. It is this: organic systems are inherently fractal, while SBSs are inherently Euclidian.

(B2) A profound difference between ourselves and any SBS is that, due to our evolutionary design, we are *intrinsically fractal, both structurally and functionally* (cf. Iannaccone and Khokha, 1996; West, 2006; Aguirre et al., 2009; Werner, 2010).[17] SBSs are not.[18] I will argue that this principal design asymmetry is decisive in

[17] We are fractal in many ways. A few structural examples (and see West et al. (1999)): we grow through fractal self-subdivision (i.e,. cell division). Each proper part of us is self-similar (i.e, each cell contains the same DNA). Larger structures in us are structurally fractal (lungs, circulatory system, central nervous system (Werner, 2010), etc.). Lastly, our overall structure is a nested hierarchy, e.g., (body (organs (tissues (cells (DNA))))). For sensorimotor examples, see Kelso (1995) and West (2006). For cognitive examples, see MacCormac and Stamenov (1996) and Werner (2010).

[18] This is not to say that silicon-based systems cannot be designed with some fractal structure. In fact, engineers have already created low-dimensional fractal antennae. The problem, for CTC advocates, however, is that there is a profound and categorical complexity gulf between the fractal dimensionality of organic and inorganic systems, on the one hand, and between organic, bottom-up processes and inorganic, top-down ones (Collier, 2008). Simply put, these differences stem from the fact that *organic systems, by dint of their self-organizing nature, continually increase their fractal dimensionality across many orders of magnitude as they grow*. In the brain, for example, cognitive development is underpinned by the fractal self-organization of neural networks (Sporns, 2011). *SBSs*

refuting CTCs, for it betokens an unbridgeable implementation gulf, which implies an unbridgeable semantic/intentional gap (which implies an unbridgeable phenomenological gap).

The key fractal (F) argument is the following:

(FP1) Consciousness depends upon fractal functional properties (cf. MacCormac and Stamenov, 1996; Gong and Leeuwen, 2002; West, 2006; Werner, 2010; Sporns, 2011).

(FP2) Fractal functional properties are generated by energy throughput within fractal structures (Michelitsch et al., 2009).

(FL1) Consciousness depends upon energy throughput within fractal structures.

(FP3) Energy throughput within fractal structures is a MDP (ibid.).

(FC1) Consciousness depends on at least one MDP.[19]

Numerous studies have shown fractal functional properties in consciousness data, suggesting that they are a dynamic correlate of consciousness (MacCormac and Stamenov, 1996; Werner, 2010). Fractal functional properties include scale-invariant power-law dynamics (such as statistical self-similarity across a wide variety of scales), hierarchical nesting dynamics, chaotic bifurcation dynamics, and nonlocality (Aguirre et al., 2009; Michelitsch et al., 2009).[20]

It has been mathematically demonstrated that fractal functional properties, such as nonlocal coherence between system elements, are generated by energy throughput within fractal structures (Michelitsch et al., 2009). Given this, realizing such *real* fractal functional properties requires implementing *real* electrodynamically-sensitive fractal structures (cf. Bieberich, 2002).

A paradigmatic electrodynamically-sensitive fractal structure is that of the human CNS (ibid., Werner, 2010). In the human CNS, with the proper energy throughput – the electrodynamics of which are at least partially regulated by neurochemical circuits (Rose, 2006) – consciousness arises. There is strong evidence that human CNSs generate fractal functional properties in proportion to the complexity of their structural fractal dimensionality (Honey et al., 2010; Sporns,

---

have no equivalent capacity for novel hardware complexification, and this underpins the gulf in implementation properties that explain the failure of CTCs.

[19] This suggests the following *Fractal Neurodynamics Theorem*: The degree of consciousness is given by the fractal dimensionality of the transiently sustainable electrodynamical energy throughput.

[20] These properties can be speculated to underlie numerous phenomena associated with consciousness, such as the self-reflexivity and unity of consciousness, on the phenomenological side, and the remarkable *zero-lag* synchronization (Pockett, 2000; John, 2001) among distant parts of the central nervous system (CNS) that appears to generate conscious experience, on the functional side.

2011).[21] Accordingly, it appears reasonable to infer that the realization of consciousness requires the generation of *real* highly complex fractal functional properties (cf. Bieberich, 2002; Walling and Hicks, 2003; Werner, 2010).

However, given that SBSs are not constructed from self-similar iterative processes on multiple scales (as occurs in organic systems through complex, self-organization processes), it is arguable that complex organic systems are of categorically higher fractal dimensionality than SBSs.[22] It is also plausible that the highest electrodynamical fractal structures are those comprising parts of the CNS (Walling and Hicks, 2003). Given these considerations, there are very good reasons to believe that SBSs cannot instantiate the highest fractal functional properties of the human CNS.[23] Therefore, if adult human consciousness is generated by the highest class of fractal functional properties of the human CNS (cf. Pocket, 2000; Bieberich, 2002; Werner, 2010), then *CTCs will fail on account of there being no viable implementation substrates that are complex enough* (to realize electrodynamically-sensitive structural properties of sufficient fractal dimensionality).

The basic argument of (B2) can be easily summarized. As revealed by empirical experiment, consciousness depends on fractal functional properties. As shown by mathematical demonstration, fractal functional properties emerge from fractal structural properties. Accordingly, computations would have to be implemented with the same relevant fractal structure as obtain in humans to generate the same consciousness-realizing fractal properties. However, while complex, non-equilibrium processes naturally generate fractal structures of variable complexity, equilibrium artifacts, such as SBSs, do not. Thus, there is a categorical fractal dimensionality asymmetry that grounds a fundamental implementation/realization asymmetry. It is this asymmetry that invalidates CTCs.

In sum, as the arguments for (B) show, empirical evidence suggests that consciousness depends on MDPs that cannot be implemented in SBSs. More specifically, (B) argues that consciousness depends upon fractal and electrodynamical field properties, both of which place severe medium-dependent constraints on the realization of consciousness, constraints that can not be met in SBSs. Having derived P2*, let us turn to P3*.

(P3*) P3* holds that entities characterized by MIPs can not instantiate entities characterized by MDPs. This is a very intuitive premise. It holds, e.g., that math equations explaining sound waves are not, themselves, audible – or that descriptions of buoyancy do not confer buoyancy thereby.

---

[21] Evidence shows that fractal dimensionality increases in healthy growth and declines in infirmity, disease and old age (West, 2006).

[22] Organic systems are those in which real fractal functionality derives from real fractal structural properties changing over time. This distinction is important for separating causally real fractal systems from causally-inert mathematical objects (e.g., Sierpinski gasket) or man-made 2D fractal antennae.

[23] This is unsurprising given the antithetical (artificial vs. natural selection) design properties. A categorical fractal dimensionality difference is expected given the self-organizing nature of organic, but not inorganic, systems.

The most important issue for our purposes is whether a fundamentally medium-independent process like computation can realize medium-dependent properties. There seem to be two possibilities. (1) On the one hand, it may be that computational processes can only realize MIPs. (2) On the other hand, it may be that computational processes can realize MDPs in virtue of their other properties. *Prima facie*, it is unclear how rule-following functionality or state-transitional "structure" could permit the realization of MDPs. This is perhaps most clearly appreciated in the context of mathematics. Mathematical properties generally imply rule-following functionality and state-transitional structure, but, on their own, fail to instantiate real causal properties. Thus, such computational properties can not be expected to generate MDPs. One cogent reply, though, is that computations can realize MDPs in virtue of implementation properties. In fact, as we've seen, this is exactly how Chalmers argues that computations can have semantic properties.

Obviously, option (1) is unattractive for computationalists, as it would severely limit what computational processes can be expected to do. Certainly, option (2) provides much more wiggle room. The problem with the second reply, however, is that it gives up the ship to a type-physicalist characterization. This is because it grants computations the power to realize MDPs *only in virtue of the physical properties* of their implementations. Accordingly, if consciousness is medium-dependent (i.e., constituted by MDPs), computations will only be able to realize consciousness in virtue of the physical properties of their implementations, which is certainly not an *abstract* individuation condition.

(C1*) C1* asserts that no CTC can be true. The argument is straightforward. Computation, as such, is medium-independent, because many structures can implement the same state-transitional sequence. Consciousness, moreover, is arguably medium-dependent, on account of the electrodynamical field and fractal properties that generate it. Accordingly, if computational processes are to generate consciousness, they must generate the requisite medium-dependent properties in virtue of implementation properties. However, the specific structural properties implicated in human consciousness do not appear to be potential substrates of non-biological implementations. Therefore, the medium-independence individuation condition of CTCs cannot be met, and no CTC can succeed.

7. General Argument Against CTCs: Objections and Replies

Several objections can be leveled against such a broad conclusion. Perhaps most obviously, the evidence used herein might be disputed, or at least questioned.

However, there is nothing inherently contentious or surprising about the empirical claims involved. Natural processes aggregate by fractal dynamics (Iannaccone and Khokha, 1996; Aguirre et al., 2009, etc.), and we are natural processes *par excellence*. Given that functions supervene on structures (Honey et al., 2010), it is natural to reason that as the fractal dimensionality of a structure increases, so too does the functional complexity the structure is capable of subvening (Waller and Hicks, 2003). Accordingly, it is no surprise that our mental lives should be made possible by highly fractal central nervous system properties (Werner, 2010; Sporns, 2011). In addition, living organisms depend on electrodynamics for most biological activities (Adey, 1984). Unquestionably, the brain is the predominant electrodynamical organ. Hence, it is unsurprising that our mental lives would depend on electrodynamical states, such as EFPs (Pockett, 2000). So, while one might question the empirical claims, plausibility is in their favor.

A second question is that raised as (ii) in § 5 (above): Perhaps, e.g., "electrophysiological properties…are just the sorts of causal relations that…[a C_C]…can handle…" That is, why can't a CTC account for MDPs in virtue of a causal theory of individuation? Why can't the causal effects of fractal and electrodynamical organization be "written into" the computational description? There are several relevant considerations.

As we've seen, for Chalmers, a viable computational account rests on producing the "same pattern of causal interactions" (Chalmers, 1994; 1996; p. 248). The short answer to the above question, then, is that the actual pattern of causal interactions in neurons depends on their omni-electro-responsiveness. This means that patterns of causation between neurons involve the continuous fluctuation of EM fields in which they are embedded (Pockett, 2000; Rose, 2006). Not only are the ions within and around neurons sensitive to (and move due to) electrical fields, the graded potentials within changing neural membranes are near-continuously sensitive to changes in the electrical properties of extracellular matrix (ibid.). The EM properties of this space (and hence the causal organization of the neurons) are also largely affected by glial cells, which keep the EM field at the edge of responsiveness by absorbing and releasing neurochemicals (Fields, 2009). Thus, this extracellular space provides a common causal influence that is not represented in the discrete modeling of any given neuron. Accordingly, part of our functional organization is constituted by field properties that cannot be reduced to the sum of discrete states of neurons. These causally efficacious fields are (computationally) emergent properties.

Since the foregoing neurobiological explanation might be accused of begging the question, let's approach the question from a computational point of view, returning to Chalmers' paradigm theory to consider what he writes about the implementation conditions for a combinatorial state automata (CSA), the key computational system for the TCS: A "system implements a given CSA if there exists…a vectorization of states of the system, and a mapping from elements of those vectors onto corresponding elements of the vectors of the CSA, such that the state-transition relations are isomorphic in the obvious way." (Chalmers, 1994; p. 5) In general, the key problems with implementing nonlocal fractal and field properties are that there is no way to capture such physically holistic properties in a system 1.) whose essential syntax is in terms of system sub-states (i.e. parts), and 2.) which changes in a local, discrete and sequential manner, as opposed to a nonlocal, global and holistic one.

But what about the possibility of writing a computation such that each state change introduces changes in, potentially, all other variables? It might be argued that this would properly recreate the "global" state effects argued to be essential.

A first, epistemic, difficulty that arises is "How do we know what are the relevant physical states?" By what physical taxonomy might we properly carve up our own structural underpinnings? A second, further, problem is "How do we know, given the chaotic dynamics that are well-established aspects of our neurodynamics, what the relevant causal transitions are?" It is unclear that such obstacles are surmountable. For the sake of argument, though, let us imagine such difficulties can be overcome.

A third, programming, problem awaits. How could we create programs that contained the vast numbers of states that must be included? Mental states are often contextually cued; accordingly, all the potential environments that might be encountered would have to be programmed, along with all the appropriate, but contingent, responses. Note that it is quite unclear how environments and their dispositional architectures should be programmed. Additionally, these contingent responses would have to be written in a non-transitive manner, such that the order of experiences would be relevant to state determination. Again, the programming burdens increase.

It is easy to underestimate the difficulty of the task implicated. Even a single dendritic spine is the size of an amoeba, which contains approximately a million chemicals in semi-ordered configurations (Rose, 2006). The activity of individual dendrites influences graded potentials, which themselves influence action potentials, which are widely assumed to be predominant, though not exclusive, mental causes. There are over 100 billion neurons in the human CNS, some of which have thousands of sites of active chemical transmission with upwards of thousands of other neurons. The number of potential state combinations and permutations quickly become mathematically intractable (NP-Hard) for the calculation of the diachronic development of such a network. Now, even if such a program could be written, what machine is powerful enough to store all the contingent states and effectively process the program? Given the complexity of the causal organization of the CNS, combinatorial explosion problems seem insurmountable.

Of course, if nonlocal processing is required to instantiate our mental states, as Jerry Fodor (2008) suggests, then Chalmers' discrete and discontinuous TCS would fail on those grounds. But let us bracket this problem, granting, for the sake of argument, that either Fodor is wrong in his speculation or that some remedy might be found – if not one to which the TCS has access, then one that other CTCs might utilize.

Bracketing all the aforementioned problems, might a CTC advocate claim that nonlocal and global properties can be written into computational descriptions? It seems not. There still remains what we can call the *semantic* problem. In order to avoid reification fallacies, computationalists must hold that computations need to be implemented in order to generate real (non-abstract) properties. The implementation properties generate the real physical causality that provides the semantic properties necessary for realizing real mental states. But here we have the introduction of physical properties. The necessary semantics are generated by real forces; moreover, the semantic properties depend upon the relevant physical properties. But this means that simply modeling, and not instantiating, certain physical properties (nonlocality, holism) will fail to generate the necessary semantics, just as modeling a storm will fail to generate rain. The point of the implementation requirements is that the relevant properties must be instantiated, not merely described. This is why nonlocal properties, as arise from field and fractal dynamics in the brain, cannot be merely written into computational descriptions. They would thereby fail to have the semantic properties that make them important in the first place.

A third question concerns the nature of the implementation dynamics required. Would not a form of analog computation suffice to account for the implementation properties required to realize consciousness?

Let us first note that this strategy is unavailable for Chalmers (1994), who specifically stipulates a discrete and discontinuous account of computation (section 3.4). Chalmers argues that analog computation is unnecessary because continuous and chaotic dynamics will wash out, at a basal enough level, as *noise*. Moreover, since whatever turns out to be the basal functional level will be describable with a fine-enough grain of analysis, such continuous and chaotic dynamics will fail to "add anything new." This strategy runs into the following problems. One, defining neuro-activity as "noise" must post-date, rather than precede, a theory of neurodynamics, which is something we don't yet have (cf. Werner, 2010; p. 14). It is a somewhat empty gesture to turn an argument upon "neural noise" *sans* a theory of neurodynamics. Two, assuming that complex and chaotic dynamics cannot "add anything new" is a very ironic reply, given that epistemic emergence is a *sine qua non* of complex dynamics in general. That is, the one prediction one can make about complex systems (with a probability that rises in proportion to system complexity) is unpredictability. And if you can be sure you can't predict the behavior of a system, it rings hollow to claim *a priori* that the system will not generate novel behaviors. Lastly, and decisively, discrete and discontinuous computational processes cannot realize nonlocal properties, as we've seen.

But what of analog CTCs? Can they achieve what digital CTCs cannot? Analog computation can be defined as that which takes continuously varying, and not discrete, variables as its inputs. Such inputs, however, are still manipulated in an essentially medium-independent state-transitional rule-following manner. In order to qualify as a computation, there must be unambiguous input-output transitions. This, however, requires unambiguous input identification. Accordingly, the analog nature of the input is partial and implementation-dependent (in that greater input specificity requires more complex and resource-expensive hardware).

Now, in restricting input, part of the causal organization is discarded. However, for certain properties – archetypically, for MDPs – holistic causal preservation is an individuating condition. Thus, for archetypal MDPs, such as consciousness, analog computation fails to instantiate holistic causal organization. Given these constraints, it can be seen that analog CTCs fare no better. Their medium-independence causes them to suffer from the same problems as digital CTCs.

Lastly, it might be argued that novel engineering progress might permit the generation of non-biological machines capable of consciousness. There are two main considerations. First, novel machines might be of a heretofore unknown structure and generate heretofore unknown functions. My arguments cannot be weighed against this possibility; to do so would commit the fallacy of ignorance. Second, it might be thought that novel SBSs are possible that might surmount the limitations I have enumerated. However, I maintain that there are no good reasons to hold that non-biological systems (systems without biological systems as proper parts) can generate the requisite structural and functional properties. Why? Very simply: the antithetical design principles by which living and non-living systems are generated imply profound structural and functional asymmetries. As I have argued, the most relevant asymmetries for the question of human-like consciousness are electrodynamical field and fractal structural properties.

In sum, I have argued that defenders of CTCs face a dilemma. If they attempt to eschew computational implementation, then, given the essentially MIP-nature of computations, computations will fail to generate MDPs, which is problematic given the strong reasons to think semantic content (*a fortiori*, consciousness) depends on MDPs. This forces advocates of CTCs to claim that computations can generate MDPs in virtue of implementation properties. This latter strategy, though, accedes that MDPs obtain only in virtue of real physical causation. *Given the empirical evidence that fractal structure and electrodynamical fields are causally efficacious in the generation of mental properties, the inability of silicon-based computational systems to realize the fractal electrodynamical properties of organic brains, due to fundamental design asymmetries between natural-evolutionary and artificial processes, implies no CTC can be true*.

8. CTCs and the HMSMRST

If the foregoing is correct, then no CTC can be correct. What, though, is the relation of computational processes and silicon-based systems (SBSs)?

As denoted here, SBSs refer to artifact-machines comprised of silicon and other non-living materials.[24] Computers are the archetypal SBSs; certainly they are

[24] A caveat. My argument is tangential to the possibilities of natural-artificial system hybrids (cyborgs, etc.), which have natural systems as proper parts. My argument also only holds against presently conceivable silicon-based systems. (I am agnostic about what might be possible with "high-powered" quantum computers (which are presently of negligible computing power (8 qubits?) due to physical

the most sophisticated SBSs known. Given the self-evident complexity of conscious experience, it is reasonable to expect that only the most complex machines might multiply realize (MR) it. Accordingly, computers have been the paradigm candidates for subjects of artificial consciousness/experience.

As currently conceived, all of the most complex electrodynamical SBSs can be classified as one type or another of computer (or computer-like system). Tautologically, computers instantiate computations. As we have seen, computations are insufficient for the instantiation of consciousness. Hence, computers are insufficient for the instantiation of consciousness. Since there are no candidates for complex SBSs that are not computers in one guise or another, it is reasonable to generalize the arguments against computers to SBSs.

Thus, given the assumption that only complex SBSs could even potentially MR human-like consciousness, and that computers comprise the class of the most complex SBSs, the present argument suggests that SBSs cannot MR human mental states. That is, the HMSMRST is false.

The general argument of this section is as follows:

(C1*) All CTCs are false (§ 6).

(P3*) The only plausible realization of the HMSMRST is via a CTC.

(C2*) There is no plausible realization of the HMSMRST.

(C3*) The HMSMRST is (almost certainly) false.

9. Conclusion

With the goal of shedding some light on the question of the nature of the relationship between consciousness and computation, I assessed David J. Chalmers' paradigm computational theory of consciousness, his thesis of computational sufficiency, in which the capacity of computations to capture abstract causal organization, paired with the organizational invariance of mental states, implies that the right kind of computation is sufficient for possession of a (conscious) mind.

I argued that the TCS was undermined by a critically insufficient deployment of the concept of abstract causal organization. The key problem was seen to be a dilemma arising from the tension between the concepts of "abstraction" and "causality." It forces computational approaches to depend too heavily on implementation properties, which are physical properties.

---

implementation limitations). I am not agnostic about whether such are possible: I wager the implementation limitations to engineering computationally powerful quantum computers are decisive.)

Generalizing this problem, I introduced a general argument against computational approaches to consciousness based on the medium-independent nature of computation. Specifically, I argued that consciousness is a medium-dependent phenomenon generated by medium-dependent fractal and electrodynamical field properties.

Computational systems, I argued, cannot realize human-like medium-dependent fractal and electrodynamical field properties in silicon implementation structures and so cannot generate conscious states/processes/experiences.

Finally, I argued that the failure of computational theories of consciousness (CTCs) suggests that the Human Mental State Multiple Realizability in Silicon Thesis (HMSMRST) is almost certainly false.[25]

References

Adey, W. (1984). Nonlinear Electrodynamics in Biological Systems. New York, NY: Springer.

Aguirre, J., Viana, R. and Sanjuan, M. (2009). Fractal Structures in Nonlinear Dynamics. *Reviews of Modern Physics 81*: 333-387.

Alter, T. (2010). A Defense of the Necessary Unity of Phenomenal Consciousness. *Pacific Philosophical Quarterly 91*: 19-37.

Armstrong, D. (1968). *A Materialist Theory of the Mind*. London: Routledge and K. Paul.

Babiloni, C., Vecchio, F., Rossi, S., DeCapua, A., Bartalini, S., Ulivelli, M. and Rossini, P. (2007) Human Ventral Parietal Cortex Plays a Functional Role on Visuospatial Attention and Primary Consciousness. A Repetitive Transcranial Magnetic Stimulation Study. *Cerebral Cortex 17* (6): 1486-1492.

Ballard, D. (2010). Computational Consciousness. In Gangopadhyay, N., Madary, M, Spicer, F. (eds.): *Perception, action, and consciousness: Sensorimotor dynamics and two visual systems*. Oxford: Oxford University Press, 2010.

Bayne, T. and Chalmers, D. (2003). What is the Unity of Consciousness? In Cleeremans (ed.), *The Unity of Consciousness*. Oxford: Oxford University Press.

Belot, G. (1998). Understanding Electromagnetism. *Brit. J. Phil. Sci. 49*: 531-555.

[25] I want to thank all those who gave feedback on versions of this paper, including two anonymous referees, Berit Brogaard, Gualtiero Piccinini, George Terzis, John Brunero and David Brown. Thanks, as well, to David Chalmers and Thomas Michelitsch, who answered questions in correspondence.

Bieberich, E. (2002). Recurrent Fractal Neural Networks: A Strategy for the Exchange of Local and Global Information Processing in the Brain. *BioSystems 66*: 145-164.

Blackmore, S. (2011). *Consciousness: An Introduction*. Oxford: Oxford University Press.

Block, N. (1978). Troubles with Functionalism. *Minnesota Studies in the Philosophy of Science 9*: 261-325.

Bostrom, N. (2006). Quantity of Experience: Brain-Duplication and Degrees of Consciousness. *Minds and Machines 1*6 (2): 185-200.

Brook, A. and Raymont, P. (2010). The Unity of Consciousness. *Stanford Encyclopedia of Consciousness*. Online at http://plato.stanford.edu/entries/consciousness-unity/

Chalmers, D. (1994). *A Computational Foundation for the Study of Cognition*. Unpublished paper available at http://consc.net/papers/computation.ps

Chalmers, D. (1996). *The Conscious Mind: In Search of a Fundamental Theory*. New York, NY: Oxford University Press

Chalmers, D. (1999). Materialism and the Metaphysics of Modality. *Philosophy and Phenomenological Research 59*: 473-493.

Chalmers, D. (2003). Consciousness and its Place in Nature. In Stich and Warfield (eds.), *Blackwell Guide to Philosophy of Mind*. Malden, MA: Blackwell Publishing.

Chalmers, D. (n.d.). *Responses to articles on my work*. Online at http://consc.net/responses.html

Charland, L. (1995). Emotion as a Natural Kind: Towards a Computational Foundation for Emotion Theory. *Philosophical Psychology 8* (1): 59-85.

Chella, A. and R. Manzotti, (2011), "Artificial Consciousness" in V. Cutsuridis, A. Hussain and J. G. Taylor, Eds, *Perception-Action Cycle: Models, Architectures, and Hardware*, Dordrecht, Springer**:** 637-671.

Cleermans, A. (2005). Computational correlates of consciousness. *Progress in Brain Research, 150:* 81-98.

Collier, J. (2008). A Dynamical Account of Emergence. *Cybernetics and Human Knowing 15* (3-4): 75-86.

Copeland, J. (2002). The Church-Turing Thesis. *Stanford Encyclopedia of Philosophy*. Online at http://plato.stanford.edu/entries/church-turing/

Dainton, B. (2010). Phenomenal Holism. *Royal Institute of Philosophy Supplement 67*: 113 -139.

Egan, F. (1995). Computation and Content. *The Philosophical Review 104* (2): 181-203.

Eliasmith, C. (2002). The Myth of the Turing Machine: The Failings of Functionalism and Related Theses. *Journal of Experimental and Theoretical Artificial Intelligence 14*: 1-8.

Fields, R.D. (2009). *The Other Brain*. New York, NY: Simon and Schuster.

Fingelkurts, A., Fingelkurts, A. and Neves, C. (2010). "Machine" Consciousness and "Artificial" Thought. *Brain Research* online.

Fodor, J. (2008). *LOT 2: The Language of Thought Revisited*. Oxford, U.K.: Oxford University Press.

Frisch, M. (2002). Non-locality in Classical Electrodynamics. *Brit. J. Phil. Sci 53*: 1-19.

Frolich, F. and McCormick, D. (2010). Endogenous Electrical Fields May Guide Neocortical Network Activity. *Neuron 67*: 129–143.

Gabriel, C., Gabriel, S. and Courthout, E., (1996). The dielectric properties of biological tissues: I. Literature survey. Phys. Med. Biol. 41: 2231–2249.

Gong, P. and Leeuwen, C. (2002). Scale-Invariant Fluctuations of the Dynamical Synchronization in Human Brain Electrical Activity. *Neurosci. Lett. 336*: 33-36.

Hill, C. (2009). *Consciousness*. NY, New York: Cambridge University Press.

Honey, C., Thivierge, JP and Sporns, O. (2010). Can Structure Predict Function in the Human Brain? *NeuroImage 52*: 766-776.

Horst, S. (2009). The Computational Theory of Mind. *Stanford Encyclopedia of Philosophy*. Online at http://plato.stanford.edu/entries/computational-mind/

Iannaccone, P. and Khokha, M. (eds.) (1996). *Fractal Geometry in Biological Systems: An Analytical Approach*. Boca Raton, FL: CRC Press.

John, E.R. (2001). A Field Theory of Consciousness. *Consciousness and Cognition 10*: 184-213.

Kary, M. and Mahner, M. (2002). How Would You Know if You Synthesized a Thinking Thing? *Minds and Machines 12*: 61-86.

Kauffmann, S. (2002). *Investigations*. Oxford: Oxford University Press.

Kelso, J.A. (1995). *Dynamic Patterns: The Self-Organization of Brain and Behavior*. Cambridge, MA: MIT Press.

Levin, J. (2008). Taking Type-B Materialism Seriously. *Mind & Language 23 (4)*: 402-425.

Lewis, D. (1972). Psychophysical and Theoretical Identifications. *Australasian Journal of Philosophy 50*: 249-258.

Lipkind, M. (2005). Fields in Current Models of Consciousness: A Tool for Saving the Hard Problem? *Minds and Matter 3* (2): 29-85.

MacCormac, E. and Stamenov, M. (eds.) (1996). *Fractals of Brain, Fractals of Mind*. Philadelphia, PA: John Benjamins.

Mathis, D. and Mozer M. (1996). Conscious and Unconscious Perception: A Computational Theory. In Cottrell, G. (ed.) *Proceedings of the Eighteenth Annual Conference of the Cognitive Science Society*: 324-328. Hillsdale, NJ: Erlbaum.

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

McFadden, J.J. (2002). The Conscious Electromagnetic Field Theory: the Hard Problem Made Easy. *Journal of Consciousness Studies 9* (8): 45-60.

McLaughlin, B. (2003). A Naturalist-Phenomenal Realist Response to Block's Harder Problem. *Philosophical Issues 13*: 163-204.

McLaughlin, B. (2010). Consciousness, Type Physicalism, and Inference to the Best Explanation. *Philosophical Issues 20*: 266-304.

Michelitsch, T., Maugin, G., Nicolleau, F., Nowakowski, A. and Derogar S. (2009). Dispersion Relations and Wave Operators in Self-Similar Quasi-Continuous Linear Chains. *Physical Review E 80*, 011135.

Nussbaum, C. (2003). Another Look at Functionalism and the Emotions. *Brain and Mind 4*: 353-383.

Overgaard, M., Nielsen, J. and Fuglsang-Frederiksen, A. (2004). A TMS Study of the Ventral Projections from V1 with Implications for the Finding of Neural Correlates of Consciousness. *Brain and Cognition* 54 (1):58-64.

Pelzcar, M. (2008). On an Argument for Functional Invariance. *Minds and Machines 18* (3).

Piccinini, G. (2010a). The Mind as Neural Software? Understanding Functionalism, Computationalism, and Computational Functionalism. *Philosophy and Phenomenological Research 81* (2): 269-311.

Piccinini, G. (2010b). Computation in Physical Systems. *Stanford Encyclopedia of Philosophy*. Online at http://plato.stanford.edu/entries/computation-physicalsystems/

Piccinini, G. and Scarantino, A. (2011). Information Processing, Computation, and Cognition. *Journal of Biological Physics 37* (1): 1-38.

Pockett, S. (2000). *The Nature of Consciousness: A Hypothesis*. San Jose, CA: Writer's Club Press.

Polger, T. (2011). Are Sensations Still Brain Processes? *Philosophical Psychology 24* (1): 1-21.

Rolls, E. (2007). A computational neuroscience approach to consciousness. *Neural Networks* 20 (9):962-82.

Rose, D. (2006). *Consciousness: Philosophical, Psychological and Neural Theories*. Oxford, UK: Oxford University Press.

Scheutz, M. (2001). Computational versus Causal Complexity. *Minds and Machines 11*: 543-546.

Schweizer, P. (2002). Consciousness and Computation. *Minds and Machines 12*: 143-144.

Seager, W. (1995). Consciousness, Information and Panpsychism. *Journal of Consciousness Studies 2* (3): 272-288.

Searle, J. (1980). Minds, Brains and Programs. *Behavioral and Brain Sciences 3*: 417-457.

Searle, J. (2004). *Mind: A Brief Introduction*. New York, NY: Oxford University Press.

Shagrir, O. (2006). Why We View the Brain as A Computer. *Synthese 153*: 393-416.

Shoemaker, S. (2007). *Physical Realization*. New York, NY: Oxford University Press.

Simon, H. (1973). The Organization of Complex Systems. In Pattee, H. (ed.) *Hierarchy Theory*. New York, NY: Brazillier.

Sizer, L. (2000). Towards a Computational Theory of Mood. *British J. Phil. Sci 51*: 743-769.

Sporns, O. (2011). *Networks of the Brain*. Cambridge, MA: MIT Press.

Sprevak, M. (2010). Computation, Individuation, and the Received View on Representation. *Studies in History and Philosophy of Science 41:* 260–270.

Strawson, G. (2010). Fundamental Singleness: How to Turn the 2nd Paralogism into a Valid Argument. *Royal Institute of Philosophy Supplement 67*: 61-92.

Stuart, S. (2010). Conscious machines: memory, melody and muscular imagination. *Phenomenology and Cognitive Science* 9: 37-51.

Van Heuveln, B., Dietrich, E. and Oshima, M. (1998). *Let's Dance! The Equivocation in Chalmers' Dancing Qualia Argument*. Minds and Machines 8 (2): 237-249.

Walling, P. and Hicks, K. (2003). Dimensions of Consciousness. *BUMC Proceedings 16*: 162-166.

Weisberg, J. (forthcoming). The Zombie's Cogito: Meditations on Type-Q Materialism. Philosophical Psychology.

Werner, G. (2010). Fractals in the Nervous System. *Frontiers in Physiology 1* (15): 1-28.

West, B. (2006). Complexity, Scaling and Fractals in Biological Signals. *Wiley Encyclopedia of Biomedical Engineering*.

West, G., Brown, J. and Enquist, B. (1999). The Fourth Dimension of Life: Fractal Geometry and Allometric Scaling of Organisms. *Science 284* (4): 1677-1679.