NEWSLETTER | The American Philosophical Association

Philosophy and Computers

FALL 2018 VOLUME 18 | NUMBER 1

ISSN 2155-9708

APA NEWSLETTER ON

Philosophy and Computers

PETER BOLTUC, EDITOR

VOLUME 18 | NUMBER 1 | FALL 2018

MISSION STATEMENT

Mission Statement of the APA Committee on Philosophy and Computers: Opening of a Short Conversation

Marcello Guarini
UNIVERSITY OF WINDSOR

Peter Boltuc
UNIVERSITY OF ILLINOIS, SPRINGFIELD, AND THE WARSAW SCHOOL OF ECONOMICS

A number of years ago, the committee was charged with the task of revisiting and revising its charge. This was a task we never completed. We failed to do so not for the lack of trying (there have been several internal debates at least since 2006) but due to the large number of good ideas. As readers of this newsletter know, the APA committee dedicated to philosophy and computers has been scheduled to be dissolved as of June 30, 2020. Yet, it is often better to do one's duty late rather than never. In this piece, we thought we would draft what a revised charge might look like. We hope to make the case that there is still a need for the committee. If that ends up being unpersuasive, we hope that a discussion of the activities in which the committee has engaged will serve as a guide to any future committee(s) that might be formed, within or outside of the APA, to further develop some of the activities of the philosophy and computers committee.

The original charge for the philosophy and computers committee read as follows:

> The committee collects and disseminates information on the use of computers in the profession, including their use in instruction, research, writing, and publication, and it makes recommendations for appropriate actions of the board or programs of the association.

As even a cursory view of our newsletter would show, this is badly out of date. Over and above the topics in our original charge, the newsletter has engaged issues in the ethics and philosophy of data, information, the internet, e-learning in philosophy, and various forms of computing, not to mention the philosophy of artificial intelligence, the philosophy of computational cognitive modeling, the philosophy of computer science, the philosophy of information, the ethics of increasingly intelligent robots, and other topics as well. Authors and perspectives published in the newsletter have come from different disciplines, and that has only served to enrich the content of our discourse. If a philosopher is theorizing about the prospects of producing consciousness in a computational architecture, it might not be a bad idea to interact with psychologists, cognitive scientists, and computer scientists. If one is doing information ethics, a detailed knowledge of how users are affected by information or information policy—which could come from psychology, law, or other disciplines—clearly serves to move the conversation forward.

The original charge made reference to "computers in the profession," never imagining how the committee's interests would evolve in both an inter- and multidisciplinary manner. While the committee was populated by philosophers, the discourse in the newsletter and APA conference sessions organized by the committee has been integrating insights from other disciplines into philosophical discourse. Moreover, the discourse organized by the committee has implications outside the profession. Finally, even if we focus only on computing in the philosophical profession, the idea that the committee simply "collects and disseminates information on the use of computers" never captured the critical and creative work not only of the various committee members over the years, but of the various contributors to the newsletter and to the APA conference sessions. It was never about simply collecting and disseminating. Think of the white papers produced by two committee members who published in the newsletter in 2014: "Statement on Open-Access Publication" by Dylan E. Wittkower, and "Statement on Massive Open Online Courses (MOOCs)" by Felmon Davis and Dylan E. Wittkower. These and other critical and creative works added important insights to discussions of philosophical publishing and pedagogy. The committee was involved in other important discussions as well. Former committee chair Thomas Powers provided representation in a 2015–2016 APA Subcommittee on Interview Best Practices, chaired by Julia Driver. The committee's participation was central because much of the focus was on Skype interviews. Once again, it was about much more than collecting and disseminating.

Over the years, the committee also has developed relationships with the International Association for Computing and Philosophy (IACAP) and International Society for Ethics and Information Technology. Members of these and other groups have attended APA committee sessions and published in the newsletter. The committee has developed relationships both inside and outside of philosophy, and both inside and outside of the APA. This has served us well with respect to being able to organize sessions at APA conferences. In 2018, we organized a session at each of the Eastern, Central, and Pacific meetings. We are working to do the same for 2019, and we are considering topics such as the nature of computation, machine consciousness, data ethics, and Turing's work.

In light of the above reasons, we find it important to clarify the charges of the committee still in 2018. A revised version of the charge that better captures the breadth of the committee's activities might look as follows:

> **The committee works to provide forums for discourse devoted to the critical and creative examination of the role of information, computation, computers, and other computationally enabled technologies (such as robots). The committee endeavors to use that discourse not only to enrich philosophical research and pedagogy, but to reach beyond philosophy to enrich other discourses, both academic and non-academic.**

We take this to be a short descriptive characterization. We are not making a prescription for what the committee should become. Rather, we think this captures, much better than the original charge, what it has actually been doing, or so it appears to us. Since the life of this committee seems to be coming to an end shortly, we would like to open this belated conversation now and to close it this winter, at the latest. While it may be viewed as a last ditch effort of sorts, its main goal is to explore the need for the work this committee has been doing at least for the last dozen years. This would provide more clarity on what institutional framework, within or outside of the APA, would be best suited for the tasks involved.

There have been suggestions to update the name of the committee as well as its mission. While the current name seems nicely generic, thus inclusive of new subdisciplines and areas of interest, the topic of the name may also be on the table.

We very much invite feedback on this draft of a revised charge or of anything else in this letter. We invite not only commentaries that describe what the committee has been doing, but also reflections on what it could or should be doing, and especially what people would like to see over the next two years. All readers of this note, including present and former members of the committee, other APA members, authors in our newsletter, other philosophers and non-philosophers interested in this new and growing field, are encouraged to contact us. Feel free to reply to either or both of us at:

Marcello Guarini, Chair, mguarini@uwindsor.ca

Peter Boltuc, Vice-Chair, pboltu@sgh.waw.pl

FROM THE EDITOR

Piotr Boltuc
UNIVERSITY OF ILLINOIS, SPRINGFIELD, AND THE WARSAW SCHOOL OF ECONOMICS

The topic of several papers in the current issue seems to be radical difference between the reductive and nonreductive views on intentionality, which (in)forms the rift between the two views on AI. To make things easy, there are two diametrically different lessons that can be drawn from Searle's *Chinese room*. For some, such as W. Rapaport, Searle's thought experiment is one way to demonstrate how semantics collapses into syntax. For others, such as R. Baker, it demonstrates that nonreductive first-person consciousness is necessary for intentionality, thus also for consciousness.

We feature the article on Machine Intentions by Don Berkich (the current president of the International Association for Computing and Philosophy), which is an homage to L. R. Baker—Don's mentor and our esteemed author. Berkich tries to navigate between the horns of the dilemma created by strictly functional and nonreductive requirements on human, and machine, agency. He tries to replace the Searle-Castaneda definition of intentionality, that requires first-person consciousness, with a more functionalistic definition by Davidson. Thus, he agrees with Baker that robots require intentionality, yet disagrees with her that intentionality requires irreducible first-person perspective (FPP). Incidentally, Berkich adopts Baker's view that FPP requires self-consciousness. (If we were talking of *irreducible first-person consciousness*, it would be quite clear these days that it is distinct from self-consciousness, but *irreducible first-person perspective* invokes some old-school debates.) On its final pages, the article contains a very clear set of arguments in support of Turing's critique of the Lady Lovelace's claim that machines cannot discover anything new.

In the "Logicist Remarks…" Selmer Bringsjord argues, *contra* W. Rapaport, that we should view computer science as a proper part of mathematical logic, instead of viewing it in a procedural way. In his second objection to Rapaport, Bringsjord argues that semantics does not collapse into syntax because of the reasons demonstrated in Searle's *Chinese room*. The reason being that "our understanding" is "bound up with subjective understanding," which brings us back to Baker's point discussed by Berkich.

In his response to Bringsjord on a procedural versus logicist take on computer science, Rapaport relies on Castaneda (quite surprisingly, as his is one of the influential nonreductive definitions of intentionality). Yet, Rapaport relates to Castaneda's take on philosophy as "the personal search for truth"—but he may be viewing personal search for the truth as a search for personal truth, which does not seem to be Castaneda's point. This subjectivisation looks like Rapaport is going for a draw—though he seems to present a stronger point in his interview with Robin Hill that follows. Rapaport seems to have a much stronger response defending his view on semantics as syntax, but

I'll not spoil the read of this very short paper. Bill Rapaport's interview with R. K. Hill revisits some of the topics touched on by Bringsjord, but I find the case in which he illustrates the difference between instructions and algorithms both instructive and lively.

This is followed by two ambitious sketches within the realm of theoretical logic. Doukas Kapantaïs presents an informal write-up of his formal counterexample to the standard interpretation of Church-Turing thesis. Joseph E. Brenner follows with a multifarious article that presents a sketch of a version of para-consistent (or dialectical) logic aimed at describing consciousness. The main philosophical point is that thick definition consciousness always contains contradiction though the anti-thesis remains unconscious for the time being. While the author does bring the argument to human consciousness but not all the way to artificial general intelligence, the link can easily be drawn.

We close with three papers on e-learning and philosophy. We have a thorough discussion by a professor, Fritz J. McDonald, who discusses the rare species of synchronous online classes in philosophy and the mixed blessings that come from teaching them. This is followed by a short essay by a student, Adrienne Anderson, on her experiences taking philosophy online. She is also a bit skeptical of taking philosophy courses online, but largely for the reason that there is little, if any, synchronicity (and bodily presence) in the online classes she has taken. We end with a perspective by an administrator, Jeff Harmon, who casts those philosophical debates in a more practical dimension.

Let me also mention the note from the chair and vice chair pertaining to the mission of this committee—you have probably read it already since we placed it above the note from the chair and my note.

FROM THE CHAIR

Marcello Guarini
UNIVERSITY OF WINDSOR

The committee has had a busy year organizing sessions for the APA meetings, and things continue to move in the same direction. Our recent sessions at the 2018 meetings of the Eastern, Central, and Pacific meetings were well attended, and we are planning to organize three new sessions—one for each of the upcoming 2019 meetings. For the Eastern Division meeting, we are looking to organize a book panel on Gualtiero Piccinini's *Physical Computation: A Mechanistic Account* (Oxford University Press, 2015). For the Central Division meeting, we are working on a sequel to the 2018 session on machine consciousness. For the upcoming Pacific Division meeting, we are pulling together a session on data ethics. We are even considering a session on Turing's work, but we are still working out whether that will take place in 2019 or 2020.

While it is true that the philosophy and computers committee is scheduled for termination as of June 30, 2020, the committee fully intends to continue organizing high-quality sessions at APA meetings for as long as it can. Conversations have started about how the work done by the committee can continue, in one form or another, after 2020. The committee has had a long and valuable history, one that has transcended its original charge. For this issue, Peter Boltuc (our newsletter editor and associate committee chair) and I composed a letter reviewing our original charge and explained the extent to which the committee moved beyond that charge. We hope that letter communicates at least some of the diversity and value of what the committee has been doing, and by "committee" I refer to both its current members and its many past members.

As always, if anyone has ideas for organizing philosophy and computing sessions at future APA meetings, please feel free to get in touch with us. There is still time to make proposals for 2020, and we are happy to continue working to ensure that our committee provides venues for high-quality discourse engaging a wide range of topics at the intersection of philosophy and computing.

FEATURED ARTICLE

Machine Intentions

Don Berkich
TEXAS A&M UNIVERSITY

INTRODUCTION

There is a conceptual tug-of-war between the AI crowd and the mind crowd.[1] The AI crowd tends to dismiss the skeptical markers placed by the mind crowd as unreasonable in light of the range of highly sophisticated behaviors currently demonstrated by the most advanced robotic systems. The mind crowd's objections, it may be thought, result from an unfortunate lack of technical sophistication which leads to a failure to grasp the full import of the AI crowd's achievements. The mind crowd's response is to point out that sophisticated behavior alone ought never be taken as a sufficient condition on full-bore, human-level mentality.[2]

I think it a mistake for the AI crowd to dismiss the mind crowd's worries without very good reasons. By keeping the AI crowd's feet to the fire, the mind crowd is providing a welcome skeptical service. That said, in some cases there are very good reasons for the AI crowd to push back against the mind crowd; here I provide a specific and, I submit, important case-in-point so as to illuminate some of the pitfalls in the tug-of-war.

It can be argued that there exists a counterpart to the distinction between *original intentionality* and *derived intentionality* in agency: Given its design specification, a machine's agency is at most derived from its designer's original agency, even if the machine's resulting behavior sometimes surprises the designer. The argument for drawing this distinction hinges on the notion that intentions are necessarily *conferred* on machines by their designers' ambitions, and intentions have features which immunize them from computational modeling.

In general, skeptical arguments against original machine agency may usefully be stated in the *Modus Tollens* form:

1. If X is an original agent, then X must have property P.
2. No machine can have property P.

∴ 3. No machine can be an original agent. 1&2

The force of each skeptical argument depends, of course, on the property P: The more clearly a given P is such as to be required by original agency but excluded by mechanism the better the skeptic's case. By locating property P in intention formation in an early but forcefully argued paper, Lynne Rudder Baker[3] identifies a particularly potent skeptical argument against original machine agency. I proceed as follows. In the first section I set out and refine Baker's challenge. In the second section I describe a measured response. In the third and final section I use the measured response to draw attention to some of the excesses on both sides.[4]

THE MIND CROWD'S CHALLENGE: BAKER'S SKEPTICAL ARGUMENT

Roughly put, Baker argues that machines cannot act since actions require intentions, intentions require a first-person perspective, and no amount of third-person information can bridge the gap to a first-person perspective. Baker[5] usefully sets her own argument out:

A 1. In order to be an agent, an entity must be able to formulate intentions.

2. In order to formulate intentions, an entity must have an irreducible first-person perspective.

3. Machines lack an irreducible first-person perspective.

∴ 4. Machines are not agents. 1,2&3

Baker has not, however, stated her argument quite correctly. It is not just that machines are not (original) agents or do not happen presently to be agents, since that allows that at some point in the future machines may be agents or at least that machines can in principle be agents. Baker's conclusion is actually much stronger. As she outlines her own project, "[w]ithout denying that artificial models of intelligence may be useful for suggesting hypotheses to psychologists and neurophysiologists, I shall argue that there is a radical limitation to applying such models to human intelligence. And this limitation is exactly the reason why computers can't act."[6]

Note that "computers can't act" is substantially stronger than "machines are not agents." Baker wants to argue that it is impossible for machines to act, which is presumably more difficult than arguing that we don't at this time happen to have the technical sophistication to create machine agents. Revising Baker's extracted argument to bring it in line with her proposed conclusion, however, requires some corresponding strengthening of premise A.3, as follows:

B 1. In order to be an original agent, an entity must be able to formulate intentions.

2. In order to formulate intentions, an entity must have an irreducible first-person perspective.

3. Machines necessarily lack an irreducible first-person perspective.

∴ 4. Machines cannot be original agents. 1,2&3

Argument B succeeds in capturing Baker's argument provided that her justification for B.3 has sufficient scope to conclude that machines cannot in principle have an irreducible first-person perspective. What support does she give for B.1, B.2, and B.3?

B.1 is true, Baker asserts, because original agency implies intentionality. She takes this to be virtually self-evident; the hallmark of original agency is the ability to form intentions, where intentions are to be understood on Castaneda's[7] model of being a "dispositional mental state of endorsingly thinking such thoughts as 'I shall do A'."[8] B.2 and B.3, on the other hand, require an account of the first-person perspective such that

- The first person perspective is necessary for the ability to form intentions; and
- Machines necessarily lack it.

As Baker construes it, the first person perspective (FPP) has at least two essential properties. First, the FPP is irreducible, where the irreducibility in this case is due to a linguistic property of the words used to refer to persons. In particular, first person pronouns cannot be replaced with descriptions *salve veritate*. "First-person indicators are not simply substitutes for names or descriptions of ourselves."[9] Thus Oedipus can, without absurdity, demand that the killer of Laius be found. "In short, thinking about oneself in the first-person way does not appear reducible to thinking about oneself in any other way."[10]

Second, the FPP is necessary for the ability to "conceive of one's thoughts as one's own."[11] Baker calls this "second-order consciousness." Thus, "if X cannot make first-person reference, then X may be conscious of the contents of his own thoughts, but not conscious that they are his own."[12] In such a case, X fails to have second-order consciousness. It follows that "an entity which can think of propositions at all enjoys self-consciousness if and only if he can make irreducible first-person reference."[13] Since the ability to form intentions is understood on Castaneda's model as the ability to endorsingly think propositions such as "I shall do A," and since such propositions essentially involve first-person reference, it is clear why the first person perspective is necessary for the ability to form intentions. So we have some reason to think that B.2 is true. But, apropos B.3, why should we think that machines necessarily lack the first-person perspective?

Baker's justification for B.3 is captured by her claim that "[c]omputers cannot make the same kind of reference to themselves that self-conscious beings make, and this difference points to a fundamental difference between humans and computers—namely, that humans, but not computers, have an irreducible first-person perspective."[14] To make the case that computers are necessarily handicapped in that they cannot refer to themselves in the same way that self-conscious entities do, she invites us to consider what would have to be the case for a first person perspective to be programmable:

a) FPP can be the result of information processing.

b) First-person episodes can be the result of transformations on discrete input via specifiable rules.[15]

Machines necessarily lack an irreducible first-person perspective since both (a) and (b) are false. (b) is straightforwardly false, since "the world we dwell in cannot be represented as some number of independent facts ordered by formalizable rules."[16] Worse, (a) is false since it presupposes that the FPP can be generated by a rule governed process, yet the FPP "is not the result of any rule-governed process."[17] That is to say, "no amount of third-person information about oneself ever compels a shift to first person knowledge."[18] Although Baker does not explain what she means by "third-person information" and "first-person knowledge," the point, presumably, is that there is an unbridgeable gap between the third-person statements and the first-person statements presupposed by the FPP. Yet since the possibility of an FPP being the result of information processing depends on bridging this gap, it follows that the FPP cannot be the result of information processing. Hence it is impossible for machines, having only the resource of information processing as they do, to have an irreducible first-person perspective.

Baker's skeptical challenge to the AI crowd may be set out in detail as follows:

C 1. Necessarily, X is an original agent only if X has the capacity to formulate intentions.

2. Necessarily, X has the capacity to formulate intentions only if X has an irreducible first person perspective.

3. Necessarily, X has an irreducible first person perspective only if X has second-order consciousness.

4. Necessarily, X has second-order consciousness only if X has self-consciousness.

∴ 5. Necessarily, X is an original agent only if X has self-consciousness. 1,2,3&4

6. Necessarily, X is a machine only if X is designed and programmed.

7. Necessarily, X is designed and programmed only if X operates just according to rule-governed transformations on discrete input.

8. Necessarily, X operates just according to rule-governed transformations on discrete input only if X lacks self-consciousness.

∴ 9. Necessarily, X is a machine only if X lacks self-consciousness. 6,7&8

∴ 10. Necessarily, X is a machine only if X is not an original agent. 5&9

A MEASURED RESPONSE ON BEHALF OF THE AI CROWD

While there presumably exist skeptical challenges which ought not be taken seriously because they are, for want of careful argumentation, themselves unserious, I submit that Baker's skeptical challenge to the AI crowd is serious and ought to be taken as such. It calls for a measured response. It would be a mistake, in other words, for the AI crowd to dismiss Baker's challenge out of hand for want of technical sophistication, say, in the absence of decisive counterarguments. Moreover, counterarguments will not be decisive if they simply ignore the underlying import of the skeptic's claims.

For example, given the weight of argument against physicalist solutions to the hard problem of consciousness generally, it would be incautious of the AI crowd to respond by rejecting C.8 (but see[19] for a comprehensive review of the hard problem). In simple terms, the AI crowd should join the mind crowd in finding it daft at this point for a roboticist to claim that *there is something it is like to be her robot*, however impressive the robot or resourceful the roboticist in building it.

A more modest strategy is to sidestep the hard problem of consciousness altogether by arguing that having an irreducible FPP is not, contrary to C.2, a necessary condition on the capacity to form intentions. This is the appropriate point to press provided that it also appeals to the mind crowd's own concerns. For instance, if it can be argued that the requirement of an irreducible FPP is too onerous even for persons to formulate intentions under ordinary circumstances, then Baker's assumption of Castaneda's account will be vulnerable to criticism from both sides. Working from the other direction, it must also be argued the notion of programming that justifies C.7 and C.8 is far too narrow even if we grant that programming an irreducible FPP is beyond our present abilities. The measured response I am presenting thus seeks to moderate the mind crowd's excessively demanding conception of intention while expanding their conception of programming so as to reconcile, in principle, the *prima facie* absurdity of a programmed (machine) intention.

Baker's proposal that the ability to form intentions implies an irreducible FPP is driven by her adoption of Castaneda's[20] analysis of intention: To formulate an intention to A is to endorsingly think the thought, "I shall do A." There are, however, other analyses of intention which avoid the requirement of an irreducible FPP. Davidson[21] sketches an analysis of what it is to form an intention to act: "an action is performed with a certain intention if it is caused in the right way by attitudes and beliefs that rationalize it."[22] Thus,

> If someone performs an action of type A with the intention of performing an action of type B, then he must have a pro-attitude toward actions of type B (which may be expressed in the form: an action of type B is good (or has some other positive attribute)) and a belief that in performing an action of type A he will be (or probably will be) performing an action of type B (the belief may be expressed in the obvious way). The expressions of the belief and desire entail that actions of type A are, or probably will be, good (or desirable, just, dutiful, etc.).[23]

Davidson is proposing that S A's with the intention of B-ing only if

i. S has pro-attitudes towards actions of type B.

ii. S believes that by A-ing S will thereby B.

The pro-attitudes and beliefs S has which rationalize his action cause his action. But, of course, it is not the case that S's having pro-attitudes towards actions of type B and S's believing that by A-ing she will thereby B jointly implies that S actually A's with the intention of B-ing. (i) and (ii), in simpler terms, do not jointly suffice for S's A-ing with the intention of B-ing since it must be that S A's because of her pro-attitudes and beliefs. For Davidson, "because" should be read in its causal sense. Reasons consisting as they do of pro-attitudes and beliefs cause the actions they rationalize.

Causation alone is not enough, however. To suffice for intentional action reasons must cause the action in the right way. Suppose (cf[24]) Smith gets on the plane marked "London" with the intention of flying to London, England. Without alarm and without Smith's knowledge, a shy hijacker diverts the plane from its London, Ontario, destination to London, England. Smith's beliefs and pro-attitudes caused him to get on the plane marked "London" so as to fly to London, England. Smith's intention is satisfied, but only by accident, as it were. So it must be that Smith's reasons cause his action in the right way, thereby avoiding so called wayward causal chains. Hence, S A's with the intention of B-ing if, and only if,

i. S has pro-attitudes towards actions of type B.

ii. S believes that by A-ing S will thereby B.

iii. S's relevant pro-attitudes and beliefs cause her A-ing with the intention of B-ing in the right way.

Notice that there is no reference whatsoever involving an irreducible FPP in Davidson's account. Unlike Castaneda's account, there is no explicit mention of the first person indexical. So were it the case that Davidson thought animals could have beliefs, which he does not,[25] it would be appropriate to conclude from Davidson's account that animals can act intentionally despite worries that animals would lack an irreducible first-person perspective. Presumably robots would not be far behind.

It is nevertheless open to Baker to ask about (ii): S believes that by A-ing S will thereby B. Even if S does not have to explicitly and endorsingly think, "I shall do A" to A intentionally, (ii) requires that S has a self-referential belief that by A-ing he himself will thereby B. Baker can gain purchase on the problem by pointing out that such a belief presupposes self-consciousness every bit as irreducible as the FPP.

Consider, however, that a necessary condition on Davidson's account of intentional action is that S believes that by A-ing S will thereby B. Must we take 'S' in S's belief that by A-ing S will thereby B *de dicto*? Just as well, could it not be the case (*de re*) that S believes, of itself, that by A-ing it will thereby B?

The difference is important. Taken *de dicto*, S's belief presupposes self-consciousness since S's belief is equivalent to having the belief, "by A-ing I will thereby B." Taken (*de re*), however, S's belief presupposes at most self-*representation*, which can be tokened without solving the problem of (self) consciousness.

Indeed, it does not seem to be the case that the intentions I form presuppose either endorsingly thinking "I shall do A!" as Castaneda (and Baker) would have it or a *de dicto* belief that by A-ing I will B as Davidson would have it. Intention-formation is transparent: I simply believe that A-ing B's, so I A. The insertion of self-consciousness as an intermediary requirement in intention formation would effectively eliminate many intentions in light of environmental pressures to act quickly. Were Thog the caveman required to endorsingly think "I shall climb this tree to avoid the saber-toothed tiger" before scrambling up the tree he would lose precious seconds and, very likely, his life. Complexity, particularly temporal complexity, constrains us as much as it does any putative original machine agent. A theory of intention which avoids this trouble surely has the advantage over theories of intention which do not.

In a subsequent pair of papers[26] and a book,[27] Baker herself makes the move recommended above by distinguishing between weak and strong first-person phenomena (later recast in more developmentally discerning terms as "rudimentary" and "robust" first-person perspectives), on the one hand, and between minimal, rational, and moral agency, on the other. Attending to the literature in developmental psychology (much as many in the AI crowd have done and would advise doing), Baker[28] argues that the rudimentary FPP is properly associated with minimal—that is, non-reflective—agency, which in turn is characteristic of infants and pre-linguistic children and adult animals of other species. Notably, the rudimentary FPP does *not* presuppose an *irreducible* FFP, although the robust FPP constituitively unique to persons does. As Baker puts it,

> [P]ractical reasoning is always first personal: The agent reasons about what to do on the basis of her own first-person point of view. It is the agent's first-person point of view that connects her reasoning to what she actually does. Nevertheless, the agent need not have any first-person concept of herself. A dog, say, reasons about her environment from her own point of view. She is at the origin of what she can reason about. She buries a bone at a certain location and later digs it up. Although we do not know exactly what it's like to be a dog, we can approximate the dog's practical reasoning from the dog's point of view: Want bone; bone is buried over there; so, dig over there. The dog is automatically (so to speak) at the center of the her world without needing self-understanding.[29]

Baker further argues in these pages[30] that, despite the fact that artifacts like robots are intentionally made for some purpose or other while natural objects sport no such teleological origin, "this differences does not signal any ontological deficiency in artifacts *qua* artifacts." Artifacts suffer no demotion of ontological status insofar as they are ordinary objects regardless of origin. Her argument, supplemented and supported by Amie L. Thomasson,[31] repudiates drawing on the distinction between mind-dependence and mind-independence (partly) in light of the fact that,

> [A]dvances in technology have blurred the difference between natural objects and artifacts. For example, so-called digital organisms are computer programs that (like biological organisms) can mutate, reproduce, and compete with one another. Or consider robo-ratsrats with implanted electrodesthat direct the rats movements. Or, for another example, consider what one researcher calls a bacterial battery: these are biofuel cells that use microbes to convert organic matter into electricity. Bacterial batteries are the result of a recent discovery of a micro-organism that feeds on sugar and converts it to a stream of electricity. This leads to a stable source of low power that can be used to run sensors of household devices. Finally, scientists are genetically engineering viruses that selectively infect and kill cancer cells and leave healthy cells alone. Scientific American referred to these viruses as search-and-destroy missiles. Are these objects—the digital organisms, robo-rats, bacterial batteries, genetically engineered viral search-and-destroy missilesartifacts or natural objects? Does it matter? I suspect that the distinction between artifacts and natural objects will become increasingly fuzzy; and, as it does, the worries about the mind-independent/mind-dependent distinction will fade away.[32]

Baker's distinction between rudimentary and robust FPPs, suitably extended to artifacts, may cede just enough ground to the AI crowd to give them purchase on at least *minimal* machine agency, all while building insurmountable ramparts against the AI crowd to defend, on behalf of the mind crowd, the special status of persons, enjoying as they must their computationally intractable robust FPPs. Unfortunately Baker does not explain precisely how the minimal agent enjoying a rudimentary FPP develops into a moral agent having the requisite robust FPP. That is, growing children readily, gracefully, and easily scale the ramparts simply in the course of their normal development, yet how remains a mystery.

At most we can say that there are many things a minimal agent cannot do rational (reflective) and moral (responsible) agents can do. Moreover, the mind crowd may object that Baker has in fact ceded no ground whatsoever, since even a suitably attenuated conception of intention cannot be programmed under Baker's conception of programming. What is her conception of programming? Recall that Baker defends B.3 by arguing that machines cannot achieve a first-person perspective since machines gain information *only* through rule-based transformations on discrete input and no amount or combination of such transformations could suffice for the transition from a third-person perspective to a first-person perspective. That is,

D 1. If machines were able to have a FPP, then the FPP can be the result of transformations on discrete input via specifiable rules.

2. If the FPP can be the result of transformations on discrete input via specifiable rules, then there exists some amount of third-person information which compels a shift to first-person knowledge.

3. No amount of third-person information compels a shift to first-person knowledge.

∴ 4. First-person episodes cannot be the result of transformations on discrete input via specifiable rules. 2&3

∴ 5. Machines necessarily lack an irreducible first-person perspective. 1&4

The problem with D is that it betrays an overly narrow conception of machines and programming, and this is true even if we grant that we don't presently know of *any* programming strategy that would bring about an irreducible FPP.

Here is a simple way of thinking about machines and programming as Argument D would have it. There was at one time (for all I know, there may still be) a child's toy which was essentially a wind-up car. The car came with a series of small plastic disks, with notches around the circumference, which could be fitted over a rotating spindle in the middle of the car. The disks acted as a cam, actuating a lever which turned the wheels when the lever hit a notch in the side of the disk. Each disk had a distinct pattern of notches and resulted in a distinct route. Thus, placing a particular disk on the car's spindle "programs" the car to follow a particular route.

Insofar as it requires that programming be restricted to transformations on discrete input via specifiable rules, Argument D treats all machines as strictly analogous to the toy car and programming as analogous to carving out new notches on a disk used in the toy car. Certainly Argument D allows for machines which are much more complicated than the toy car, but the basic relationship between program and machine behavior is the same throughout. The program determines the machine's behavior, while the program itself is in turn determined by the programmer. It is the point of D.2 that, if an irreducible FPP were programmable, it would have to be because the third-person information which can be supplied by the programmer suffices for a first-person perspective, since all the machine has access to is what can be supplied by a programmer. Why should we think that a machine's only source of information is what the programmer provides? Here are a few reasons to think that machines are not so restricted:

- Given appropriate sensory modalities and appropriate recognition routines, machines are able to gain information about their environment without that information having been programmed in advance.[33] It would be as if the toy car had an echo-locator on the front and a controlling disk which notched itself in reaction to obstacles so as to maneuver around them.

- Machines can be so constructed as to "learn" by a variety of techniques.[34] Even classical conditioning techniques have been used. The point is merely that suitably constructed, a machine can put together information about its environment and itself which is not coded in advance by the programmer and which is not available other than by, for example, trial and error. It would be as if the toy car had a navigation goal and could adjust the notches in its disk according to whether it is closer or farther from its goal.

- Machines can evolve.[35] Programs evolve through a process of mutation and extinction. Code in the form of so-called genetic algorithms is replicated and mutated. Unsuccessful mutations are culled, while successful algorithms are used as the basis for the next generation. Using this method one can develop a program for performing a particular task without having any knowledge of how the program goes about performing the task. Strictly speaking, there is no programmer for such programs. Here the analogy with the toy car breaks down somewhat. It's as if the toy car started out with a series of disks of differing notch configurations and the car can take a disk and either throw it out or use it as a template for further disks, depending on whether or not a given disk results in the car being stuck against an obstacle, for instance.

- Programs can be written which write their own programs.[36] A program can spawn an indefinite number of programs, including an exact copy of itself. It need not be the case that the programmer be able to predict what future code will be generated, since that code may be partially the result of information the machine gathers, via sensory modalities, from its environment. So, again, in a real sense there is no programmer for these programs. The toy car in this case starts out with a disk which itself generates disks and these disks may incorporate information about obstacles and pathways.

Indeed, many of the above techniques develop Turing's own suggestions:

Let us return for a moment to Lady Lovelace's objection, which stated that the machine can only do what we tell it to do.

> Instead of trying to produce a programme to simulate the adult mind, why not rather try to produce one which simulates the child's? If this were then subjected to an appropriate course of education one would obtain the adult brain. Presumably the child brain is something like a notebook as one buys it from the stationer's. Rather little mechanism, and lots of blank sheets. (Mechanism and writing are from our point of view almost synonymous.) Our hope is that there is so little mechanism in the child brain that something like it can be easily programmed. The amount of work in the education we can assume, as a first approximation, to be much the same as for the human child.
>
> We have thus divided our problem into two parts. The child programme and the education process. These two remain very closely connected. We cannot expect to find a good child machine at the first attempt. One must experiment with teaching one such machine and see how well it learns...
>
> The idea of a learning machine may appear paradoxical to some readers. How can the rules of operation of the machine change? They should describe completely how the machine will react whatever its history might be, whatever changes it might undergo. The rules are thus quite time-invariant. This is quite true. The explanation of the paradox is that the rules which get changed in the learning process are of a rather less pretentious kind, claiming only an ephemeral validity. The reader may draw a parallel with the Constitution of the United States.[37]

As Turing anticipated, machines can have access to information and utilize it in ways which are completely beyond the purview of the programmer. So while it may not be the case that a programmer can write code for an irreducible FPP, as Argument D requires, it still can be argued that the sources of information available to a suitably programmed robot nevertheless enable it to formulate intentions when intentions do not also presuppose an irreducible FPP.

Consider the spectacularly successful Mars rovers Spirit and Opportunity. Although the larger goal of moving from one location to another was provided by mission control, specific routes were determined in situ by constructing maps and evaluating plausible routes according to obstacles, inclines, etc. Thus the Mars rovers were, in a rudimentary sense, gleaning information from their environment and using that information to assess alternatives so as to plan and execute subsequent actions. None of this was done with the requirement of, or pretense to having, an irreducible FPP, yet it does come closer to fitting the Davidsonian model of intentions. To be sure, this is intention-formation of the crudest sort, and it requires further argument that propositional attitudes themselves are computationally tractable.

A LARGER POINT: AVOIDING EXCESSES ON BOTH SIDES

Baker closes her original article by pointing out that robots' putative inability to form intentions has far-reaching implications:

> So machines cannot engage in intentional behavior of any kind. For example, they cannot tell lies, since lying involves the intent to deceive; they cannot try to avoid mistakes, since trying to avoid mistakes entails intending to conform to some normative rule. They cannot be malevolent, since having no intentions at all, they can hardly have wicked intentions. And, most significantly, computers cannot use language to make assertions, ask questions, or make promises, etc., since speech acts are but a species of intentional action. Thus, we may conclude that a computer can never have a will of its own.[38]

The challenge for the AI crowd, then, is to break the link Baker insists exists between intention formation and an irreducible FPP in its robust incarnation. For if Baker is correct and the robust FPP presupposes self-consciousness, the only way the roboticist can secure machine agency is by solving the vastly more difficult problem of consciousness, which so far as we presently know is a computationally impenetrable problem. I have argued that the link can be broken, provided a defensible and computationally tractable account of intention is available to replace Castaneda's overly demanding account.

If my analysis is sound, then there are times when it is appropriate for the AI crowd to push back against the mind crowd. Yet they must do so in such a way as to respect so far as possible the ordinary notions the mind crowd expects to see employed. In this case, were the AI crowd to so distort the concept of intention in their use of the term that it no longer meets the mind crowd's best expectations, the AI crowd would merely have supplied the mind crowd with further skeptical arguments. In this sense, the mind crowd plays a valuable role in demanding that the AI crowd ground their efforts in justifiable conceptual requirements, which in no way entails that the AI crowd need accept those conceptual requirements without further argument. Thus the enterprise of artificial intelligence has as much to do with illuminating the efforts of the philosophers of mind as the latter have in informing those working in artificial intelligence.

This is a plea by example, then, to the AI crowd that they avoid being overly satisfied with themselves simply for simulating interesting behaviors, unless of course the point of the simulation simply is the behavior. At the same time, it is a plea to the mind crowd that they recognize when their claims go too far even for human agents and realize that the AI crowd is constantly adding to their repertoire techniques which can and should inform efforts in the philosophy of mind.

NOTES

1. With apologies to BBC Channel 4's "The IT Crowd," airing 2006–2010.
2. Consider John Searle's article in the February 23, 2011, issue of the *Wall Street Journal*, aptly entitled, "Watson Doesn't Know It Won on Jeopardy!"
3. L. R. Baker, "Why Computer's Can't Act," *American Philosophical Quarterly* 18 (1981): 157–63.
4. This essay is intended in part to serve as a respectful homage to Lynne Rudder Baker, whose patience with unrefined, earnest graduate students and unabashed enthusiasm for rigorous philosophical inquiry wherever it may lead made her such a valued mentor.
5. Baker, "Why Computer's Can't Act," 157.
6. Ibid.
7. H-N. Castaneda, *Thinking and Doing: The Philosophical Foundations of Institutions* (Dordrecht: D. Reidel Publishing Co., 1975).
8. Baker, "Why Computer's Can't Act," 157.
9. Ibid.
10. Ibid., 158.
11. Ibid.
12. Ibid.
13. Ibid.
14. Ibid., 159.
15. Ibid.
16. Ibid., 160.
17. Ibid.
18. Ibid.
19. D. Chalmers, "Consciousness and Its Place in Nature," *Philosophy of Mind: Classical and Contemporary Readings*, 247–72 (Oxford: Oxford University Press, 2002).
20. Castaneda, *Thinking and Doing: The Philosophical Foundations of Institutions*.
21. D. Davidson, "Intending," *Essays on Actions and Events*, 83–102 (Oxford: Clarendon Press, 1980).
22. Ibid., 87.
23. Ibid., 86–87.
24. Ibid., 84–85.
25. D. Davidson, "Thought and Talk," *Inquiries into Truth and Interpretation*, 155–70 (Oxford: Clarendon Press, 1984).
26. L. R. Baker, "The First-Person Perspective: A Test for Naturalism," *American Philosophical Quarterly* 35, no. 4 (1998): 327–48; L. R. Baker, "First-Personal Aspects of Agency," *Metaphilosophy* 42, nos. 1-2 (2011): 1–16.
27. L. R. Baker, *Naturalism and the First-Person Perspective* (New York: Oxford University Press, 2013).
28. Baker, "First-Personal Aspects of Agency."

29. Baker,*Naturalism and the First-Person Perspective*, 189.
30. L. R. Baker, "The Shrinking Difference Between Artifacts and Natural Objects," *APA Newsletter on Philosophy and Computers* 07, no. 2 (2008): 2–5.
31. A. L. Thomasson, "Artifacts and Mind-Independence: Comments on Lynne Rudder Baker's 'The Shrinking Difference between Artifacts and Natural Objects'," *APA Newsletter on Philosophy and Computers* 08, no. 1 (2008): 25–26.
32. Baker, "The Shrinking Difference Between Artifacts and Natural Objects," 4.
33. R. C. Arkin, *Behavior Based Robotics* (Cambridge, MA: MIT Press, 1998).
34. R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*, 3rd ed. (Cambridge, MA: MIT Press. A Bradford Book, 1998).
35. D. H. Ballard, *An Introduction to Natural Computation* (Cambridge, MA: MIT Press, 1997).
36. Ibid.
37. A. M. Turing, "Computing Machinery and Intelligence," *Mind* 59 (1950): 454–58.
38. Baker, "Why Computer's Can't Act," 163.

REFERENCES

Arkin, R. C. *Behavior Based Robotics*. Cambridge, MA: MIT Press, 1998.

Baker, L. R. "First-Personal Aspects of Agency." *Metaphilosophy* 42, nos. 1-2 (2011): 1–16.

———. *Naturalism and the First-Person Perspective*. New York: Oxford University Press, 2013.

———. "The First-Person Perspective: A Test for Naturalism." *American Philosophical Quarterly* 35, no. 4 (1998): 327–48.

———. "The Shrinking Difference Between Artifacts and Natural Objects." *APA Newsletter on Philosophy and Computers* 07, no. 2 (2008): 2–5.

———. "Why Computer's Can't Act." *American Philosophical Quarterly* 18 (1981): 157–63.

Ballard, D. H. *An Introduction to Natural Computation*. Cambridge, MA: MIT Press, 1997.

Castaneda, H-N. *Thinking and Doing: The Philosophical Foundations of Institutions*. Dordrecht: D. Reidel Publishing Co., 1975.

Chalmers, D. "Consciousness and Its Place in Nature," 247–72. Philosophy of Mind: Classical and Contemporary Readings. Oxford: Oxford University Press, 2002.

Davidson, D. "Intending," 83–102. Essays on Actions and Events. Oxford: Clarendon Press, 1980.

———. "Thought and Talk," 155–70. Inquiries into Truth and Interpretation. Oxford: Clarendon Press, 1984.

Sutton, R. S., and A. G. Barto. *Reinforcement Learning: An Introduction*. 3rd. Cambridge, MA: MIT Press. A Bradford Book, 1998.

Thomasson, A. L. "Artifacts and Mind-Independence: Comments on Lynne Rudder Baker's 'The Shrinking Difference between Artifacts and Natural Objects'." *APA Newsletter on Philosophy and Computers* 08, no. 1 (2008): 25–26.

Turing, A.M. "Computing Machinery and Intelligence." *Mind* 59 (1950): 433–60.

LOGIC AND CONSCIOUSNESS

Consciousness as Process: A New Logical Perspective

Joseph E. Brenner
INTERNATIONAL CENTRE FOR THE PHILOSOPHY OF INFORMATION, JIAOTONG UNIVERSITY, XI'AN, CHINA

1. INTRODUCTION

A NEW LOGICAL APPROACH

I approach the nature of consciousness from a natural philosophical-logical standpoint based on a non-linguistic, non-truth-functional logic of real processes—Logic in Reality (LIR). As I will show, the LIR logic is strongly anti-propositional and anti-representationalist, and gives access to a structural realism that is scientifically as well as logically grounded. The elimination I effect is not that of the complex properties of human consciousness and reasoning but of the chimerical entities that are unnecessary to and interfere with beginning to understand it. I point to the relation of my logic to personal identity, intuition, and anticipation, viewed itself as a complex cognitive process that embodies the same logical aspects as other forms of cognition.

A TYPE F MONISM

In his seminal paper of 2002, David Chalmers analyzed several possible conceptions of consciousness based on different views of reality.[1] Type F Monism "is the view that consciousness is constituted by the intrinsic properties of fundamental physical entities: that is, by the categorial bases of fundamental physical dispositions. On this view, phenomenal or proto-phenomenal properties are located at the fundamental level of physical reality, and in a certain sense, underlie physical reality itself." Chalmers remarks that in contrast to other theories, Type F monism has received little critical examination.

LIR and the theory of consciousness I present in this paper are based on the work of Stéphane Lupasco (Bucharest, 1900–Paris, 1988). It could be designated as a Type F or Neutral Monism[2] *provided* that several changes are introduced into the standard definition: a) in complex systems, properties have processual as well as static characteristics. Much of the discussion about consciousness is otiose because of its emphasis on entities, objects, and events rather than processes; b) properties and processes, especially of complex phenomena like consciousness, are constituted by both actual and potential components, and both are causally efficient; c) properties do not underlie reality; they are reality. The first two points eliminate the attribution of panpsychism. This theory allows consciousness-as-process to be "hardware,"[3] albeit in a different way than nerves and computers. FPC is not information processing in the standard computationalist sense, since information itself, as well as FPC, is conceived of as a process.[4] For hardware we may also read, for FPC, proper ontological status.

2. THE PROBLEM OF LOGIC

I propose that the principles involved in my extension of logic to real phenomena, processes, and systems enable many problems of consciousness to be addressed from a new perspective. As a non-propositional logic "of and in reality," LIR is grounded in the fundamental dualities of the universe and provides a rationale for their operation in a dialectical manner at biological, cognitive, and social levels of reality. Application of the principles of LIR allows us to cut through a number of ongoing debates as to the "nature" of consciousness. LIR makes it possible to essentially deconstruct the concept of any mental entities—including representations, qualia, models and concepts of self and free will—that are a substitute for, or an addition to, the mental processes themselves. I have accomplished this without falling back into an identity theory of mind, as described in the *Stanford Encyclopedia of Philosophy*.[5] Recent developments in the Philosophy of Information by Floridi, Wu, and others support the applicability of LIR to consciousness and intelligence.[6]

I characterize the science of consciousness today as

- embodying a process ontology and metaphysics, following the work of Bickhard and his colleagues.
- integrating the obvious fact that consciousness is an emergent phenomenon, and that arguments against emergence, such as those of Kim, are otiose.
- placing computational models of mind in the proper context.

The brain is massively complex, parallel, and redundant, and a synthesis of multiple nested evolutionary processes. To further capture many of the essential aspects of consciousness, in my view, one still must:

- ground consciousness in fundamental physics, as a physical phenomenon;
- define the path from afferent stimuli to the conscious mind and the relation between conscious and unconscious;
- establish a basis for intentionality and free will as the basis for individual moral and responsible behavior;
- from a philosophical standpoint, avoid concepts of consciousness based on substance metaphysics.

Valid insights into the functioning of *some* groups or modules of neurons and their relation to consciousness have come from the work of Ehresmann using standard category theory.[7] Standard category and set theories, as well as computational models of consciousness, however, suffer from the inherent limitations for the discussion of complex phenomena imposed by their underlying bivalent propositional logics.

3. PROCESS METAPHYSICS; INTERACTIVISM

The fundamental metaphysical split between two kinds of substances, the factual, non-normative world and the mental, normative and largely intensional world, goes back to Descartes. In Mark Bickhard's succinct summary, substance metaphysics makes process problematic, emergence impossible, and normativity, including representational normativity, inexplicable.

The discussion of nature of consciousness is facilitated as soon as one moves from the idea that consciousness is a thing or structure, localized or delocalized to some sort of process view. This has been demonstrated by Mark Bickhard and his associates at Lehigh University in Pennsylvania in a paper entitled quite like mine, "Mind as Process"[8] and subsequently. Arguments can be made[9] to model causally efficient ontological emergence within a process metaphysics that deconstructs the challenges of both Kim (metaphysical) and Hume (logical). For example, Kim's view is that all higher level phenomena are causally epiphenomenal, and causally efficacious emergence does not occur. This argument depends on the assumption that fundamental particles participate in organization, but do *not* have organization of their own. The consequence is that organization is not a locus of causal power, and the emergence assumption that new causal power can emerge in new organization would contradict the assumption that things that have no organization hold the monopoly of causal power. Bickhard's counter is that particles as such do not exist; "everything" is quantum fields; such fields are processes; processes are organized; all causal power resides in such organizations; and different organizations can have different causal powers and consequently also novel or emergent causal power.

Representations have had a major role to play in discussions of the nature of consciousness. Interactivism, Bickhard's interactivist model of representation, is a good point to start our discussion since it purports to link representation, anticipation, and interaction. Anticipatory processes are emergent and normative, involving a functional relationship between the allegedly autonomous organism and its environment. The resulting interactive potentialities have truth values for the organism, constituting a minimal model of representation. Representation, whose evolutionary advantages are easy to demonstrate, is of future potentialities for future action or interaction by the organism, and Bickhard shows that standard encoding, correspondence, isomorphic, and pragmatic views of representation, such as that of Drescher, lead to incoherence. The major problem with this process view is that it still defines its validity in terms of the truth of propositions, without regard to the underlying real processes that constitute existence. Further, the ontological status of representations can by no means be taken for granted, as I will discuss. The interactivist movement towards a process ontology is to be welcomed, many of its underlying ontological assumptions regarding space, time, and causality embody principles of bivalent propositional logic or its modal, deontic, or paraconsistent versions. Such logics fail to capture critical aspects of real change and, accordingly, of emergent complex processes, especially consciousness. The extension of logic toward real phenomena attempts to do just that. The increase in explanatory power for the characteristics of processes is therefore, in this view, a new tool in the effort to develop a science of consciousness. It complements systemic approaches, computational approaches to anticipation, such as those of Daniel Du Bois and the informational approaches of Floridi.

4. LOGIC IN REALITY (LIR)

The concept of a logic particularly applicable to the science and philosophy of consciousness as well as other complex cognitive phenomena will be unfamiliar to most readers. I will show that this has been due to the restriction of logic to propositions or their mathematical equivalents, and an alternative form of logic is both possible and necessary. Someone to whom I described my physicalist, but non-materialist theory of consciousness commented, "But then mind is just matter knowing itself!" The problem with this formulation is that it appears illogical, perhaps even unscientific. The logical system I will now propose is a start on naturalizing this idea.

LIR is a new kind of logic, grounded in quantum physics, whose axioms and rules provide a framework for analyzing and explaining real world processes.[10] The term "Logic in Reality" (LIR) is intended to imply both 1) that the principle of change according to which reality operates is a *logic* embedded in it, *the* logic in reality; and 2) that what logic really *is* or should be involves this same real physical-metaphysical but also logical principle. The major components of this logic are the following:

- The foundation in the physical and metaphysical dualities of nature
- Its axioms and calculus intended to reflect real change
- The categorial structure of its related ontology
- A two-level framework of relational analysis

DUALITIES

LIR is based on the quantum mechanics of Planck, Pauli, and Heisenberg, and subsequent developments of twentieth-century quantum field theory. LIR states that the characteristics of energy—extensive and intensive; continuous and discontinuous; entropic and negentropic—can be formalized as a structural logical principle of dynamic opposition, an antagonistic duality inherent in the nature of energy (or its effective quantum field equivalent), and, accordingly, of all real physical and non-physical phenomena—processes, events, theories, etc. The key physical and metaphysical dualities are the following:

- Intensity and Extensity in Energy
- Self-Duality of Quantum and Gravitational Fields
- Attraction and Repulsion (Charge, Spin, others)
- Entropy: tendency toward Identity/ Homogeneity (2nd Law of Thermodynamics)
- Negentropy: tendency toward Diversity/ Heterogeneity (Pauli Exclusion Principle)
- Actuality and Potentiality
- Continuity and Discontinuity
- Internal and External

The Fundamental Postulate of LIR is that every element e always associated with a non-e, such that the actualization of one entails the potentialization of the other and vice versa, alternatively, without either ever disappearing completely. This applies to all complex phenomena, since without passage from actuality to potentiality and vice versa, no change is possible. Movement is therefore toward (asymptotic) non-contradiction of identity or diversity, or toward contradiction. The midpoint of semi-actualization and semi-potentialization of both is a point of maximum contradiction, a "*T*-state" resolving contradiction (or "counter-action"), from which new entities can emerge. Some examples of this are the following:

- Quantum Level: Uncertainty Principle
- Biological Level: Antibody/Antigen Interactions
- Cognitive Level: Conscious/Unconscious
- Social Level: Left–Right Swings

AXIOMS AND CALCULUS

Based on this "antagonistic" worldview, I have proposed axioms which "rewrite" the three major axioms of classical logic and add three more as required for application to the real world:

LIR1: *(Physical) Non-Identity*: There is no A at a given time that is identical to A at another time.

LIR2: *Conditional Contradiction*: A and non-A both exist at the same time, but only in the sense that when A is actual, non-A is potential, reciprocally and alternatively.

LIR3: *Included (Emergent) Middle*: An included or additional third element or T-state (T for "*tiers inclus*," included third term) emerges from the point of maximum contradiction at which A and non-A are equally actualized and potentialized, but at a higher level of reality or complexity, at which the contradiction is resolved.

LIR4: *Logical Elements*: The elements of the logic are all representations of real physical and non-physical entities.

LIR5: *Functional Association*: Every real logical element **e**—objects, processes, events—is always associated, structurally and functionally, with its anti-element or contradiction, non-**e**, without either ever disappearing completely; in physics terms, they are conjugate variables. This axiom applies to the classical pairs of dualities, e.g., identity and diversity.

LIR6: *Asymptoticity*: No process of actualization or potentialization of any element goes to 100 percent completeness.

The nature of these real-world elements can be assumed to be what are commonly termed "facts" or extra-linguistic entities or processes. The logic is a logic of an *included* middle, consisting of axioms and rules of inference for determining the state of the three dynamic elements involved in a phenomenon ("dynamic" in the physical sense, related to real rather than to formal change, e.g., of conclusions).

In the notation developed by Lupasco, and as far as I know used only by him, where **e** is any real-world element involved in some process of change; e_A means that **e** is predominantly actual and implies $\bar{e}_P$ meaning that non-**e** is predominantly potential; e_T and $\bar{e}_T$ mean that **e** in a **T-state** implies non-**e** in a **T-state**; and e_A means that non-**e** is predominantly actual implying e_P, that is, that **e** is potential. In the LIR calculus, the reciprocally determined "reality" values of the degree of actualization A, potentialization P and T-state T replace the truth values in standard truth tables, as summarized in the following notation where the symbol **T** refers *exclusively* to the **T-state**, the logical included middle defined by Axiom **LR3**.

These values have properties similar to non-standard probabilities. When there is actualization and potentialization of logical elements, their non-contradiction is always partial. Contradiction, however, cannot take place between two classical terms that are rigorously or totally actualized or absolute, that is, where the axiom of non-contradiction holds absolutely. The consequence is that no real element or event can be rigorously non-contradictory; it always contains an irreducible quantity of contradiction.

The semantics of LIR is non-truth-functional. LIR contains the logic of the excluded middle as a limiting case, approached asymptotically but only instantiated in simple situations and abstract contexts, e.g., computational aspects of reasoning and mathematical complexity. Paraconsistent logics do mirror some of the contradictory aspects of real phenomena, as Priest has shown in his work on inconsistency in the material sciences. However, in LIR the "contradiction" is conditional. In paraconsistent logics, propositions are "true" and "false" at the same time; in LIR, only in the sense that when one is actual, the other is potential. Truth is the truth of reality. I recall here Japaridze's subordination of truth in computability logic as a zero-interactivity-order case of computability.

LIR is a logic applying to processes, in a process-ontological view of reality, to trends and tendencies, rather than to "objects" or the steps in a state-transition picture of change. Relatively stable macrophysical objects and simple situations are the result of processes of processes going in the direction of a "non-contradictory" identity. Starting at the quantum level, it is the potentialities as well as actualities that are the carriers of the causal properties necessary for the emergence of new entities at higher levels. The overall theory is thus a metaphysics of energy, and LIR is the formal, logical part of that metaphysical theory. LIR is a non-arbitrary method for including contradictory elements in theories or models whose acceptance would otherwise be considered as invalidating them entirely. It is a way to "manage" contradiction, a task that is also undertaken by paraconsistent, inconsistency-adaptive, and ampliative-adaptive logics. More relevant Hegelian dialectic logics as "precursors" of LIR are reviewed briefly below.

CATEGORIAL NON-SEPARABILITY IN THE ONTOLOGY OF LIR

The third major component of LIR is the categorial ontology that fits its axioms. In this ontology, the sole material category is Energy, and the most important formal category is Dynamic Opposition. From the LIR metaphysical standpoint, for real systems or phenomena or processes in which real dualities are instantiated, their terms are *not* separated or separable! Real complex phenomena display a contradictional relation to or interaction between themselves and their opposites or contradictions. On the other hand, there are many phenomena in which such interactions are not present, and they, and the simple changes in which they are involved, can be described by classical, binary logic or its modern versions. The most useful categorial division that can be made is exactly this: phenomena that show non-separability of the terms of the dualities as an essential aspect of their existence at their level of reality and those that instantiate separability.

LIR thus approaches in a new way the inevitable problems resulting from the classical philosophical dichotomies, appearance and reality, as well as the concepts of space, time, and causality as categories with *separable categorial features*, including, for example, final and effective cause. Non-separability underlies the other metaphysical and phenomenal dualities of reality, such as determinism and indeterminism, subject and object, continuity and discontinuity, and so on. This is a "vital" concept: to consider process elements that are contradictorially linked as separable is a form of category error. I thus claim that non-separability at the macroscopic level, like that being explored at the quantum level, provides a principle of organization or structure in macroscopic phenomena that has been neglected in science and philosophy.

Stable macrophysical objects and simple situations, which can be discussed within binary logic, are the result of processes of processes going in the direction of non-contradiction. Thus, LIR should be seen as a logic applying to processes, to trends and tendencies, rather than to "objects" or the steps in a state-transition picture of change.

Despite its application to the extant domain, LIR is neither a physics nor a cosmology. It is a logic in the sense of enabling stable patterns of inference to be made, albeit not with reference to propositional variables. LIR resembles inductive and abductive logics in that truth preservation is not guaranteed. The elements of LIR are not propositions in the usual sense, but probability-like metavariables as in quantum logics. Identity and diversity, cause and effect, determinism and indeterminism, and time and space receive non-standard interpretations in this theory.

The principle of dynamic opposition (PDO) in LIR extends the meaning of contradiction in paraconsistent logics (PCL), *defined* such that contradiction does not entail triviality. LIR captures the logical structure of the dynamics involved in the non-separable and inconsistent aspects of real phenomena, e.g., of thought, referred to in the paraconsistent logic of Graham Priest. LIR thus applies to all real dualities, between either classes of entities or two individual elements. Examples are theories and the data of theories, or facts and meaning, syntax and semantics. Others are interactive relations between elements, relations between sets or classes of elements, events, etc., and the descriptions or explanations of those elements or events.

LIR does not replace classical binary or multivalued logics, including non-monotonic versions, but reduces to them for simple systems. These include chaotic systems which are not mathematically incomprehensible but also computational or algorithmic, as their elements are *not* in an adequately contradictorial interactive relationship. LIR permits a differentiation between 1) dynamic systems and relations *qua* the system, which have no form of internal representation (e.g., hurricanes), to which binary logic can apply; and 2) those which do, such as living systems, for which a ternary logic is required. I suggest that the latter is the privileged logic of complexity, of consciousness and art, of the real mental, social, and political world.

ORTHO-DIALECTIC CHAINS OF IMPLICATION

The fundamental postulate of LIR and its formalism can also be applied to logical operations, answering a potential objection that the operations themselves would imply or lead to rigorous non-contradiction. The LIR concept of real processes is that they are constituted by series of series of series, etc., of alternating actualizations and potentializations. However, these series are not finite, for by the Axiom **LIR6** of Asymptoticity they never stop totally. However, in reality, processes *do* stop, and they are thus not infinite. Following Lupasco, I will use the term "transfinite" for these series or chains, which are called ortho- or para-dialectics.

Every implication implies a contradictory negative implication, such that the actualization of one entails the potentialization of the other and that the non-actualization non-potentialization of the one entails the non-potentialization non-actualization of the other. This leads to a tree-like development of chains of implications. This development in chains of chains of implications must be finite but unending, that is, transfinite, since it is easy to show that if the actualization of implication were infinite, one arrives at classical identity (tautology): $(e \supset e)$. Any phenomenon, insofar as it is empirical or diversity or negation, that is, not attached, no matter how little, to an identifying implication of some kind, $(\bar{e} \supset e)$ suppresses itself. It is a theorem of LIR that both identity and diversity must be present in existence, to the extent that they are opposing dynamic aspects of phenomena and consequently subject to its axioms.

STRUCTURAL REALISM

Some form of structural realism, such as those developed by Floridi and Ladyman[11] and their respective associates, is also required for a logico-philosophical theory of consciousness of the kind I will propose. In the Informational Structural Realism of Luciano Floridi, the simplest structural objects are *informational objects*, that is, cohering clusters of *data*, not in the alphanumeric sense of the word, but in an equally common sense of *differences de re*, i.e., mind-independent, concrete points of lack of uniformity. In this approach, a datum can be reduced to just a lack of uniformity, that is, a binary difference, like the presence and the absence of a black dot, or a change of state, from there being no black dot at all to there being one. The relation of difference is binary and symmetric, here static. The white sheet of paper is not just the necessary background condition for the occurrence of a black dot as a datum; it is a constitutive part of the datum itself, together with the fundamental relation of inequality that couples it with the dot. In this specific sense, nothing is a datum *per se*, without its counterpart, just as nobody can be a wife without there being a husband. It takes two to make a datum. So, ontologically, data (as still unqualified, concrete points of lack of uniformity) are purely relational entities.

Floridi's informational ontology proposes such partially or completely unobservable informational objects at the origin of our theories and constructs. Structural objects work epistemologically like constraining affordances: they allow or invite constructs for the information systems like us who elaborate them. Floridi's ISR is thus primarily epistemological, leaving the relation to the energetic structure of the universe largely unspecified, even if, correctly, the emphasis is shifted from substance to relations, patterns and processes. However, it points at this level toward the dynamic ontology of LIR in which the data are the processes and their opposites or contradictions.

In the Information-Theoretic Structural Realism of James Ladyman and Don Ross and their colleagues, the notion of individuals as the primitive constitutents of an ontology is replaced by that of real patterns. A real pattern is defined as a relational structure between data that is informationally projectable, measured by its logical depth, which is a normalized quantitative index of the time required to generate a model of the pattern by a near-incompressible universal computer program, that is, one not itself computable as the output of a significantly more concise program. In replacing individual objects with patterns, the claim that relata are constructed from relations does not mean that there are no relata, but that relations are logically prior in that the relata of a relation always turn out to be relational structures themselves.

An area of overlap between OSR and LIR is Ladyman's definition of a "pattern" as a carrier of information about the real world. A pattern is real iff it is projectable (has an information-carrying possibility that can be, in principle, computed) and encodes information about a structure of events or entities *S* which is more efficient than the bit-map encoding of *S*. More simply: "A pattern is a relation between data." Ladyman's position is that what exist are just real patterns. There are no "things" or hard relata, individual objects as currently understood. It is the real patterns that behave like objects, events, or processes and the structures of the relations between them are to be understood as mathematical models.

Lupasco's question "What is a structure?" now appears, but the only answer to it is not a set of equations! The indirect answer of Ladyman and Ross is in terms of science as describing modal structures including unobservable instances of properties. What is not of serious ontological account are unobservable *types* of properties. Thus, seeing phenomena not as the "result" of the existence of things, but their (temporary) stability as part of the world's modal structure, necessity *and* contingency, is something that is acceptable in the LIR framework, provided that the dynamic relation of necessity and contingency is also accepted. There is information carried by LIR processes from one state (of actualization and potentialization) to another, describable by some sort of probability-like non-Kolmogorovian inequalities, although it may not be Turing-computable.

DIALECTICAL LOGICS

Because of the parallels to Hegel's dialectics, logic, and ontology, I have shown in some detail how LIR should be differentiated from Hegel's system.[12] Hegel distinguished between dialectics and formal logic, which was for him the Aristotelian logic of his day. The law of non-contradiction holds in formal logic, but it is applicable without modification only in the limited domain of the static and changeless. In what is generally understood as a dialectical logic, the law of non-contradiction fails. Lupasco considered that his system included and extended that of Hegel. One cannot consider Lupasco a Hegelian or neo-Hegelian without specifying the fundamental difference between Hegel's idealism and Lupasco's realism, which I share. Both Hegel and Lupasco started from a vision of the contradictorial or antagonistic nature of reality; developed elaborate logical systems that dealt with contradiction and went far beyond formal propositional logic; and applied these notions to the individual and society, consciousness, art, history, ethics, and politics.[13]

Among more recent (and lesser-known) dialectical logicians, I include the Swiss philosopher and mathematician Ferdinand Gonseth who discussed the philosophical relevance of experience.[14] The system of Gonseth has the advantage of providing a smooth connection to science through mutual reinforcement of theoretical (logical in the standard sense), experimental and intuitive perspectives. Its "open methodology" refers to openness to experience. The interactions implied in Gonseth's approach can be well described in Lupascian terms. In a prophetic insight in 1975, in his "open methodology" he described the immersion of the individual in "informational processes." (As it turns out, Gonseth was also critical of Lupasco's system, considering it insufficiently rigorous.) More congenial and very much in the spirit of Lupasco was the work of the Marxian Evald Ilyenkov.[15] In a section entitled "The Materialist Conception of Thought as the Subject Matter of Logic," Ilyenkov wrote, "At first hand, the transformation of the material into the ideal consists in the external first being expressed in language, which is the immediate actuality of thought (Marx). But language itself is as little ideal as the neuro-physiological structure of the brain. It is only the form of expression (JEB: dynamic form) of the ideal, its material-objective being."

NON-DUALISM

Non-dualism attempts to relate key insights of Eastern Asian thought to Western thought about life and mind. it establishes a "working" relationship between opposites. Eastern and Western thought processes have been discussed in a series of compendia to which I have contributed.[16] Non-dualism has been criticized as being non-scientific, perhaps for the wrong reasons, but Logic in Reality can be considered a "non-standard" non-dualism in that it recognizes the existence of the familiar physical and meta-physical dualities. However, the additional interactive, oppositional feature it ascribes to them as a logic avoids introducing a further unnecessary duality between it and Eastern non-dualism. Let us now turn to the Lupasco theory of consciousness as such.

5. THE LIR THEORY OF CONSCIOUSNESS

As Lupasco proposed in the mid-twentieth century, the opportunity and the possibility of characterizing consciousness as a complex process, or set of processes, arise from consideration of the details of perception and action.[17] Such consideration allows one to include, from the beginning, a complementary structure of processes that corresponds to what is loosely referred to as the unconscious, to the relation between the conscious and the unconscious, and to the emergence of a second order consciousness of consciousness. Higher level cognitive functions are perhaps easier to characterize as processes than "having consciousness," but consciousness of consciousness is active enough. It remains to demonstrate the evidence for their also resulting from contradictorial interactions of the kind described as fundamental in LIR.

The analysis of the processes of consciousness in LIR starts with that of the initial reception of external stimuli and the consequent successive alternations of actualization and potentialization leading to complex sequences of T-states, as follows:

- An initial internal state of excitation, involving afferent stimuli.
- An internal/external (subject-object) state in which afferent and efferent (motor) mechanisms interact.
- The above states interacting in the brain to produce higher level T-states: ideas, images, and concepts.
- Further interactions leading to consciousness and unconsciousness (the unconscious) as T-states, memory, and forgetting.
- At the highest level, the emergence of consciousness of consciousness, knowledge, intuition, and overall psychic structure.

The originality of this picture does not reside in its identification of a consciousness, a consciousness of consciousness (sometimes designated as awareness), and an unconscious. Rather, it is in its emphasis on the logical origin of these higher-level dynamic structures in a principle of opposition at the level of basic physics that provides the mechanism for their emergence and the subsequent complexification of their interactions. Thus, it can be shown that there are, in addition and as a consequence, three (types of) the other mental processes of memory, forgetting, imagination, and creativity. Only *via* a system complex enough to incorporate these aspects might one be able to arrive at a meaningful, real dynamics of consciousness.

To try to disentangle the various issues involved in consciousness and cognition, I therefore will first position some of the entities involved in this picture of consciousness in the key categories of the ontology of LIR:

- Energy: light; thermal, chemical, and electrochemical gradients;
- Process: chemical and ion flows; chemical synthesis; structural changes of molecules; actions and behavior; remembering and forgetting;
- Dynamic Opposition: activation/excitation and passivation/inhibition;
- Subject and Object: the phenomenological subject-object;
- T-states and Emergence: control states; feelings; concepts; ideas.

THE DIALECTICS OF AFFERENT AND EFFERENT SYSTEMS

The next step is to look in more detail at the dialectics, in human perception, of afferent nerve impulses moving from peripheral receptors toward the central nervous system and efferent impulses moving toward the peripheral, especially, motor systems. Prior to excitation by internal or external stimuli, let us assume that the afferent system is in a state of potentiality, maintained by the antagonistic actualization of the polarization or electrostatic equilibrium.[18] Excitation results in a new actualization, potentializing the ionic equilibrium, the reception of an equivalent to a heterogeneity of sensations. The new equilibrium state of perception appears, in its homogeneity, as something objective, exterior, an identity of which one can have "knowledge," while sensations, although really belonging to the external world, *appear* interior to the senses and more subjective. The dialectics established in and by the afferent process is, accordingly, between the conscious mind, the "knower" as such, actualizing a series of energetic heterogeneities, and the "known" displaced to the exterior, in the potentiality of energetic homogeneity. This conception could be called "pan-energetics," but it is not a pan-psychism; the mind appears as an aspect of the structuring and functioning of energy itself, like the physical and biological, but admittedly the most complex one.

Following re-equilibration (re-polarization) of the excited nerve cells in a T-state, efferent stimuli leave the brain in the direction of organs of movement (of course, with the possibility of many intermediate feedback loops), with a dialectics that is the inverse of the afferent system. Its actualization looks like a *plan*, an operation of active structural homogenization, which will be opposed by the heterogeneity of the external world in which it will operate, and the dialectics involve thus the imposition of this plan on the external world, and the potentialization of this heterogeneity. There is thus a dialectic of the contradictory and antagonistic dialectics of perception and action, which implies, since one does not exist without the other, that each succeeds the other, but neither is very far, in the nervous system, from the T-state. The difference between actualizations that potentialize and potentializations that actualize is not continuous, and the pauses in the process, in the T-state, are what can be considered states of control. These constitute the dialectic of the psyche itself, which *becomes* what is generally called consciousness.

CONSCIOUSNESS AND UNCONSCIOUSNESS AND THEIR DIALECTICS

The next step in the explanation is to identify what appears in the most primitive consciousness as the objects capable of satisfying physiological and biological needs—food, the sexual partner—*in potential form*, through the actualization of the biochemical phenomena of those needs. The consciousness of hunger is not the *consciousness of* an alimentary need, but *is* the need in a potential state. The actualization of this need projects, by antagonism and contradiction, the missing objects into potentiality, and it is this potentiality that is or constitutes consciousness. In other words, the same concept of parallelism is to be rejected here as in the case of energy itself: not here energy and there its properties, intensity and extensity. The needs, the operations of biological satisfaction are not on one side and the consciousness of those needs on the other mediated by some enigmatic entity, leading to the common, but misleading expression "consciousness of." There is, in the LIR theory, no such "consciousness of," no reification or objectification, only *that* which occupies the conscious mind, *that* which is potentiality itself is what is commonly called consciousness.

A potentiality is a conscious energetic state that contains *that* which will be actualized, the need, and its opposite, the lack of which is the need, and is unconscious. However, in contrast to standard theories of mind, there is no actualized structure corresponding to the conscious mental state: it *is* delocalized potentialities. When the lack (hunger) is replaced by the sensation of satiety, the missing elements (food) are eliminated from consciousness and replaced by the potentiality of satiety, which in turn creates a consciousness of satiety and rejection of operations leading to food intake.

What it is actualized, then, does not disappear totally but disappears into the unconscious mind. The next step is to see that there were present, on the one hand, the consciousness of the need and what the need required for its actualization and the unconsciousness of the lack of this requirement, and on the other hand, consciousness of the satisfaction by the actualization and consequently unconsciousness of the disappearance of the lack. Thus, there are two consciousnesses and two un-consciousnesses that alternate. There is a constant dialectical movement between what occupies and *is* consciousness and that, which, by its actualization, leaves the domain of potentiality and "falls" into the unconscious. But, as always, these moves are never total; there is always some potentiality or relative consciousness in unconsciousness and vice versa.

The extent of movement into the unconscious is normally inversely proportional to the importance of the event. After locking my car, I will in general not find it necessary to remember that I locked it, but there is always a probability of the belief that I did not lock it. If this is actualized, I will go back and check it unless I remember enough of the diffuse (diverse) circumstances at the time of my locking to convince me I did so. The dualism in consciousness can be captured in the example of perception of an object, say, a chair, which shows at the same time how the concepts of internal and external can be understood. When one is conscious that something is a chair, one says that one is conscious of *it*, rather than of its detailed form and color. In this view, I "am" the chair in a potential state; it is the potentiality of the chair *qua* chair that is the content of my consciousness of the chair, that is, consciousness itself. But the identity, permanence, and so on of the chair are also actual, although I am unconscious of them. Everything happens as if the chair were my representation of it and at the same time external to me. There is double consciousness and double unconsciousness, of an external world as if made up of objects, that is, of identities, whose location is my consciousness, and an external world of sensations, actualizations of my sense organs, which, as actualizations, disappear into my unconscious. My consciousness is polarized by the object of perception, but this object is only potentialized relative to my senses. The key difference between this description of external and internal reality is in the relation of the internal representation with the potentiality that appears in the perception of the object, "of the chair, of this chair, in back of the heterogeneous actualizations of my senses, my receptors and brain centers, contradictorially associated with the chair, with this chair." (These are the examples essentially as presented by Lupasco in reference 27.) This consequently permits the elaboration of a complex system of two consciousnesses of homogeneity and two sub-consciousnesses of heterogeneity, one of each stronger and the other weaker, succeeding one another dialectically. Table 4.2 illustrates this.

Table 4.2

	Stronger	**Weaker**
Consciousness (of homogeneity of)	Object of perception (potential)	Afferent stimuli
Subconsciousness (of homogeneity(ies) of perception)	Afferent stimuli (actual)	Object

The two inverse dialectics, of consciousness and sub- or un-consciousness, are themselves antagonistic and contradictory, involved in a dialectic, succeeding, interfering, and dependent on one another, with the result that there exist, in a waking state, no less than *eight* mental structures, four consciousnesses and four subconsciousnesses, with different gradients of homogeneity and heterogeneity.

CONSCIOUSNESS OF CONSCIOUSNESS

I suggest, as have many others, that what distinguishes individual human awareness from animal or primitive consciousness is consciousness of consciousness. The best-known formulation of self-awareness is Descartes's *cogito, ergo sum*. Descartes also said, "we cannot doubt of our existence while we doubt," but Lupasco emphasized that it is through doubting that one becomes conscious of thought and therefore conscious of one's consciousness. Lupasco said, specifically, *dubito ergo sum*. Doubting implies being aware of oneself as the locus of the contradictory consciousnesses referred to above, and of their T-states of the semi-actualization and semi-potentialization of each, which also includes the corresponding processes in the subconscious. One then possesses, in effect, two consciousnesses, each of which is aware of the other, of their contradiction, of their antagonism and accordingly of themselves, through a consciousness of consciousness, via an internal dialectics of control. (Such a control state is, admittedly, an hypothesis; no control state has been identified, although it may be implied by recent work on latency and response times.) This dialectic of dialectics is thus at the same time a dialectic of consciousness of consciousness and consciousness of subconsciousness, and constitutes what is generally called the mind or psyche as such.

The "sequence" of events in consciousness in this picture is the following:

Level 1: When a set of perceptions is actualized by the afferent system, two things happen: the heterogeneous actualizations as such, which constitute a primitive subject, without self-awareness, disappear into an unconscious (or subconscious, SC1A). The corresponding potentializations constitute a primitive consciousness C1A, also lacking self-awareness, in which the perceptions appear as largely homogenous objects, OA. When a set of actions is initiated by the efferent system, the related, actualized homogeneous plan of action becomes another subconscious (SC1E), and its heterogeneous objects OE constitute another consciousness C1E. At this level, the resting state of equilibrium in the absence of afferent and efferent influx is defined as a T-state of control (see above).

Level 2: Self-awareness develops out of the dynamic opposition between the above two consciousnesses C1A and C1E, at the point of equilibrium of semi-actualization and semi-potentialization of each, producing, *always as an energetic pattern*, a T-state which is a consciousness of consciousness. Interaction of the latter with the unconscious or sub-consciousnesses SC1A and SC1E result in a consciousness of sub-consciousness.

Level 3: The two dynamically opposed general subjects of Level 1, SA and SE, and the corresponding general objects OA and OE are not isolated entities, but overlap and interact. At Level 2, via the corresponding semi-actualizations and semi-potentializations, a consciousness of the consciousness of the subject and object will develop as well as the corresponding consciousness of the unconsciousness of subject and object. Out of these and their related dynamic oppositions develop the higher functions of images, concepts, and creativity. Memories are present as delocalized potential events in the consciousness of consciousness and as re-actualized events in the consciousness of subconsciousness.

A key property of the interactions that I have described as obtaining between a real mental element and its opposite or contradiction, and between both and any emergent included middle (T-state) is bidirectionality. Since all elements, conscious or sub-conscious, are present in the same configuration space-time, the LIR picture of the reciprocity between A and non-A, A potentializing non-A followed by non-A potentializing A is a description of a bidirectional process. In the two-level LIR framework, the two elements may be at different levels. This picture finds support in the apparent irreducibility of sub-conscious psychophysical computation to neuronal brain activation. Bentwich has found that while most models of brain computational processes propose that neurochemical activity causes cognitive, behavioral, or physiological processes (PCP), the opposite does not take place.[19] In one case, he shows that this assumption results in a contradiction. This leads to the conclusion that PCP takes place at another (higher) computational level that is not reducible to the lower neural level and has received the term "Duality Principle." Bentwich suggests that the Duality Principle may apply to other brain-related computational processing. Although I criticize some aspects of computational models in relation to intentionality, this work fits closely the dynamics that I propose between the potentialized elements in the brain and the actualized, observable ones at the neuronal level. A duality principle of this kind should be a preferred heuristic.

For comparison with the LIR system, Table 4.3 lists the four types of consciousness defined by Daniel Dubois in 1990.[20]

Although there are a number of differences with the LIR view of consciousness, I fully agree with Dubois's statement that global consciousness is constituted by interactive loops between the different types of consciousnesses. The LIR types are differentiated according to the ontological features of identity and diversity, and development from the systems of perception or action, rather than a left-right brain division. The other difference is that in LIR, one is never fully conscious of acts, and there is no separate self to be conscious of either. At this point, I can only say that the LIR view is not incompatible with non-reductionist informational approaches to the fundamental information processing components of brain function such as those of Pedro Marijuan.[21] The "topodynamic" duality principles leading to minimization of the ratios of excitation to inhibition allows for continuity between the nervous system, the cellular signaling system, and consciousness. To conclude this rapid overview, it is clear that only a phenomenological differentiation, at least as complex as that which Dubois and I have described, will enable a potential relation of consciousness to the underlying neuroscience to be made.

Table 4.3

Primary Localization: Left Hemisphere	**Primary Localization: Right Hemisphere**
Objective Psychological Consciousness	Subjective Psychological Consciousness
Consciousness of Acts Meta-Consciousness = Consciousness of Consciousness = Conscious Consciousness	Consciousness of Self Meta-Self-Consciousness = Self-Consciousness of Self-Consciousness
Unconsciousness = Unconscious Consciousness	

MEMORY AND FORGETTING

No theory of consciousness is complete or acceptable unless it accounts for memory, images, concepts, qualia, intentionality, and creativity, and I will just make a few comments here on the LIR view of memory. The contradictorial picture of the processual interactions, in the brain, of macrophysical, biological, and neuro-psychical systems provides the basis for a new explication of memory, in which a distinction is made between conscious "information" or remembrance (souvenir) consciously present and memory-as-such. In the LIR point of view, memory is not a receptacle, a box containing past events in potential form, and the souvenir their actualization, which springs up all at once for one reason or another into consciousness. Memory is the actuality and actualization themselves, albeit, as actualizations, in the unconscious. The souvenir, on the other hand, is the potential event as it emerges in consciousness, occupies, and creates it. Memory and souvenir are thus also antagonistic and contradictory to one another.

In the LIR picture, like that of Bergson, one is subject to the constant interference and antagonism, a dialectics of dialectics of memory and souvenir in opposition, of the two physical (mechanical) and biological (organic) memories that correspond to the chemical and neurophysiological constitutents of the brain plus a third cognitive memory, constituted by the consciousness of sub-consciousness and consciousness of consciousness in a T-state, an emergent included middle. This third memory is equivalent to self-awareness, a memory "that knows it is a memory." It is not to be found as an actualized structure but as an incessant internal contradictorial process, greatest when afferent and efferent operations are cut off from external, relatively non-contradictory contacts.

A forgetting is also a neuro-energetic process activity that prevents an actualization in the unconscious or a semi-actualization in the subconscious from potentializing itself, that is, emerging at the level of consciousness, becoming conscious. There are, accordingly, three forgettings, one for each of the three kinds of memories: the forgetting of identities, actualized by unconscious homogenizing forces; the forgetting of variations and diversities actualized by unconscious heterogenizing forces; and a third forgetting at the same time of the identities and diversities in the subconscious, even though the relevant force of actualization is only a semi-actualization in the T-state. As noted below in the section on creativity, it is from this third memory and its corresponding forgetting that emerge discoveries and inventions, all the riches of the creative imagination, all the new combinations of images, concepts, and ideas.

Mechanistic biology looks for a precise location, a fully actual location for memory. Edelman's neural network picture is one of many possible examples from current neuroscience and cognitive science. One general problem with this picture is the well-known turnover of brain structure at the molecular level; something different must be maintaining the relative stability of the memory, and I see this as explained by the LIR conception of the persistence of non-localized but also potential properties of the physico-chemical brain structures themselves and of the higher levels built up from them, as suggested above.

THE BENNETT AND HACKER VIEW OF CONSCIOUSNESS

The Lupasco logic of real processes allows an interpretation of many of the criteria proposed by Bennett and Hacker for a theory of consciousness, without getting into the details of their neurological model.[22] Bennett and Hacker focus on the human being as a psychophysical unity, without attributing thought or knowing to the brain or its parts, such as its hemispheres. The following summarizes where I have found their approach useful:

- Mind and Self

  The self, defined as something that is identical with me, as something I have or as something in me is an aberration. There is no such thing, and "I" does not refer to an "Ego" owned by me. One has, as arguments, the formal one from infinite regress and the phenomenological one from our existence as human beings, not brains or minds. There is no such thing as my perceiving, rather than having, my own thoughts. The LIR appearance/reality dialectics is useful here. The mind is not an entity or a thing or a "domain"; this term refers idiomatically to a wide range of human powers and their exercise.

- Representation

  In this view, it is a mistake to say that what we or some "mind" perceive is an image or representation of an object, or that perception involves *having* an image of the object. The so-called binding problem is a false problem, since the brain does not construct a perceived world, but enables an animal to see a visible scene. Damasio was mistaken in his distinction between having and feeling an emotion, as if emotions were some sort of somatic image or marker.

- Qualia

  The term of "qualia" was introduced to signify the alleged private character of experience, its phenomenal qualities, or qualia for short. This led to Nagel's strategy of explaining the subjective or qualitative feel of experience in terms of their being something it is like to have it. However, as Bennett and Hacker show, neuroscientists such as Damasio and Edelman shift the sense of the term "quale" from the qualitative character of experience to the qualitative character of objects. The term "quale" equivocates between what it is like to have an experience and the experience itself. The indexical approach clarifies the problem, although in the LIR conception of human psychological types, it will never convince everybody. The question "Why is seeing red like seeing *this* (Wittgensteinian pointing to a sample) is misguided because seeing red does not *resemble* seeing *this*, it *is* seeing this. The alleged incommunicability of the subjective qualities of an experience is confused. One cannot describe a quality in the same way as one describes an object by specifying its qualities; what one needs is a better vocabulary. A description is not a substitute for experience.

There are two points where one can criticize the Bennett and Hacker approach: (1) I agree that to perceive is not to form a hypothesis or make an inference, but I disagree that inferences are not mental processes, but transformations of propositions in accordance with a rule. LIR extends inference to process. (2) The authors say correctly that it is the task of neuroscience to investigate empirical nature of consciousness, while that of philosophy is to elucidate its defining its concepts and connections with related ones such as anticipation, thought, and so on. They also say, however, that philosophy can contribute nothing to the scientific theories about the neural basis of consciousness, although the two activities are complementary, not competitive or mutually exclusive. Perhaps standard philosophy cannot, but I submit that the logic and metaphysics of LIR cannot be separated from science; complementarity implies interaction, and the concepts of LIR are thus pertinent to a science of consciousness.

Intransitive consciousness is a condition for various forms of occurrent transitive consciousness—that is, for being conscious of something at a given time. Transitive consciousness is a form of knowledge, or, preferably, knowing. Above all, what one is perceptually conscious of is not something over and above some of the things one perceives. One is conscious of what occupies one's mind at a particular time, and Lupasco emphasized the dynamics of *change* from conscious to unconscious as one from (primarily) actual to (primarily) potential.

Bennett and Hacker urge us to avoid taking "mysterian" positions that start by *trying* to see First Person Consciousness as outside nature. It is the richness of cognitive processes in human beings that are the *logical criteria* for a creature's being conscious. "Cognitive neuroscience operates across a categorial 'divide' between the psychological and the neural (which is a particular case of the physical). There is nothing mysterious about this divide." It is constituted by the logico-grammatical differences as well as the connections between the characteristic concepts of neuroscience and those of psychology. "Cognitive science has constantly to cross this logical divide." This is ample justification, in my mind, for a reworking of the underlying logic.

6. REVISITING SOME KEY QUESTIONS

THE MAJOR APPROACHES OF COGNITIVE SCIENCE. REPRESENTATIONALISM

The major, related approaches of cognitive science to the phenomenological data of consciousness are as follows:

- *Representationalism*, according to which internal mental entities stand for or correspond to real external properties and events. Representationalism includes the next two approaches, in which the entities involved are the symbols and properties, respectively.
- *Computationalism*, the view that thinking is basically a form of computation in the sense of computer science, an algorithmically determined process of manipulation of symbols in a neural network;
- *Connectionism*, which sees cognitive function as the operation of the system of neural networks, not with isolated symbols, but with vector distributions of properties according to a dynamical model and following rules for non-linear dynamic systems.
- *Functionalism*, of which computationalism is a variety, the view that thinking is wholly defined by its function in a physical system as it interacts with other internal and external processes.

These approaches, alone or in various combinations, all seem to me to have one or more of the following weaknesses and, accordingly, are fair targets for debate:

1. Reification of consciousness, neglecting its *process aspects*, equivalent to a classical substance ontology;
2. Reference to actual entities to the exclusion of potential ones;
3. Functional separation of external and internal aspects of consciousness, despite reference, as in discussions of biological phenomena, to the environment;
4. Absence of adequate complexification of conscious and unconscious processes, suggested by the Lupasco scheme above.

The "*Working Hypothesis of Neurophenomenology*" is that phenomenological accounts of the structure of experience (or structured phenomenological accounts of experience) and their counterparts in cognitive science relate to each other through reciprocal constraints.[23]

Although from my point of view these authors' attempts to find a specific locus for the bridge between the two domains were unsuccessful, the concept that the mutual constraints would need to be operationally generative, that is, *directly* link "appearances" to specific emergent biological processes, points directly toward the positions of what I have called Logic in Reality (LIR). LIR is a way of joining, to use the term used by Roy et al., both the two types of data separated by a "wavy line" at a level of description that is sufficiently general, rather than abstract, to provide for functional interactions between the elements on both sides of the line.

If one assumes, on the other hand, as in the computational form of representationalism, that there is a symbolic entity between neurobiological and phenomenological data, a host of secondary problems arise as to the properties and relations of the symbols involved. In representationalist theories, internal entities of some sort stand for or correspond in some way to external processes and events. These mental representations explain or are explanatory devices for cognition in that they are, or correspond to (this vagueness is typical) intentional states, instances of intentionality considered as embodying the irreducible first-person properties that are alleged to characterize consciousness, reasoning, and qualia. This account of mental processes suffers from the need to introduce additional entities due to the lack of a principled categorial method of relating its critical concepts contradictorially. LIR on the other hand supports not only the "truth" of first-person consciousness[24] but its ontological existence. A mental phenomenon is not something other than the physical processes with emergent properties. It only "displays" its contradictorial origins in appearing to have symbolic and non-symbolic aspects, and being closer or farther from the center of attention at a particular time.

My approach is thus fundamentally anti-representationalist, bringing it into conflict with the semiotics of Peirce and his current followers, especially Sören Brier. I have proposed LIR as an ontological substitute for Peirce's theory of signs both in the field of information and more generally.[25] Very rapidly, I conclude that signs are both 1) ontologically dependent on the phenomena of which they are the signs and 2) incapable of reflecting the dynamic and value-laden interactions involved in real phenomena, such as personal identity.

PERSONAL IDENTITY

For human beings, the concepts of consciousness and personal identity are inextricably linked. I have described the LIR view of consciousness of consciousness, and I can now claim that recent philosophical work by Dan Kolak supports this picture, especially as regards the origin of individual and collective responsibility. The characterization of personal identity is thus a key issue for science as well as philosophy. The logical perspective of LIR, which sees identity also as a process of identification, accompanied by its opposite, permits a naturalization of concepts of personal identity such as that of Kolak. LIR can be seen as a bridge between philosophy and science that places this view in logic and therefore in science and society. The picture that emerges from this analysis is an ethical one. It supports and explicates another of the insights of Dan Kolak in his major book, *I am You; The Metaphysical Foundations for Global Ethics*.[26]

In the June 2008 volume of *Synthese* dedicated to the subject of personal identity, Kolak wrote that "(a) consciousness makes personal identity and (b) in consciousness alone personal identity exists." His analysis of public vs. first-person perspectives, using cases from neuropsychiatry, provides the scientific, mathematical, and logical frameworks for what he calls a new theory of self-reference wherein consciousness, self-consciousness, and the "I" are precisely defined in terms, close to Sartre, of the subject and the subject-in-itself. In Kolak's approach, the critical move is to avoid a separation of the subject that is the bearer of personal identity from its psychological object identifications. LIR supports the argument by providing the rules for the relative, alternating dominance of the two perspectives: personal identity and the intuition of personal identity, the reality of subject-dependence and the appearance of subject-independence of experience are dynamically, dialectically related in the LIR logic. Logical, psychological, and metaphysical perspectives intersect in this view. In LIR terms, Kolak's statement that one's essential subjectivity is obscured by the intuition of one's own existence and identity is that the former is potentialized by the latter. The conjoined personality experience by the subject from the inside as the identified self that expresses itself as "*I am I*," not my brain, not my body, and not even my "self." LIR thus allows a principled *ontological* process view of consciousness. It is constituted by systems of systems of past and present mental processes following the LIR dynamics of alternating actualization an potentialization from which personal identity is constituted as an emergent structure. LIR offers no explanation of why I am *this* I and you are another one, but nobody has yet done so, as far as I know.

KNOWLEDGE AND INTUITION

As Kolak states, the above line of reasoning gives a privileged status to the role, function, and nature of intuition. I would like to expand on it to illustrate the application of LIR to one of the controversial functional properties of human consciousness. The problem is that there has been no obvious way to make an absolute differentiation between knowledge, or knowledge-as-such, and intuition as regards how they arise and their respective functions as protagonists in the drama of knowledge. Let us postulate that knowledge-as-such and intuition or intuitive knowledge are indeed two forms of knowledge or better knowing. Actualization and potentialization constitute, at the same time, the mechanisms of both knowledge and existence (logical becoming), both involving alternation between states in which one term is (almost) fully actualized and then the other is (almost) fully potentialized. Then, as stated by Lupasco in his State Thesis of 1935, given any cognitive process, a logical becoming is involved since knowing means inhibiting one antagonistic factor by another. The knowledge associated with the strongly actualized terms is the identifying knowledge-as-such, the major content of consciousness. The statistical process of oscillation "leaves behind," however, a minor, accidental knowledge or known that can be designated as intuition. Intuition is thus an embryonic non-identity, always an unexpected and brief "irrational" invasion of consciousness, discontinuous, without a direct relation to it. In terms of cognitive power, the difference between intuition and knowledge as such is only one of degree, and their relation can be described by saying that what is given intuitively is the inverse of what is given to knowledge; the content of knowledge is contradictory in the sense of being dynamically opposed to the content of intuition, and the existentiality of one is a function of that of the other. Intuition in this dynamic aspect must be seen as a *logical* process, subject to the rules of LIR applied to knowledge. I claim that both types of intuition, sensible and intellectual, are direct experiences, actions, or processes and have a place in a theory of mind. From the point of view of difference in function, what is primarily retained in the conscious mind is a kind of identity and synthetic rationality, and what constitutes intuition is the knowledge of movement, time, intensity, the heterogeneous, etc. Thus, one does not "see" change itself, but rather one identity replacing another. Change is "felt," i.e., known intuitively. Other functional examples that can be developed are those of intellectual consciousness vs. active consciousness, the first the consequence of vital becoming, where science dominates and intuition is avoided; the second of material becoming and "action," in which intuition is essential and the role of formal knowledge is reduced.

The position taken by Levy regarding the distinction between knowledge-that and knowledge-how supports my anti-propositional view of logic in real processes in general.[27] The argument is succinct: knowledge-how requires both propositional knowledge and motor representations in the mind. But motor representations are not mere dispositions to behavior; they have some representational content. Since that content is not propositional, propositional knowledge is not sufficient for knowledge-how. Neither propositions nor representations are required in the LIR approach: if motor representations play a central role in realizing the intelligence in knowledge-how, or more simply, are a form of knowledge hence of consciousness, the concept of a representation as a separate entity can be replaced by that of process.

ANTICIPATION

Anticipation is primarily a property of conscious living systems. That anticipation can play a role in systems that involve substantial abstract modeling rather than self-representation at their level of reality is simply another case, in my view, of the projection of aspects of the real world, reality, into a configuration space of lower dimensionality. The clearest example of this notion is to be found in the work of Gödel. The Gödel theorems and logic—as written—do not apply to physical or mental emergent phenomena, but LIR views the principle involved, the duality of consistency and completeness, axiomatically, as another instantiation of the fundamental duality of the universe. Gödel rejected, correctly in my view, the more idealist implications of many-world pictures of reality, but did not make the extension of his own ideas to it. The logical and ontological development undertaken in LIR provides a bridge between prior definitions of the principle of dynamic opposition and Gödelian dualism and illuminates Gödelian dualism as another expression of the fundamental dynamic opposition at the heart of energy and phenomena.

I have argued that potential states and processes, of which consciousness is an example, are causally effective and not epiphenomenal. If this is accepted, then the naturalization of anticipation follows logically, at least in my logic. One needs to differentiate, however, between anticipation in living beings and anticipation in machines, or, rather, between anticipatory systems that are and are not computable. I, in fact, assimilate anticipation at the cognitive level to particularly well-formed, homogeneous potential states that are opposed to the general fuzziness of the "stream of consciousness." I differ with Dubois, however, in that I do not assign a separate subjective or objective character to anticipation or to a particular hemisphere to the exclusion of the other. I believe it is important to focus on all high-level properties as properties of the whole human being, of whom the alleged parts are convenient abstractions for analysis. I am thus not saying that there are some anticipatory systems that are not computable, with which I am sure we can all agree. I am saying that conscious anticipation is not fully computable. What distinguishes anticipatory processes is a higher degree of potentiality, but anticipation does not define all processes. Anticipatory processes are thus a sub-class of a broader group of processes that constitute "consciousness."

My key difference with Dubois can be summarized as follows:

- Dubois: anticipation is the potential future value of a system's variables
- LIR: anticipation is the current potential value of some systems' variables

Dubois has criticized Rosen's concept of anticipation as "quasi-anticipation" as failing to account for feedback. The LIR model does not require full predictability. Nevertheless, I am sympathetic with Rosen's intuitions about life in general, expressed in his emphasis on semantic aspects of entailment and organization, but he does not provide a basis for the relational aspects of organization and complexity.[28]

Leydesdorff and Dubois have also looked at anticipation in social systems, but their analytical model is orthogonal to the contradictorial LIR view of individual-group interactions. This basically states that the individual and the group share some of each other's properties.

7. CONCLUSION

I have proposed a logic of and in reality as a new perspective on the nature of consciousness. My arguments have covered issues in fundamental physics, mechanisms of perception, and the emergence of consciousness, and implications of this view of consciousness for man in society. I am aware that the "transport dialectique," to use the term of Gilles Deleuze, may have been a difficult one. My vision of the world and theories of the world as related, consistent, *and* inconsistent conflicts with much received wisdom. I ask, to begin with, that the reader renounce, for the sake of a science of consciousness, some standard (and cherished) notions not only of logic, but also set theory, category theory, causality, and accept concepts from the latest quantum field views of the secondary ontological status of spacetime. The methodology of LIR means looking for structures in nature that are potential as well as actual, in a sense that is neither more nor less than that a certain sequence of amino acids in an enzyme has the potential for binding with specific substrates under the appropriate conditions in the appropriate medium.

In the LIR epistemology, we as knowers are not totally external to what is known by us and not completely different from it. I must know, then, that if there are other knowers, as there are, they must be part of my known and vice versa. The source of human dignity *is* in ourselves as knowers, but if we avoid the error of solipsism, the origin of the sense of moral responsibility can only come from the relation to other knowers, in other words, all human beings, and by extension, other beings and perhaps even, as suggested by Lorenzo Magnani, certain non-living entities. *A contrario*, one cannot find responsibility in oneself as an isolated agent. Since we are both a "not-other" *and* an "other" at the same time, a self-interest argument for morality holds. Two or more human individuals and their relations constitute interactive systems in the LIR categorial sense of non-separable subjects and objects, sharing in part one another's characteristics. An individual is no more isolated logically, psychologically, or morally than he or she is economically. The fact that potential or potentialized states exist does not, in a deterministic universe, mean that we have the capacity to make a choice among them that is independent of our genetic and experiential background. Every individual is indeed unique, but this should not be taken to mean that his or her mind is independent, since each incorporates a portion of the subjective experience of other brains. As Bennett and Hacker point out, our ability to know the states of other persons' minds is not folk psychology, but a natural consequence of the evolution of our species. LIR simply adds the logical consequence as an origin of individual moral responsibility.

My claim is that the LIR contradictorial picture of consciousness is a *form* of identity theory of mind which avoids the difficulties of both standard identity and dualist theories by the introduction of the principle of dynamic opposition at all levels of perception, mental processing, and action. No new, independent entities of the kind postulated in the various forms of representationalism are required, due to the availability, in LIR, of a dynamic relation between internal and external, actual and potential, and identical and diverse aspects of phenomena. It is the alternating actualizations and potentializations derived from initial energetic inputs that *are* our ideas, images, beliefs, etc. Some further phenomenological classification of these process elements (such as that made by Husserl) is possible, but it does not change the overall structure of my proposed picture. One of my objectives to further this work is to find mathematical formalizations of the systems aspects of LIR that would render them both more accessible and more rigorous. I would be grateful for suggestions along these lines.

NOTES

1. Chalmers,"Consciousness and Its Place in Nature."
2. Stubenberg, "Neutral Monism."
3. Boltuc, "First-Person Consciousness as Hardware."
4. Wu and Brenner, "Philosophy of Information: Revolution in Philosophy. Towards an Informational Metaphilosophy of Science."
5. Smart, "The Identity Theory of Mind."
6. Floridi, *The Philosophy of Information*; Wu, "The Interaction and Convergence of the Philosophy and Science of Information"; Wu and Brenner, "Philosophy of Information."
7. Ehresmann, in *Memory Evolutive Systems. Hierarchy, Emergence, Cognition*.
8. Bickhard, "Mind as Process."
9. Bickhard, "Process and Emergence: Normative Function and Representation."
10. Brenner, *Logic in Reality*.
11. Ladyman, *Every Thing Must Go*.
12. Brenner, *Logic in Reality*.
13. Taylor, *Hegel*.
14. Pouget, Le rôle de la contradiction dans l'œuvre de Ferdinand Gonseth; Cohen-Tannoudji, "Ferdinand Gonseth."
15. Ilyenkov, *Dialectical Logic; Essays on Its History and Theory*.
16. Brenner, "Linking the Tao, Biomathics and Information Through a Logic of Energy."
17. Lupasco, *Du Devenir Logique et de l'Affectivité, Vol. 2. Essai d'une Nouvelle Théorie de la Connaissance*. Paris: Vrin, 1935.
18. Lupasco's expression was "electro-chemical gradient of equilibrating antagonism."
19. Bentwich, "The Duality Principle: Irreducibility of Sub-threshold Psychophysical Computation to Neuronal Brain Activation."
20. Dubois, "Breakthrough in the Human Decision Making Based on Unconscious Origin of Free Will."
21. Del Moral, "From Genomics to Scientomics: Expanding the Bioinformation Paradigm."
22. Bennett and Hacker, *Philosophical Foundations of Neuroscience*.
23. Formulated by Roy et al., "Beyond the Gap. An Introduction to Naturalizing Phenomenology."
24. Petitmengin, "Listening from Within."
25. Brenner, "On Representation in Information Theory."
26. Dan Kolak, *I Am You. The Metaphysical Foundations for Global Ethics*.
27. Levy, "Embodied Savoir-Faire: Knowledge-how Requires Motor Representations."
28. Rosen, *Life Itself*.

BIBLIOGRAPHY

Bennett, M. R., and P. M. S. Hacker. *Philosophical Foundations of Neuroscience*. Cambridge, MA: Blackwell Publishing, 2003.

Bentwich, Jonathan. "The Duality Principle: Irreducibility of Sub-threshold Psychophysical Computation to Neuronal Brain Activation." *Synthese* 153, no. 3 (2006): 451–55.

Bickhard, M. "Mind as Process." January 2002. https://www.lehigh.edu/~mhb0/mindasprocess.pdf.

Bickhard, M. "Process and Emergence: Normative Function and Representation." In *Process Theories: Crossdisciplinary studies in dynamic categories*, edited by J. Seibt, 121–55. Netherlands: Kluwer Academic Publishers, 2003.

Boltuc, P. "First-Person Consciousness as Hardware." *APA Newsletter on Philosophy and Computers* 14, no. 2 (2015): 11–15.

Brenner, J. *Logic in Reality*. Dordrecht: Springer, 2008.

———. "On Representation in Information Theory." *Information* 2 (2011): 260–78.

Brenner, J. "Linking the Tao, Biomathics and Information Through a Logic of Energy." *Progress in Biophysics and Molecular Biology* 131 (2017):16–33.

Chalmers, D. "Consciousness and Its Place in Nature." In *Philosophy of Mind: Classical and Contemporary Readings*. Oxford: Oxford University Press, 2002.

Cohen-Tannoudji, G., and E. Emery. "Ferdinand Gonseth." In *Dictionnaire des philosophes*. Paris: Presses Universitaires de France, 1993.

Del Moral, R., M. González, J. Navarro, and P. Marijuán. "From Genomics to Scientomics: Expanding the Bioinformation Paradigm." *Information* 2 (2011): 651–71.

Dennett, Daniel. "I Could Not Have Done Otherwise – So What?" *The Journal of Philosophy* LXXXI 10 (1984): 553–67.

Dubois, D. "Breakthrough in the Human Decision Making Based on Unconscious Origin of Free Will." *Acta Systemica* X 1 (2010): 13–18.

Ehresmann, A., and J.-P. Vanbremeersch. In *Memory Evolutive Systems. Hierarchy, Emergence, Cognition*. Amsterdam: Elsevier, 2007.

Floridi, L. *The Philosophy of Information*. Oxford: Oxford University Press, 2010.

Floridi, L. "Singularitiarians, Altheists, and Why the Problem with Artificial Intelligence is H.A.L. (Humanity At Large), not HAL." *APA Newsletter on Philosophy and Computers* 14, no. 2 (2015): 8–11.

Ilyenkov, E. *Dialectical Logic; Essays on Its History and Theory*, translated by H. Creighton. Moscow: Progress Publications, 1977 (in English).

Kolak, D. *I Am You. The Metaphysical Foundations for Global Ethics*. Dordrecht: Springer Books, 2005.

———. "Room for a View: On the Metaphysical Subject of Personal Identity." *Synthese* 162 (2008): 341–72.

Ladyman, J., and D. Ross. *Every Thing Must Go*. Oxford: Oxford University Press, 2007.

Levy, N. "Embodied Savoir-Faire: Knowledge-how Requires Motor Representations." *Synthese 194* (2017): 511–30.

Lupasco, S. *Du Devenir Logique et de l'Affectivité, Vol. 2. Essai d'une Nouvelle Théorie de la Connaissance*. J. Vrin: Paris (1973).

———. *Le principe d'antagonisme et la logique de l'énergie*. Paris: Editions du Rocher, (1987). (Originally published in Paris: Éditions Hermann, 1951).

Petitmengin, C., and M. Bitbol. "Listening from Within." *Journal of Consciousness Studies* 16 (2009): 10–12.

Pouget, P. Le rôle de la contradiction dans l'œuvre de Ferdinand Gonseth. *Bulletin de l'Association F. Gonseth* 166 (2017): 15–28.

———. *L'énergie et la matière psychique*. Paris: Editions le Rocher (1987a). (Originally published in Paris: Julliard, 1974).

Rosen, R. *Life Itself*. New York: Columbia University Press, 1991.

Roy, J.-M., et al. "Beyond the Gap. An Introduction to Naturalizing Phenomenology." In *Naturalizing Phenomenology*, edited by J. Peitot et al. Stanford: Stanford University Press, 1999.

Smart, J. "The Identity Theory of Mind." *The Stanford Encyclopedia of Philosophy* (Fall 2000 Edition), edited by Edward N. Zalta. http://plato.stanford.edu/archives/fall2000/entries/mind-identity/.

Stubenberg, L. "Neutral Monism." *The Stanford Encyclopedia of Philosophy* (Summer 2018 Edition), edited by Edward N. Zalta. http://plato.stanford.edu/archives/sum2018/entries/neutral-monism/.

Taylor, C. *Hegel*. Cambridge: Cambridge University Press, 1975.

Wu, K. "The Interaction and Convergence of the Philosophy and Science of Information." *Philosophies* 1 (2016): 228–44.

Wu, K., and J. Brenner. "Philosophy of Information: Revolution in Philosophy. Towards an Informational Metaphilosophy of Science." *Philosophies* 2 (2017): 1–30.

A Counterexample to the Church-Turing Thesis as Standardly Interpreted

(Theoretical outline and technical results)[1]

Doukas Kapantaïs
ACADEMY OF ATHENS, RESEARCH CENTRE FOR GREEK PHILOSOPHY

The standard (extensional) interpretation of the Church-Turing thesis is that every mechanically calculable function is Turing Machine computable. A counterexample to this interpretation is that proof of the totality of a function ***K*** (from Kapantaïs) constructed upon addition, iteration and a family of programs on programs can be performed by a human, but not a Turing Machine. The human computer, assisted by some instructor, can construct a hierarchy of functions with broad similarities to the finite part of existing fast growing function hierarchies on the sole basis that addition is a mechanical item, which can be mechanically operated on according to some mechanical programs on programs and iteration. She can then calculate the values for any assignments to functions of this hierarchy, and can also arrive at proving that function ***K*** (a function that lies outside of the same hierarchy) is total. No Turing Machine can do the same, since a Peano Arithmetic proof of the totality of function ***K*** would have required Induction up to ε_0.

BACKGROUND

Kapantaïs 2016 criticized a strong (intensional) interpretation of the Church-Turing thesis, according to which what the Thesis says is that everything which can be effectively calculated can be computed by a Turing Machine equivalent, with all maximal models of computation being equivalent to Turing Machines up to isomorphism.[2] This interpretation has been criticized on the grounds that a human computer (computor, henceforth)[3] can compute the values of the Ackermann function in a mechanical way, which cannot be translated into a Turing Machine computation in an isomorphism preserving manner.[4] The proof that these specific calculations by the computor and the Turing Machine are not isomorphic rests on two particular assumptions. The first is the clearly true presumption that addition is a mechanical item, which can be mechanically operated on. The second is that a "for loop" mechanical operation that initially operates upon addition exists and iteration of it yields mechanical operations for all functions upon the Knuth up-arrow notation hierarchy. Such an operation can clearly be specified and justification given for why it should be regarded as mechanical. On these assumptions, it can be proved both that (i) a computor can employ a program made out of "for loops" uniquely for effectively calculating the values of the Ackermann function and (ii) that no Turing Machine can employ a program made out of "for loops" uniquely for the same task. Since there is a way a calculator can compute the values of the Ackermann function, which has no isomorphic counterpart in Turing Machines, there are models of computation, which are not Turing Machine equivalent up to isomorphism.

The present proof builds on a counterexample to the standard (extensional) interpretation of the Church-Turing thesis. The standard interpretation does not contain the claim that all maximal models of computation are isomorphic. Its sole claim is that Turing Machines is one of them. That is to say that what matters concerning the extensional interpretation is the output(s) of the effective calculation(s), not how one arrives at this (these) output(s).

The counterexample in the present proof rests upon a generalization of the algorithm in Kapantaïs 2016. It is a counterexample to the Thesis in its standard interpretation insofar as it shows that a proof of the totality of a function ***K*** can be performed by a computor but not by a Turing Machine.

STIPULATIONS ABOUT THE CONCEPT OF A MECHANICAL PROCESS

Following Turing, Post, Kleene, and others, the terms "mechanical" and/or "effective" remain formally undefined in our proof.[5] "Mechanical/effective" is taken to be the intuitive pre-formal notion all putative (maximal) models of computation aim at capturing. The test of whether they capture it or not must always be decided according to actual evidence. That is to say that in case one has no empirical evidence of a mechanical process that outbids Turing Machines one is compelled to abide with the Thesis, but, once such a model is put forward, one has to abandon it. This also means that the Thesis can only be refuted and never be proved, which, in turn, implies by no means that the Thesis might not be true. If it is true, no such model will ever be found and we will always be compelled to abide with the Thesis. However, we will never be in position to ascertain this fact from within our historical perspective.

Pre-theoretical concepts are not formally defied, and yet pre-theoretical concepts are systematically discussed, which means that the absence of a formal definition over a concept does not condemn one to say nothing about the concept. As far as "mechanical" is concerned one is far from being at a loss on what to say. Direct evidence for the above is that there are lots of processes, about which one can be certain that they are mechanical, and lots of processes, of which one can be certain that they are not. Hence, even if the concept is formally undefined, we take it for granted that there is (ought to be) some general consensus on what is *definitely* mechanical and what is *definitely* nonmechanical.

Stipulations 1 and 2 below aim at reflecting part of this general consensus on what is definitely mechanical and what is definitely nonmechanical. Keep in mind, however, that we do not suggest that each one of these or their conjunction is a *definition* of "mechanical"—even more so, since the term "mechanical" appears both in subject and predicate position in some of these *Stipulations*.

We distinguish between "mechanical" as an object predicate ("mechanical item") and "mechanical" as a process predicate ("mechanical operation").

Stipulation1: All mechanical items are such that they can be manipulated by purely mechanical means.

Examples of mechanical items according to *Stipulation1*:

(a) Natural numbers in the form of finite sequences of strokes are mechanical items. Apprehending such a sequence and mechanically operating upon it does not require any interpretation of what the sequence stands for.

(b) Computer programs are mechanical items. They signify nothing (at least for the machine) and yet the machine can, by purely mechanical means, "understand" what they expect it to do and do it.

Stipulation2: All mechanical operations are such that (i) they engage purely mechanical means, e.g., no appeal to intuition, only finitary methods etc., (ii) the items they operate upon are mechanical, and (iii) the items they transform the items they operate upon are mechanical.

Examples of mechanical operations according to *Stipulation2*:

(a) Turing Machine numerical operations.

(b) Mechanical operations on programs.

Examples of nonmechanical items and operations according to *Stipulations1-2*:

(a) Transfinite ordinals. They cannot be apprehended as finite sequences of strokes or by other mechanical/finitary means.

(b) Implicit representations of transfinite ordinals within Peano Arithmetic by functions. The "representation" is a trick, requiring interpretation, i.e., these functions correspond to transfinite ordinals only for the proof theorist, who knows the trick/interpretation.

(c) Transfinite arithmetical operations.

So, some items/operations are definitely mechanical and some are definitely nonmechanical. Operations/items which are neither clearly mechanical nor nonmechanical must be classified on an individual basis. This applies also to all methods/operations within our proof. To begin with, they certainly do not betray any of the above criteria. Our claim is that they do not betray any other sensible criterion for "mechanical" either.[6]

***

Our proof is performed within the following general setting. An instructor provides mechanical guidance to a computor, in order to enable her to construct a hierarchy of programs and to prove that a program lying outside of this hierarchy returns an output for every input to its input placeholders. This program is shown to calculate the values of a function, which, in order to be proved total by standard formal Arithmetic, requires Induction up to ε_0.

STRUCTURE OF THE PROOF

STAGE 1

Initially, a hierarchy *H* of infinite hierarchies of programs is generated. All programs in *H* are generated upon (i) a preexisting mechanical program for addition, (ii) iteration, and (iii) a family of programs on programs. All programs in *H* are named after a recursively trackable system of coordinates of the $n<\omega$ dimensional space. All programs in *H* correspond to recursive functions (i.e., they calculate values of specific recursive functions).[7] While constructing *H*, the computor mimics *ad indefinitum* the way that the Ackermann function governs the Knuth up-arrow notation hierarchy, and it does so by systematically constructing new hierarchies upon functions that govern previous hierarchies.

A function ***H'*** governing *H* itself is proved total, *relative to all assignments to its variables*. ***H'*** takes as arguments functions in *H* together with assignments to their variables and returns as values the values that these functions yield for the latter assignments. The clause "total, *relative to all assignments to its variables*" is key with respect to this and all our subsequent results.[8] This is why: Most programs in *H* are noneffective, since most of them are either non effective pointers or "for loop" programs depending on noneffective pointers. The pointers in question are noneffective because they have infinite pointing scopes. The way that our calculator proves that functions calculated by these pointers are total is by mechanically showing that *for any assignment to the variables* of the same functions there is a mechanical method to construct an appropriate *for this assignment* finite initial segment of the infinite pointing scope of the pointer in question, such that the pointer becomes effective *relative to this assignment to the variables*. So this is the core of our proof in Stage 1: It shows that there is a mechanical operation such that, on input a function in *H* and an assignment to its variables, makes the computor able to mechanically construct the programs needed for the calculation of the value of the same function for the same assignment.

NB. ***H'*** can be proved total by other means too. The most standard one employs Mathematical Induction up to ω^ω. This proof, instead of turning the noneffective pointers into effective *relative to specific assignments to the variables*, substitutes the pointers for other programs that calculate extensionally equivalent functions, and continues by μ-minimization in order to show these latter total. This is, in a sense, the "canonical" way to prove ***H'*** total. Now, the height of the Induction employed in this, i.e., ω^ω, suggests that ***H'*** as well as all functions governed by it can be proved total by Turing Machines too. So, thus far, this is not a refutation of the Church-Turing thesis, as standardly interpreted.

STAGE 2

We proceed by repeating the same general mechanical routine for the generation of infinite hierarchies of functions, until we reach function ***K***, which, in order to be proved total by Peano Arithmetic, would have required Induction up to ε_0. We prove ***K*** total by Induction $<\varepsilon_0$. We claim that this proof does not betray any among the intuitive criteria for "mechanical", as stated in the first part, and we also claim that it does not betray any other expressly stated and sensible criterion for "mechanical" either.

The proof in Stage 2 proceeds as follows.

Following Stage 1, which ends by the program for function ***H′***, the computor repeats *ad indefinitum* the same procedure. That is to say that, just like in Stage 1 the computor constructs all programs by employing (i) a program for addition, (ii) a family of programs on programs, and (iii) iteration, she now constructs similar programs by employing the same (ii) and (iii), but, this time, she does so upon the program for ***H′***, not addition; she also proves that a function ***H′′*** governing these latter programs is total. This is straightforward because our proof in Stage 1 was relying solely on the fact that the program for addition is effective. Now, since the program for ***H′*** is proved effective, the same procedure can be repeated on the basis of ***H′*** being effective, and end with the proof of ***H′′*** being effective. This whole process/algorithm can be repeated at will, so as to produce the sequence of programs: ***H′***, ***H′′***,…, that governs the sequence of hierarchies: *H*, *H′*,…

Following this sub-stage, another program is constructed, which is a pointer to the entire sequence of the *H*, *H′*,… hierarchies. This program suggests another general process/algorithm for the generation of yet some other hierarchies of programs.

In general, each new and more complex general algorithm for generating hierarchies of programs produces hierarchies that are more complex than the previous ones, and which are governed by more complex programs, and so on *ad indefinitum*. For example, just like the sequence of hierarchies *H*, *H′*,… is governed by a program that is more complex than any of the programs of the same hierarchy, this program suggests a general algorithm that produces a sequence of hierarchies of such hierarchies. These latter are governed by yet another more complex program, which suggests an even more complex general algorithm, and so on.

Each such general algorithm is made to correspond to a distinct general iterative circle along our way to function ***K***.

We distinguish between "inner" and "outer" iterative such circles, but in order to see what these circles are, and how they matter in our proof, a digression is needed.

All functions from within hierarchy *H* can be mechanically named by the computor, since their names come from the previously mentioned mechanical method of attributing coordinates of the n-dimensional space. For the rest of programs, i.e., the ones beyond ***H′***, the same method won't do, since the computor has no means to mechanically represent coordinates of the $>\omega$ dimensional space. For naming these programs, we employ a hybrid notational method, which uses the coordinates of the n-dimensional space still, albeit together with the entire ω vocabulary. The method resembles, in several respects, the Archimedean technique of enriching the depository of available numerals beyond an initial cluster. Moreover, the method is such that the name assigned to the program suggests (or at times is) the ordinal that is needed for the Peano Arithmetic proof of the totalness of the function corresponding to the program.[9] E.g., the Peano Arithmetic proof for the totalness of ***H′*** requires Induction up to ω^ω. The method used by the computor is such that the name of the program for ***H′*** happens to be "ω^ω". Notice here that the use of the ω vocabulary is entirely for the needs of our proof and does not suggest any use of transfinite arithmetic tools on the part of the computor. Only the instructor knows the correspondence between names and ordinals. For the computor, all ω expressions are senseless names that have been given to the programs by a purely mechanical technique and upon a recursively definable vocabulary.

Now, an "inner circle" in our way towards ***K*** is defined by the distance separating the pair of programs that govern two consecutive general iterative circles, as described above. In other words, to each distinct general iterative algorithm, there corresponds an inner circle. As for the specific names of programs located at the boundaries of the inner circles, they are such that they notationally represent the distance between two limit ordinals. For example, the distance between programs ω^ω and $\omega^\omega+\omega$ constitutes an inner circle, and program $\omega^\omega+\omega$ is meant to be a pointer to all programs of the hierarchy ω^ω, $\omega^\omega+1$, $\omega^\omega+2$,…

An "outer circle," on the other hand, is represented by the distance separating a program $\omega^{(\omega\hat{}\ldots\hat{}\omega)n}$ and a program $\omega^{(\omega\hat{}\ldots\hat{}\omega)n+1}$, where the ω tower of the latter exceeds the ω tower of former by one ω. (Obviously, each outer circle is also an inner circle.)

Finally, program ***K*** is a pointer to all programs of the sequence of ω towers, ω^ω, $\omega^{\omega\hat{}\omega}$, $\omega^{\omega\hat{}\omega\hat{}\omega}$,… Exactly because ***K*** is this pointer, and also because of the specific program-naming technique we have employed, it is obvious that the Peano Arithmetic proof for ***K***'s totalness would have required Induction up to ε_0. For notice that the Peano Arithmetic proof of the totalness of the function calculated by program $\omega^{\omega\hat{}\omega}$ is $\omega^{\omega\hat{}\omega}$, the Peano Arithmetic proof of the totalness of the function calculated by program $\omega^{\omega\hat{}\omega\hat{}\omega}$ is $\omega^{\omega\hat{}\omega\hat{}\omega}$, and, so, the Peano Arithmetic proof of the totalness of the function calculated by ***K***, which is a pointer to all programs within the above sequence, must be ε_0.

Our main proof consists in showing by Induction $<\varepsilon_0$ that ***K*** is total.

The Induction for our main result is on the length of ω-s in the above towers. It proceeds as follows. First, we use the result of Stage 1 and take for granted that the program ω^ω (i.e., the program for ***H′***) returns an output for every input. Then, we prove that all programs from ω^ω to $\omega^{\omega\hat{}\omega}$, this one included, return an output for every input. Finally, we assume that some program $\omega^{(\omega\hat{}\ldots\hat{}\omega)n}$ with arbitrary n returns an output for every input, and show that, on this assumption, $\omega^{(\omega\hat{}\ldots\hat{}\omega)n+1}$ returns an output for every input.

This essentially completes the proof, because program ***K*** is a pointer to all programs of the sequence of ω towers. So, since ***K*** is a pointer to all such programs, the claim that it returns an output for all inputs is equivalent with the claim that (i) all programs at the boundaries of the outer circles return an output for all inputs and (ii) ***K*** has effective means to reduce any assignment to its input placeholders to an assignment to the input placeholders of (i). We show (ii) in a separate proof.

In some more detail, the core of the main proof consists in showing that there is a construction chain of programs that leads from $\omega^{(\omega\hat{}\ldots\hat{}\omega)n}$ to $\omega^{(\omega\hat{}\ldots\hat{}\omega)n+1}$, such that "effectiveness" is hereditary upon it. We show this by some subsidiary Induction that takes place within our main Inductive step, i.e., within the circle that begins by $\omega^{(\omega\hat{}\ldots\hat{}\omega)n}$ and ends with $\omega^{(\omega\hat{}\ldots\hat{}\omega)n+1}$. This subsidiary Induction is on the inner circles in between $\omega^{(\omega\hat{}\ldots\hat{}\omega)n}$ and $\omega^{(\omega\hat{}\ldots\hat{}\omega)n+1}$. Hereditariness of "effectiveness" concerns programs within these inner circles. For example, as said previously, the distance between programs ω^ω and $\omega^\omega+\omega$ constitutes an inner circle. Now, this inner circle happens to be within the outer circle ω^ω to $\omega^{\omega\hat{}\omega}$. In order to prove that effectiveness is hereditary in between ω^ω and $\omega^\omega+\omega$, we first assume that program ω^ω is effective and then show that the pointer $\omega^\omega+\omega$, which has the sequence ω^ω, $\omega^\omega+1$, $\omega^\omega+2$,…, as its pointing scope is effective too. Here again, the clause "relative to all assignments" is key to our proof. For the pointer $\omega^\omega+\omega$ is not effective in the standard use of the term, since it has an infinite scope. The pointer *becomes* effective *relative to all assignment to its input placeholders*, because the computor can acquire a mechanical method, which reduces any assignment to the input placeholders of $\omega^\omega+\omega$ to an assignment to the input placeholders of a program from within the sequence ω^ω, $\omega^\omega+1$, $\omega^\omega+2$,…, and can also learn how to mechanically construct this program. Again, this because the computor learns from the instructor how to mechanically construct any initial segment of the sequence of programs ω^ω, $\omega^\omega+1$, $\omega^\omega+2$,… By similar sub-proofs, the entire distance between ω^ω and $\omega^{\omega\hat{}\omega}$ is traversed. *Mutatis mutandis*, so can be traversed the distance between $\omega^{(\omega\hat{}\ldots\hat{}\omega)n}$ and $\omega^{(\omega\hat{}\ldots\hat{}\omega)n+1}$ of our main Inductive step.

***

If our general claim is correct, and nothing but mechanical means are employed in our proof, this proof consists in a refutation of the Church-Turing thesis, as standardly interpreted. This is because in standard formal arithmetic the proof of function ***K***'s totalness requires Induction up ε_0. So, our proof is a refutation of the Church-Turing thesis, because a consequence of the Church-Turing thesis is that Turing Machines and equivalent formalisms exhaust the limits of "mechanically computable." Therefore, since Peano Arithmetic is such a formalism and Peano Arithmetic (if consistent) cannot prove ***K*** total, there is a mechanical operation that cannot be performed by Turing Machines: the mechanical proof of the totalness of function ***K***.

NOTES

1. This is a summary of both our general philosophical argument and of the specific technical results.
2. Dershowitz and Gurevich, "A Natural Axiomatization of Computability and Proof of Church's Thesis"; Boker and Dershowitz, "The Church-Turing Thesis Over Arbitrary Domains"; Gurevich, "Sequential Abstract State Machines Capture Sequential Algorithms"; heavily relying on Gandy, "The Confluence of Ideas in 1936"; and Sieg, "Step By Recursive Step: Church's Analysis of Effective Calculability" and "Calculations by Man and Machine: Conceptual Analysis"; see also Sieg, "Church without a Dogma: Axioms for Computability."
3. Following Gandy, "The Confluence of Ideas in 1936", we call "computor" a human that computes, i.e., she is not allowed to employ nonmechanical means during her calculations.
4. Kapantaïs, "A Refutation of the Church-Turing Thesis According to Some Interpretation of What the Thesis Says."
5. Turing, "On Computable Numbers, with an Application to the *Entscheidungsproblem*," ch. 9; Post, "Finite Combinatory Processes – Formulation" Kleene, *Introduction to Metamathematics*.
6. For an indicative picture with respect to these, see Piccinini, "The Physical Church-Turing Thesis: Modest or Bold?"
7. In this proof-sketch, we will suppress the difference between functions and programs constructed by the computor for calculating the values of the same functions. We hope that which is which is always clear from the context.
8. See also Stage 2: main inductive step.
9. In this summary, we will encounter no hybrid names, though there are plenty of such in the hierarchies beyond *H*.

REFERENCES

Boker, U., and N. Dershowitz. "The Church-Turing Thesis Over Arbitrary Domains." In *Pillars of Computer Science: Essays Dedicated to Boris (Boaz) Trakhtenbrot*, edited by A. Avron, N. Derschowitz, and A. Rabinovitz, 199–229. Lecture Notes in Computer Science. Berlin: Springer, 2008.

Dershowitz, N., and Y. Gurevich. "A Natural Axiomatization of Computability and Proof of Church's Thesis." *The Bulletin of Symbolic Logic* 14 (2008): 299–350.

Gandy, R. "The Confluence of Ideas in 1936." In *The Universal Turing Machine: A Half-Century Survey*, edited by R. Herken, 55–111. Oxford: Oxford University Press, 1988.

Gurevich, Y. "Sequential Abstract State Machines Capture Sequential Algorithms." *ACM Transactions on Computational Logic* 1 (2000): 77–111.

Kapantaïs, D. "A Refutation of the Church-Turing Thesis According to Some Interpretation of What the Thesis Says." In *Computing and Philosophy*: Proceedings of *IACAP* 2014, edited by V. Müller, 45–62. Heidelberg: Springer, 2016.

Kleene, S. C. *Introduction to Metamathematics*. New York: D. Van Nostrand, 1952.

Piccinini, G. "The Physical Church-Turing Thesis: Modest or Bold?" *The British Journal for the Philosophy of Science* 62 (2011): 733–69.

Post, E. "Finite Combinatory Processes – Formulation." *Journal of Symbolic Logic* 47 (1936): 103–105.

Sieg, W. "Step By Recursive Step: Church's Analysis of Effective Calculability." *Bulletin of Symbolic Logic* 3 (1997): 154–80.

Sieg, W. "Calculations By Man and Machine: Conceptual Analysis." In *Reflections on the Foundations of Mathematics: Essays in Honor of Solomon Feferman*, edited by W. Sieg, R. Sommer, and C. Talcott, 396–415. Association for Symbolic Logic, 2002.

Sieg, W. "Church without a Dogma: Axioms for Computability." In *New Computational Paradigms*, edited by B. Lowe, A. Sorbi, and B. Cooper, 139–52. Berlin: Springer, 2008.

Turing, A. "On Computable Numbers, with an Application to the *Entscheidungsproblem*." *Proceedings of the London Mathematical Society* 2, no. 42 (1936): 230–65; and in Davis, M., ed., *The Undecidable: Basic Papers on Undecidable Propositions, Unsolvable Problems and Computable Functions*. Hewlett, NY: Raven Press, 1965.

RAPAPORT Q&A

Logicist Remarks on Rapaport on Philosophy of Computer Science+

(in the context of his Barwise Prize)

Selmer Bringsjord
RENSSELAER POLYTECHNIC INSTITUTE

INTRODUCTORY REMARKS

I shall restrict my brief remarks herein to William "Bill" Rapaport on philosophy of computer science (PCS) and some intimately related topics (which are gestured at by the superscripted + in my title), guided by his ever-expanding, online *Philosophy of Computer Science* (*PCS*); and I'll begin (in the next section) with some comments on this restriction itself. The present commentary is informed by a recent, sustained dialogue with Rapaport, one undertaken to inform my remarks (and, I confess, to allow me to somewhat selfishly enjoy some philosophical debate).[1] Unfortunately, and I wrap up the present essay by returning to this issue, our dialogue, at least by my lights, needs to continue, because important societal issues in the context of the philosophico-history of computer science and AI have been left unanalyzed, and more importantly (at least as I see things), because Rapaport (and his readers) would be well-served by having some errors that infect his *PCS*, beyond those touched upon herein, remedied. In particular, since—for reasons to be shortly seen—he views CS through the obfuscating lens of *algorithms* (first do *A*; now do *B*; if condition *C* holds, do *A* again; and so on), rather than as a part of reasoning in a well-defined logical system, it's especially important that Rapaport's account of PCS, which seems destined to be highly influential, be modified. I suppose it's possible that despite sustained discussion with him subsequent to what informs the present essay, he may resist such modification; but I hold out hope that he will engage in the discussion and see the light.

THE VASTNESS OF RAPAPORT'S REACH VS. WHAT I TREAT

As the reader will well know, much if not all of the field of philosophy is composed of sub-parts long traditionally designated by the phrase "philosophy of *X*," where instantiations of *X* include, for instance, "mind," "art," "economics," "religion," and "language."[2] In cases where a sub-part of philosophy is designated without this syntax, as, for example, "epistemology" or "metaphysics," there can be little doubt that no accuracy is sacrificed if the Po*X* template is employed (though elegance, I concede, is threatened). Rapaport has made contributions in many a philosophy of $X \neq$ 'computer science' area, but my interest, in keeping with his recent Barwise Prize, and with the venue that the present discussion is bound for, is PCS, to which, arguably, Rapaport is *the* greatest contributor—and at any rate he certainly stands minimally as one of the five or so greatest authorities on PCS today, when the whole of CS, from theory to concrete practice, is considered. The restriction to PCS means, in particular, that very little will be said herein about philosophy of artificial intelligence (PAI),[3] another Po*X* subject on which Rapaport is a world-class authority, in no small part because of his being a longtime leader in a seminal team at the University at Buffalo devoted to AI and computational cognitive science. Multiple essays of the present sort could be written on the work of this group in connection with PCS and PAI, a group long led as well by AI pillar Stuart "Stu" Shapiro. Shapiro and Rapaport have long labored to advance the SNePS system, which can be used to build artificial agents that know, reason, plan, and act. Obviously, upon hearing my implicit claim (expressed by the previous sentence) that such artificial agents are on a planet that's lucky, AI-wise, to build self-driving cars that only *occasionally* kill people, readers who are philosophers will pay attention. Are such agents in fact with us *already*? it will be asked by such readers. I think an affirmative reply would come from Rapaport and Shapiro, and I suggest that philosophers of CS, AI, mind, and logic study the work in question, deeply.

So the target is Rapaport on PCS. In our context, this target should strike the alert reader as pregnant. I'm writing for the Philosophy and *Computers* Committee (P&CC) of the APA; note my emphasis. It follows that I'm writing for a committee whose mission centers on the relationship between philosophy on the one hand, and "computers" on the other. But what is the meaning of "computers" in this mission? This is very much like the question with which Rapaport has wrestled, when, for instance, he deliberated about what title to use for his *PCS* book. It turns out that he isn't particularly happy with the phrase "philosophy of computer science." He finds the continuous string "computerscience" to be helpful, because (to brutally simplify the issue and his thinking) this neologism is easier to view as something that picks out a domain over which to philosophize that isn't in any way narrowly restricted to computers, or to what must be a science, and so on. One would think that a similar attitude is wise to adopt regarding the title and nature of the P&CC: Surely this committee's mission isn't any such narrow thing as exploring, sorting out, and charting for the APA the relationship between philosophy and, literally, computers, as in laptops and desktops. Surely "computers" here is to mean that vast space of all philosophical things computational and computation-based, from all that Rapaport deals with in the bordering-on-1,000-page *PCS* volume, to rigorously characterizing what privacy is by the standards of philosophy (which includes characterizations in its analytic side that at least aspire to jointly necessary and sufficient conditions) in an age of social media, where interaction on the shoulders of computation has led to philosophical problems as thorny as most longstanding ones, an issue to which I return when wrapping up.

Please note that in confining attention to Rapaport on PCS, the target remains enormous. This is true for the simple reason that PCS itself is gigantic. It's perhaps not uninteresting that in philosophy today, still, PCS is often thought of as some kind of Lilliputian curiosity off to the side, with the center proudly occupied by the venerable giants (ethics, epistemology, metaphysics, etc.) continuing to go merrily along as they have since Socrates. Those with this attitude should read *PCS*, and then think objectively about whether this traditional center-side conceptualization is accurate and/or sensible today. We have reached a time, now, when the prospect of artificial agents (which after all consist in things whose essence is *computing* over input to produce output) that are ethical agents unto themselves, with radical forms of autonomy (e.g., the ability to write the very programs that power them), seem to many imminent. Understanding these creatures, and what they mean for us and the cosmos, will be impossible without a prior understanding of PCS.

ACTUALLY, COMPUER SCIENCE IS A (SMALL) PROPER PART OF LOGIC

In *PCS* Rapaport bravely gives a distilled answer to "What is computer science?" The answer is given at the very end of the chapter whose title is the very question, and, verbatim, Rapaport's summative reply is this:

> Computer science is the scientific (or STEM) study of:
>
> what problems can be solved,
>
> what tasks can be accomplished,
>
> and what features of the world can be understood . . .
>
> . . . computationally, that is, using a language with only:
>
> 2 nouns ('0', '1'),
>
> 3 verbs ('move', 'print', 'halt'),
>
> 3 grammar rules (sequence, selection, repetition), and nothing else,
>
> and then to provide algorithms to show how this can be done:
>
> efficiently, practically, physically, and ethically.

This answer has a certain flair, I think. After all, by it, a great, big, daunting philosophical question is answered crisply and confidently in nothing more than a flash. Unfortunately, this is an account of computer science ferociously biased in the procedural direction. (The account is very nicely elaborated in *PCS*, and is explicitly aligned with (similarly biased) accounts of so-called "computational thinking," the cultivation of which, at least in the US, is sought by its federal government, by many states as well, and by funders like the Gates Foundation.) Yet this is not *my* answer to the question, nor is it even *approximately* in line with my answer; and I doubt whether it's the answer that would be given by anyone who thinks of computation as a proper part of reasoning and nothing more, not as a do-this-step-do that-step-do-this-step (DTS) process. Moreover, for philosophy and philosophers, I think a DTS account of CS is particularly unwise. The reason is simply that philosophers, if they do nothing else, reason; and to teach philosophy is therefore naturally to in no small part teach how to reason. (Such pedagogy is of course self-evidently in operation in the case of logic as taught and pursued under philosophy.) In my experience, sometimes philosophers with little exposure to CS are surprised to learn that computation can be studied and mastered, without loss of formal generality or of practical functioning, as reasoning, but some illumination can be provided quickly by presenting the rudiments of standard logic programming. I personally have found that the instant a rigorously trained philosopher without any prior exposure to computer science/computation is shown the underlying theory of logic programming for Prolog (a programming language in the logic-programming fold), a light snaps on.[4] In fact, sometimes the coming on of that mental light is more akin to a sort of explosive eureka moment. "Wait, you mean a valid deduction by the machine from this set *A* of formulae expressed in something that looks quite like first-order logic, to that particular formula *p*, is what execution of my 'program' consists in?!?" That is correct. No need to write any DTS thingie here, at all. The traditional coverage of logic programming in mathematical logic isn't based on inference schemata that philosophers learn (e.g., *modus tollens*, universal elimination, etc.), but rather on inference schemata in the proof theories based on schemata conforming to **resolution**, but regardless, this is a far superior way to understand what computation is, in my opinion—yet this way is utterly alien in the DTS landscape of *PCS*.[5]

SEMANTICS AS SEMANTICS, AND SEARLE

I have been intrigued for years by Rapaport's longstanding desire to portray semantics as syntax, and accordingly took up for the present project his 2016 "Semantics as Syntax" (which was wisely solicited by editor Boltuc) to study. Rapaport, as far as it goes, is entirely correct, at least spiritually speaking. (I'm limited to saying only that Rapaport is in *spirit* right, because were details discussed here, too much space would be consumed.) For my money, one major reason he's right is that the fundamental observations upon which proof-theoretic semantics (in any form thereof) is motivated by, and possibly even rests upon directly, can't be denied.[6] A simple example comes by way of considering the standard extensional semantics of a conditional with *p* as antecedent and *q* as consequent. We are standardly told in this case that the semantics for a material conditional $p \Rightarrow q$ consists in that such a conditional holds if and only if (iff), if *p*, then *q*. That is, expressed a bit more succinctly, $p \Rightarrow q$ iff if *p* then *q*. When you think about it, this is quite extraordinarily one-dimensional. Does it not directly give semantics via syntax? Consider the conditional $(p \,\&\, q) \Rightarrow q$. Does this conditional have the semantic value TRUE? Certainly. Why? Because it's TRUE iff if *p* and *q*, then *q*. Well, is it in turn TRUE that if *p* and *q*, then *q*? Absolutely:

Proof: Suppose that *p* holds, along with *q*. We can deduce *q* directly. Hence our supposition implies *q*. **QED**

We are here using the standard textbook semantics for elementary extensional deductive logic, in use in classrooms across the globe, and what just happened? What happened is that we pinned down the meaning of the syntactic formula via a perfectly, indeed purely, syntactic process.[7] I view Rapaport as having found this phenomenon at work in a deep and intricate way, far and wide.

Yet why do I say that Rapaport's "sem-by-syn" view is correct only *as far as it goes*? The reason is that Rapaport is spot on with respect to *one* sense of "semantics," and dead wrong with regard to *another* sense of the term. The first sense aligns with proof-theoretic semantics, in general; we have just seen this sense in operation on a simple specimen; *and* it aligns with any formal dyad covering syntax on the one hand and semantics on the other. Unfortunately, the second sense can't be separated from *understanding on the part of a mind*; this is the Searlean sense of semantics, and is what stands at the heart of Searle's justly famous Chinese Room Argument (CRA), whose kernel, as a slogan, is that syntax doesn't produce semantics. Rapaport believes that the sem-by-syn view can be extended in order to allow syntactic expressions (e.g., "hamburger") to be "internalized," and hence CRA to be dodged. He writes:

> In the case of a real human being, [a] *representative* is the end result of, say, the visual process of seeing a hamburger . . . resulting in a "mental image" of a hamburger. . . . More precisely, the biological neural network in the human's brain has neurons whose firing represent the word 'hamburger', and it has neurons whose firings represent the actual hamburger. Both of these sets of neuron firings are in the same "language"—the same syntactic system.[8]

This quote does nothing beyond communicating the faith of computationalist materialists, and/or (with the "neuron" here, e.g., mapped to artificial neurons in artificial neural networks so in vogue again these days) Strong AIniks. Can't we imagine this more elaborate syntactic dance happening in the complete and utter absence of our understanding, bound up with subjective awareness as it is, of the shout by a grillmaster that our redolent burger is done? Of course we can. What Rapaport is in the end doing is ingeniously (but to a degree unwittingly) working out the sem-by-syn paradigm in and for AI—but not for *us*.

HYPERCOMPUTATION

Rapaport's *PCS* includes a chapter on hypercomputation (which is, harshly encapsulated, forms of information-processing more powerful than the operation of standard Turing machines); coverage of the topic therein is what most would no doubt classify as "steadfastly balanced." I somewhat less charitably classify this chapter as noncommittal, and in being so, well, irrational. However, the chapter is also, even in its present, not-fully-polished form, the absolute best overview of the topic available in one place, over one digestible-in-one-sitting stretch of content. Indeed, I suspect that even most aggressive fans of hypercomputation will regard the chapter's wishy-washy maybe-maybe-not position on hypercomputation to be fully redeemed by its laconic erudition, right down to the lucid presentation of some key theorems. After all, *PCS* is intended to be a broad-coverage textbook, not a polemical position statement.

Nonetheless, I've declared the chapter to be irrational. Why? In short, because there can be no denying, in light of the relevant logico-mathematics, that hypercomputation is as real and robust as can be, in the context of the fact that even if (like me) we count Leibniz as having discovered general-purpose computation in the seventeenth century, the human race has really only been at this modern computation thing for about three centuries. The late twentieth century, and the beginning of the third millennium, have revealed that computation absolutely, positively cannot be rationally restricted to what standard Turing machines and their equivalents (which Rapaport lists and often discusses in *PCS*) can compute. I can't here review in any detail my own writings on this subject, and will rest content to mention but two things. To wit:

One: Rapaport respectfully cites and discusses Martin Davis's "The Myth of Computation." While there can be no denying that Davis is the author of much brilliant work, this paper is far from his finest hour; it may, in fact, be his worst. Calling a spade a spade (and I did have the opportunity to do so orally, in debating the issue with Davis in person), joined by my colleague N. S. Govindarajulu, we wrote something I recommend to Rapaport, his readers, and readers of the present essay: "The Myth of 'The Myth of Hypercomputation'," [9] in which is shown that Davis's arguments are anemic at best and stunningly fallacious at worst. I confess to being deeply surprised that Rapaport is content, at least at present, to leave the impression that Davis may have succeeded in revealing that hypercomputation is to be placed alongside, say, Hercules and Odin.

Two: It's a logico-mathematical fact that hypercomputation is as real as can be. In the logicist interpretation of computer science adumbrated above, we have only to consider, for a few minutes, any number of computing machines vastly more powerful than standard Turing machines and their equivalents, specified via the use of formal logic. Not wanting (again) to cite my own work in this connection, I can simply rely on infinite-time Turing machines;[10] they provably exceed standard Turing machines, and yet *are* Turing machines; end of story. An even-more-direct route is simply to take note of the fact that formal logic includes infinitary logics, and some reasoning (e.g., proof discovery) in even the smallest of these (which allow infinitely long formulae and infinitely long proofs) is logic-style hypercomputation. Of course, some myopic empiricists may deny the reality of hypercomputation because they affirm the dogma that what is *real* is only what is *physical*. But this position is not only at odds with such mathematical facts as that there is a natural number N too large to correspond to any physical entity whose components sum to N; it's also at odds with something that Rapaport leaves aside: Since we are coming to see that physics can be axiomatized (by, say, the axiom system **P**), absent a disproof of the proposition that **P** and a formal assertion of the *physical* existence of hypercomptuational machine is consistent, it's irrational to advance the claim that hypercomputation is only mathematically possible.[11]

FINAL REMARKS

Any serious dialogue with Rapaport, and engagement with his writings, could clearly continue, profitably and enjoyably, for a very long time. Yet, as is always the case, in order for a piece to be delivered and published, we must end—with, if you'll allow, a final thought: viz., that we need to hear at some point soon from Rapaport-qua-philosopher on the *history* of computation, of the fields which centrally partake of it (e.g AI, logic, mathematics, linguistics, and nowadays computing machines as ethical agents), and on the complex and philosophically charged turbulence that has now been catalyzed by so-called "social media." Rapaport's professional life shows no signs of slowing down (witness the ever-growing *PCS* book itself), which means his contributions will continue, but his professional life to this point has passed through the evolution of the computational sciences over a period of decades, during which time *a lot* has happened. Rapaport is one of only a handful of computationally informed philosophers who have seen firsthand the evolution (with an occasional spate of rapid change) of the many parts of philosophy intimately connected to computation (philosophy of mind, of language, etc.). Did he ever think for a moment, yesterday, that today's advocacy of the end of programming (in light of such phenomena as "Deep Learning") would ever arrive? That the concept of a machine which self-learns and thereby beats humans at their own games would become reality, as happened in the case of AlphaGo? Did he think, yesterday, that computation, first isolated in the minds and soon thereafter the simple, disconnected "pet" machines of Turing and von Neurmann et al., would come to mediate arguably all that Earth's technologized youth do, daily, via social-media technology? In all this, who are we? What is truth? What is fake? What is real? What control can computation be allowed to have over our interaction with each other, and over the analysis and presentation thereof? Philosophy, and anyone concerned with the intersection "philosophy and computers," is going to need to come to grips with these computation-infused questions, the lack of answers to which has already started to plague us.[12] Actually, truth be told, *I* need to come grips in this regard. Time to talk again to Rapaport . . .

ACKNOWLEDGMENTS

Some research in AI and theoretical computer science that informs my commentary was made possible by support from ONR and AFOSR, and I'm very grateful for this support. I have an enormous debt to Piotr Boltuc for his guidance and supernatural patience as I (irrationally?) poured more and more time into thinking about Rapaportian work, all the while with the clock cranking beyond a series of promised delivery dates. Rapaport's body of work, as I've said, has only been quickly touched upon herein. That body of work is endlessly stimulating, and I'm grateful to Rapaport for creating it.

NOTES

1. The current version of *PCS*, as this sentence is written, is May 2018, and is available at https://cse.buffalo.edu/~rapaport/Papers/phics.pdf. The reader should take account of the difference between PCS (the subject) and—note the italics—*PCS*, the Rapaportian book on that very subject.
2. I don't mean to imply that the sub-parts of philosophy to which I refer are self-contained. In point of fact, philosophy of language and philosophy of logic (in the Occidental case, anyway), are inseparably linked. Another inseparable link, one at the heart of any comprehensive analysis of Rapaport's PCS and his body of work, is that between PCS and PAI.
3. PAI, and for that matter AI itself from a philosophical point of view, is covered in the SEP entry *Artificial Intelligence* (https://plato.stanford.edu/entries/artificial-intelligence).
4. Wonderful introductory coverage of logic programming is provided in Ebbinghaus et al., *Mathematical Logic*.
5. I would personally have preferred to use automated theorem proving rather than Prolog's basis in what I just wrote, but the need for economy at the moment rules. This is as good a place as any to report that in my interview of Rapaport, he indicated that he opted for DTS, and the encapsulation of it that I've quoted, for pedagogical purposes. However, even taking his expression of this strategy at face value, as I've explained, even from the perspective of pedagogy, reasoning is by my lights something much more valuable to teach than DTS. And besides, even after DTS is used, we are still left with the challenge of showing that the procedural artifact we have produced is correct; and showing this can only be accomplished via reasoning. Why not simply start and end with reasoning?
6. Readers unfamiliar with proof-theoretic semantics could start with Gentzen, "Investigations into Logical Deduction." For what it's worth, nearly all my own work in intensional logic and philosophy is proof-theoretic in nature. See, e.g., Bringsjord et al., "Introducing the Doxastically Centered Approach to Formalizing Relevance Bonds in Conditionals."
7. Die-hard Tarskians might accuse me of tendentiously and unfairly passing straightaway to a proof, rather than giving a truth-table or truth-tree (or in the first-order case a model/interpretation). Balderdash. We shall need for the skeptic a proof that the result of tabular or tree-based manipulation yields TRUE.
8. Rapaport, "Semantics as Sytax," 12.
9. Govindarajulu and Bringsjord, "The Myth of 'The Myth of Hypercomputation'."
10. Hamkins and Lewis, "Infinite Time Turing Machines,"
11. Govindarajulu et al., "Proof Verification and Proof Discovery for Relativity," isn't a bad place to start reading about such matters.
12. I recommend, as a quick, non-technical start to this side of PCS, Wiesberg, "The Digital Poorhouse."

REFERENCES

Bringsjord, S., J. Licato, D. Arista, N. S. Govindarajulu, and P. Bello. "Introducing the Doxastically Centered Approach to Formalizing Relevance Bonds in Conditionals." In *Computing and Philosophy*, edited by V. Müller, 117–31. Berlin, Germany: Springer, 2016. This volume is *Synthese Library 375*, in *Studies in Epistemology, Logic, Methodology, and Philosophy of Science*.

Davis, M. "The Myth of Computation." In *Alan Turing: Life and Legacy of a Great Thinker*, edited by C. Teuscher, 195–211. Berlin, Germany: Springer, 2004.

Ebbinghaus, H., J. Flum, and W. Thomas. *Mathematical Logic*, 2nd edition. New York, NY: Springer-Verlag, 1994.

Gentzen, G. "Investigations into Logical Deduction." In *The Collected Papers of Gerhard Gentzen*, edited by M. E. Szabo, 68–131. Amsterdam, The Netherlands: North-Holland, 1935. (This is an English version of the well-known 1935 German version.)

Govindarajulu, N. S., and S. Bringsjord. "The Myth of 'The Myth of Hypercomputation'." *Parallel Processing Letters* 22, no. 3 (2012), https://doi.org/10.1142/S0129626412400129.

Govindarajalulu, N. S., S. Bringsjord, and J. Taylor. "Proof Verification and Proof Discovery for Relativity." *Synthese* 192, no. 7 (2015): 2077–94.

Hamkins, J., and A. Lewis. "Infinite Time Turing Machines." *Journal of Symbolic Logic* 65, no. 2 (2000): 567–604.

Weisberg, J. "The Digital Poorhouse." *The New York Review of Books*, June 7, 2018, online.

RAPAPORT RESOURCES

Rapaport's PCS is currently available, in the May 2018 edition, here: https://cse.buffalo.edu/~rapaport/Papers/phics.pdf.

Rapaport, W. "Semantics as Syntax." 2016. https://cse.buffalo.edu/~rapaport/Papers/synsemapa.pdf.

Comments on Bringsjord's "Logicist Remarks"

William J. Rapaport
UNIVERSITY AT BUFFALO, THE STATE UNIVERSITY OF NEW YORK

1. INTRODUCTION

I am grateful to my long-time friend and debating partner Selmer Bringsjord for the email interview that was the source of his "Logicist Remarks"[1] and for his generous comments on my research and my textbook-in-draft.[2] In this brief response, I simply wish to clarify three of my positions.

2. A PROCEDURAL-LOGICAL CONTROVERSY

Whereas I argue that computer science is fundamentally concerned with algorithms,[3] Bringsjord argues that "computation . . . [is] a proper part of reasoning *and nothing more*."[4] It's the "and nothing more" clause that I disagree with.

I agree that computation as a subject of study can be viewed "as a proper part of reasoning" or logic. But it can *just as well* be viewed as the study of (what Bringsjord somewhat dismissively characterizes as "do-this-step-do-that-step-do-this-step") procedures (loc. cit.). These are equivalent viewpoints from different perspectives. But I find the procedural perspective more perspicuous.[5]

This is exactly the same situation that we find in the theory of computation: Computation as a mathematical enterprise can be understood functionally, in terms of recursive functions or the lambda calculus (as well as in other ways, and by other formalisms), as well as procedurally, in terms of Turing machines or register machines (etc.). The multiple views (in both cases) are not rivals, but equivalent alternatives, each with its own advantages. Gödel found the Turing-machine analysis more convincing as a model of computability than even his own recursive functions.[6] Similarly, I would argue, the procedural view is more compelling (for me, as well as for my students) than the logical view with respect to what is unique and interesting about computer science and computation.

3. SYNTACTIC SEMANTICS

I have long advocated for the position that syntax suffices for semantics—that the semantic enterprise of understanding is fundamentally a syntactic one. Briefly, I take syntax as the study of the properties of, and relations among, the members of a set of objects, and I take semantics as the study of the relations *between* two sets of objects—one studied syntactically, and other providing its semantic interpretation. (The latter set can also be studies syntactically, and its syntax is its "ontology".) But when you take the union of those two sets, the formerly semantic relations become syntactic ones of the union.[7] A real-life, biological (and not merely "Strong AI") example of such a union is the neuron firings in our brain, some of which represent the objects in the external world and some of which represent the concepts (and language) that we use to understand them. But they all form one neural network. Bringsjord says that "this . . . does nothing beyond communicating the faith of computationalist materialists, and/or . . . Strong AIniks."[8] But it *does* do more than that: It shows that our sub-jective sense of understanding—the kind involved in Bringsjord's example of "the shout by a grillmaster that our redolent burger is done" (loc. cit.)—is accomplished by a single system (a single, unioned set) that is understood syntactically, not by two separate systems (a syntactically understood one and its semantic interpretation).

4. HYPERCOMPUTATION

Bringsjord's discussion of my position on hypercomputation was based on an earlier version of *Philosophy of Computer Science*, Chapter 11, than the one currently available. That earlier version was, indeed, somewhat "noncommittal."[9]

Rather than distinguishing between Turing-machine computation and hyper-computation, I prefer to think of there being three categories:[10]

Sub-Turing Computation:

> Finite-state automata, pushdown automata, primitive recursive functions, etc.

Turing-Machine Computation:

> Turing machines and their equivalents (partial recursive functions, lambda calculus, etc.)

Super-Turing Computation:

> Oracle machines, Zeus machines, Malament-Hogarth machines, analog recurrent neural networks, interactive computing, trial-and-error machines, etc.

To my mind, the only interesting kinds of super-Turing computation are not the "newer physics" kind (Zeus machines, etc.),[11] but the ones that can be modeled by Turing's own theory of oracle machines. These include interactive and trial-and-error computing. But oracle computation, studied under the rubric 'relative computability', is well-understood and not something that computer scientists have ignored (as some hypercomputationalists have suggested). Nor is it typically un-derstood as a counterexample to the Church-Turing Computability Thesis.[12]

5. CONCLUSION

Bringsjord raised a number of important questions in his "Final Remarks" (some of which I touch on in my book), observing that it was "time to talk again to Rapaport." I look forward to continuing our conversation!

NOTES

1. Bringsjord, "Logicist Remarks on Rapaport on Philosophy of Computer Science."
2. Rapaport, "Philosophy of Computer Science."
3. Rapaport, "What Is Computer Science?" 13–16; Rapaport, "Philosophy of Computer Science," §3.15
4. Bringsjord, "Logicist Remarks on Rapaport on Philosophy of Computer Science," §"Actually, Computer Science Is a (Small) Proper Part of Logic," my italics.

5. As I note in Rapaport, "Philosophy of Computer Science," §§2.3, 2.7, on the question of what philosophy is, I take philosophy to be the personal search for truth, in any field, by rational means, following Hector-Neri Castañeda, who said that philosophy should be done "in the first person, for the first person" (Rapaport, "Castañeda, Hector-Neri").
6. Gödel, "Undecidable Diophantine Propositions," 168; Shagrir, "Gödel on Turing on Computability"; Sieg, "Gödel on Computability"; Soare, "Turing Oracle Machines, Online Computing, and Three Displacements in Computability Theory," §2; Copeland and Shagrir, "Turing versus Gödel on Computability and the Mind."
7. Rapaport, "Searle's Experiments with Thought"; "Syntactic Semantics: Foundations of Computational Natural-Language Understanding"; "Understanding Understanding: Syntactic Semantics and Computational Cognition"; "How to Pass a Turing Test: Syntactic Semantics, Natural-Language Understanding, and First-Person Cognition"; "Holism, Conceptual-Role Semantics, and Syntactic Semantics"; "What Did You Mean By That? Misunderstanding, Negotiation, and Syntactic Semantics"; "How Helen Keller Used Syntactic Semantics to Escape from a Chinese Room"; "Yes, She Was! Reply to Ford's 'Helen Keller Was Never in a Chinese Room'"; "Semiotic Systems, Computers, and the Mind: How Cognition Could Be Computing"; "Semantics as Syntax"; and "Syntactic Semantics and the Proper Treatment of Computationalism."
8. Bringsjord, "Logicist Remarks on Rapaport on Philosophy of Computer Science," §"Semantics as Semantics, and Searle." I suspect that this section title is a typo for "Semantics as Syntax, and Searle."
9. Bringsjord, "Logicist Remarks on Rapaport on Philosophy of Computer Science," §"Hypercomputation."
10. I am limiting myself here to digital computing, so analog computation is another story, told best, I think, in Piccinini, *Physical Computation: A Mechanistic Account*.
11. The term "newer physics" is from Copeland and Sylvan, "Computability Is Logic-Relative," 190.
12. See Davis, *Computability and Unsolvability*, 20–24; Soare, "Turing Oracle Machines, Online Computing, and Three Displacements in Computability Theory"; Soare, "Formalism and Intuition in Computability"; and Fortnow, "What Is Computation?" for this point of view.

REFERENCES

Bringsjord, S. "Logicist Remarks on Rapaport on Philosophy of Computer Science." APA Newsletter on Philosophy and Computers 18, no. 1 (2018): 28–31.

Copeland, B. J., and O. Shagrir. "Turing versus Gödel on Computability and the Mind." In *Computability: Turing, Gödel, Church, and Beyond*, edited by B. J. Copeland, C. J. Posy, and O. Shagrir, 1–33. Cambridge, MA: MIT Press, 2013.

Copeland, B. J., and R. Sylvan. "Computability Is Logic-Relative." In *Sociative Logics and Their Applications: Essays by the Late Richard Sylvan*, edited by D. Hyde and G. Priest, 189–99. Ashgate, 2000.

Davis, M. D. *Computability and Unsolvability*. New York: McGraw-Hill, 1958.

Fortnow, L. "What Is Computation?" *Ubiquity* (December 2010). Article 5, http://ubiquity.acm. org/article.cfm?id=1921573.

Gödel, K. "Undecidable Diophantine Propositions." In *Collected Works, Vol. III*, edited by S. Feferman et al., 164–75. Oxford: Oxford University Press, circa 1938. http://tinyurl.com/ybw84oa6.

Moor, J. H., ed. *The Turing Test: The Elusive Standard of Artificial Intelligence*. Dordrecht, The Netherlands: Kluwer Academic, 2003.

Piccinini, G. *Physical Computation: A Mechanistic Account*. Oxford: Oxford University Press, 2015.

Rapaport, W. J. "Searle's Experiments with Thought." *Philosophy of Science* 53 (1986): 271–79. http://www.cse.buffalo.edu/~rapaport/Papers/philsci.pdf.

Rapaport, W. J. "Syntactic Semantics: Foundations of Computational Natural-Language Understanding." In *Aspects of Artificial Intelligence*, edited by J. H. Fetzer, 81–131. Dordrecht, The Netherlands: Kluwer Academic Publishers, 1988. http://www.cse.buffalo.edu/~rapaport/Papers/synsem.pdf; reprinted with numerous errors in Eric Dietrich, ed., *Thinking Machines and Virtual Persons: Essays on the Intentionality of Machines*, 225–73. San Diego: Academic Press, 1994.

Rapaport, W. J. "Understanding Understanding: Syntactic Semantics and Computational Cognition." In *AI, Connectionism, and Philosophical Psychology (Philosophical Perspectives, Vol. 9)*, edited by J. E. Tomberlin, 49–88. Atascadero, CA: Ridgeview, 1995. http://www.cse.buffalo.edu/~rapaport/Papers/rapaport95-uu.pdf. Reprinted in *Language and Meaning in Cognitive Science: Cognitive Issues and Semantic Theory (Artificial Intelligence and Cognitive Science: Conceptual Issues)*, Vol. 4, edited by Josefa Toribio and Andy Clark. New York: Garland, 1998.

Rapaport, W. J. "How to Pass a Turing Test: Syntactic Semantics, Natural-Language Understanding, and First-Person Cognition." *Journal of Logic, Language, and Information* 9, no. 4 (2000): 467–90. http://www.cse.buffalo.edu/~rapaport/Papers/TURING.pdf. Reprinted in Moor, *The Turing Test: The Elusive Standard of Artificial Intelligence*, 161–84.

Rapaport, W. J. "Holism, Conceptual-Role Semantics, and Syntactic Semantics." Minds and Machines 12, no. 1 (2002): 3–59. http://www.cse.buffalo.edu/~rapaport/Papers/crs.pdf.

Rapaport, W. J. "What Did You Mean By That? Misunderstanding, Negotiation, and Syntactic Semantics." *Minds and Machines* 13, no. 3 (2003): 397–427. http://www.cse.buffalo.edu/~rapaport/Papers/negotiation-mandm.pdf.

Rapaport, W. J. "Castañeda, Hector-Neri." In *The Dictionary of Modern American Philosophers, 1860–1960*, edited by J. R. Shook, 452–12. Bristol, UK: Thoemmes Press, 2005. http://www.cse.buffalo.edu/~rapaport/Papers/hncdict.tr.pdf.

Rapaport, W. J. "How Helen Keller Used Syntactic Semantics to Escape from a Chinese Room." *Minds and Machines* 16 (2006): 381–436. http://www.cse.buffalo.edu/~rapaport/Papers/helenkeller.pdf. See reply to comments, in Rapaport, "Yes, She Was! Reply to Ford's 'Helen Keller Was Never in a Chinese Room'."

Rapaport, W. J. "Yes, She Was! Reply to Ford's 'Helen Keller Was Never in a Chinese Room'." *Minds and Machines* 21, no. 1 (2011): 3–17. http://www.cse.buffalo.edu/~rapaport/Papers/Papers.by.Others/rapaport11-YesSheWas-MM.pdf.

Rapaport, W. J. "Semiotic Systems, Computers, and the Mind: How Cognition Could Be Computing." *International Journal of Signs and Semiotic Systems* 2, no. 1 (2012): 32–71. http://www.cse.buffalo.edu/~rapaport/Papers/Semiotic_Systems,_Computers,_and_the_Mind.pdf. Revised version published as Rapaport, "Syntactic Semantics and the Proper Treatment of Computationalism."

Rapaport, W. J. "Semantics as Syntax." APA Newsletter on Philosophy and Computers 17, no. 1 (2017a): 2–11. http://c.ymcdn.com/sites/www.apaonline.org/resource/collection/EADE8D52-8D02-4136-9A2A-729368501E43/ComputersV17n1.pdf.

Rapaport, W. J. "What Is Computer Science?" *APA Newsletter on Philosophy and Computers* 16, no. 2 (2017b): 2–22. Paper based on 2017 APA Barwise Award talk. http://c.ymcdn.com/sites/www.apaonline.org/resource/collection/EADE8D52-8D02-4136-9A2A-729368501E43/ComputersV16n2.pdf.

Rapaport, W. J. "Philosophy of Computer Science." 2018a. Current draft in progress at http://www.cse.buffalo.edu/~rapaport/Papers/phics.pdf.

Rapaport, W. J. "Syntactic Semantics and the Proper Treatment of Computationalism." In Empirical Research on Semiotics and Visual Rhetoric, edited by M. Danesi, 128–76. Hershey, PA: IGI Global, 2018b. References on pp. 273–307; http://www.cse.buffalo.edu/~rapaport/Papers/SynSemProperTrtmtCompnlism.pdf. Revised version of Rapaport, "Semiotic Systems, Computers, and the Mind: How Cognition Could Be Computing."

Shagrir, O. "Gödel on Turing on Computability." In *Church's Thesis after 70 Years*, edited by A. Olszewski, J. Wołenski, and R. Janusz, 393–419. Ontos-Verlag, 2006. Page references to online version at http://edelstein.huji.ac.il/staff/shagrir/papers/Goedel on Turing on Computability.pdf.

Sieg, W. "Gödel on Computability." *Philosophia Mathematica* 14 (2016): 189–207. Page references to preprint at http://repository.cmu.edu/cgi/viewcontent.cgi?article=1112&context= philosophy.

Soare, R. I. "Turing Oracle Machines, Online Computing, and Three Displacements in Computability Theory." *Annals of Pure and Applied Logic* 160 (2009): 368–99. Preprint at http://www.people.cs.uchicago.edu/~soare/History/turing.pdf; published version at http://ac.els-cdn.com/S0168007209000128/1-s2.0-S0168007209000128-main.pdf?tid=8258a7e2-

01ef-11e4-9636-00000aab0f6b&acdnat=1404309072 f745d1632bb6fdd95f711397fda63ee2. A slightly different version appears as Soare, "Interactive Computing and Relativized Computability."

Soare, R. I. "Formalism and Intuition in Computability." *Philosophical Transactions of the Royal Society* A 370 (2012): 3277–304. doi:10.1098/rsta.2011.0335.

Soare, R. I. "Interactive Computing and Relativized Computability." In *Computability: Turing, Gödel, Church, and Beyond*, edited by B. J. Copeland, C. J. Posy, and O. Shagrir, 203–60. Cambridge, MA: MIT Press, 2013. A slightly different version appeared as Soare, "Turing Oracle Machines, Online Computing, and Three Displacements in Computability Theory."

Exploring the Territory: The Logicist Way and Other Paths into the Philosophy of Computer Science

Robin K. Hill
UNIVERSITY OF WYOMING

Robin K. Hill is an adjunct professor in both the Wyoming Institute for Humanities Research and the Philosophy Department of the University of Wyoming, and a Lecturer there in Computer Science, and writes a blog on the philosophy of computer science for the online Communications of the ACM.

William J. Rapaport holds the position of Associate Professor Emeritus of Computer Science, as wells as Affiliated Faculty Emeritus in Philosophy and Linguistics, Member Emeritus of the Center for Cognitive Science, and Associate Director of the SNePS Research Group, all at the University at Buffalo.

The scholarly work on the philosophy of computer science that most nearly achieves comprehensive coverage is the "Philosophy of Computer Science" textbook, manifest as an ever-growing resource online,[1] by William J. Rapaport, winner of both the Covey Award and the Barwise Prize in 2015. His former Ph.D. student, Robin K. Hill, interviews him herein on that and related subjects. They start with a discussion of the proper perspective on logic from the philosophy of computer science, and on the philosophy of computer science from logic, in response to Selmer Bringsjord's commentary.[2] While Rapaport doubts that *the philosophy of* computer science has anything to say, or has to say anything, about logic, or that logic has anything to say, or has to say anything, about *the philosophy of* computer science, certainly logic has something to say about computer science proper, e.g., it can help us to understand the nature of algorithms or to verify programs. What about the algorithm?

Hill: The avowed logicist Selmer Bringsjord condemns the spotlight on the algorithm in the philosophy of computer science, claiming that the study of a DTS ("do-this-step-do-that-step" process) yields an inadequate account of computer science, which is, in his view, a study of reasoning. I admire reasoning as much as the next guy, so to speak, but he claims that "a valid deduction by the machine from this set A of formulae expressed in something that looks quite like first-order logic, to that particular formula p, is what execution of my program consists in." This is *not* the right description, in my view. We *do* need to emphasize the "DTS thingie," the sequence of steps, just as that process is respected in the low-level principles of programming languages that implement logic programming.

Rapaport: If you have a bunch of logically equivalent ways of representing or expressing something, and each has a very different "flavor," does it make sense to say that one of them is the best or right one? I think that it is better to say that each sheds different light on the common phenomenon that is represented or expressed or analyzed, enabling us to see it from different perspectives, each of which may be more appropriate or useful in different circumstances or for different purposes. (You can write programs that will behave the same using any programming language, but some languages will make the job easier or be more easily understood by humans, etc.) Gödel was not convinced by his own recursive functions or by Church's lambda calculus, but was by Turing's DTS analysis. So, rather than saying that any of these are avoidable or unavoidable, I'd prefer to say that the DTS approach opened up a fruitful way of understanding things. Logic programming and functional programming, like imperative programming, have their own realms where each is most fruitful.

Hill: Certainly, no one can gainsay the primacy of logic. Formal reasoning may be the most distinct accomplishment of the human race. The state of the world—not just now but at any time—supports Bringsjord's personal view that "reasoning is by my lights something much more valuable to teach than DTS." Logic served humanity long before computerized algorithms did. In this sense, the significance of computing is its novelty, which calls for research, whereas logic already enjoys high respectability in that regard. Should we resist the invitation to treat them as competitors?

Rapaport: Yes; let a thousand flowers bloom. Pedagogically, if you find it easier to explain computation logically rather than imperatively, go for it. If you as student find it easier to understand logically rather than imperatively, fine! Yet Dan Dennett, in his *Darwin's Dangerous Ideas* (which I'm just reading now for the first time) makes a point of distinguishing between logical approaches (specifically, the deductive-nomological scientific method) and an algorithmic approach of the sort that he claims Darwin took.[3]

Hill: I claim that an algorithm is an abstract imperative control structure,[4] with the *imperative* characteristic the most radical claim, and relentlessly DTS. Declarative structures, such as recursive definitions, are not algorithms under this view. One simply can't convey, teach, explain, or show an algorithm without telling some computing device to do something. Of course, my arguments are based on the big names—well-known and widely taught working algorithms such as Binary Search and Heapsort, which enjoy a robust conceptual life outside of their implementations in various Turing Machines. I would say that the algorithm is where no distinction holds between *satisfying* and *following* rules;[5] in other words, algorithms are not conceptually distinct from, and built upon, but rather are embodiments of, "the rules." The ascription of algorithms to a category incompatible with the declarative nature of logic disrupts the standard analyses. Now, yanking algorithms out of the declarative logic realm altogether might assuage Professor Bringsjord's concerns or it might irritate them. What do you think? Based on your deep dive into the nature of computer science, do you recommend any reconsideration of the received views of other aspects of computation?

Rapaport: You say that "an algorithm *is* an abstract imperative control structure." I agree, but I'd prefer to say that it *can be expressed as* one, but could also be expressed logically or functionally. Surely both binary search and heapsort can be expressed not only in an imperative language like, say, Algol, but also in Prolog or Lisp.

Hill: Your hospitality to philosophical views is welcome. The conflict between computationalists and others smacks of demagoguery, even hostility, and I have always been glad to follow your example of objective scholarship. How great a role does affect play? Some philosophers of computer science want human intellect to be demoted and computation raised to its proper place, while others want to see computation demoted and human intellect raised to its proper place. Should we, and can we, account for this somehow?

Rapaport: I'll use this question as an excuse to tell a story that I need to think about some more: My wife recently opened a restaurant and asked me to handle the paperwork and banking that needs to be done in the morning before opening (based on the previous day's activities). She wrote out a detailed set of instructions, and one morning I went in with her to see if I could follow them, with her looking over my shoulder. As might be expected, there were gaps in her instructions, so even though they were detailed, they needed even more detail. Part of the reason for this was that she knew what had to be done, how to do it, and why it had to be done, but I didn't. This actually disturbed me, because I tend to think that algorithms should really be just "Do A," not "To G, do A." Yet I felt that I needed to understand G in order to figure out how to do A. But I think the reason for that was simply that she hadn't given me an algorithm, but a sketch of one, and, in order for me to fill in the gaps, knowing why I was doing A would help me fill in those gaps. But I firmly believe that if it made practical sense to fill in all those gaps (as it would if we were writing a computer program), then I wouldn't have to ask why I was doing it. No "intelligence" should be needed for this task if the instructions were a full-fledged algorithm. If a procedure (a sequence of instructions, including vague ones like recipes) is not an algorithm (a procedure that is fully specified down to the last detail), then it can require "intelligence" to carry it out (to be able to fill in the gaps, based, perhaps on knowing why things are being done). If intelligence is not available (i.e., if the executor lacks relevant knowledge about the goal of the procedure), then the procedure had better be a full-fledged algorithm. There is a difference between a human trying to follow instructions and a machine that is designed to execute an algorithm. The machine cannot ask why, so its algorithm has to be completely detailed. But a computer (or a robot, because one of the tasks is going to the bank and talking to a teller!) that could really do the job would almost certainly be considered to be "intelligent." This neither demotes human intellect nor raises computation, but shows *how* (not necessarily *that*) human intellect can be computationally understood. I think it's a nice case study for Dennett's "Turing's strange inversion of reasoning," that is, "In order to be a perfect and beautiful computing machine it is not requisite to know what arithmetic is."[6]

Hill: You note, in your most recent *APA Newsletter* pieces,[7] that the Semantic Web is really a syntactic web, which provides a perfect example of your thesis. In the World Wide Web, a network of nodes and connections, the only handy meaning of a node is its location relative to other nodes, and the only possible meaning of a connection is its association of two nodes. In the Semantic Web, the new tags applied to markup elements, such as <date> or <component>, are not actual meanings in a different domain, but rather strings from the same category as the formatting markup tags such as <h1>; in other words, all those tags belong to the domain S of character strings. But how is it that we understand the <date> tag to be richer, more contentful, than the <h1> tag?

Rapaport: You ask, "how is it that we understand the <date> tag to be richer, more contentful, than the <h1> tag?" I think the crucial word here is "we." My first reaction is that *we* understand <date> more richly than <h1>, but the Semantic Web itself doesn't. My second reaction is to say that the Semantic Web might understand <date> more richly than <h1> in the same way that we do, namely, by the amount and complexity of the connections it can make with them. Presumably, both of us have more and richer connections with dates than with tags. If you're referring to the fact that the tag <date> contains the word "date," surely that's for our, human, benefit, not for the Semantic Web. (This is Drew McDermott's point in his "AI Meets Artificial Stupidity."[8])

Hill: In your analysis of the Chinese room, you point out that both the hamburger and the word "hamburger" map to neuron firings, which is the common domain. You are, no doubt, willing to accept another relation as part of the overarching domain U that associates the "hamburger" neuron firings with morphemes and the hamburger neuron firings with food. Is that how the richer connotation is captured?

Rapaport: I'm not sure what U is, unless it's the world itself, the world that contains not only my neuron firings but also certain morphemes (better?: certain sound waves) as well as certain real, perishable foodstuffs. But I would argue that I don't have direct access to parts of U: I only have access to those sound waves and those foods as mediated by my sensory organs (OK, OK: I eat hamburgers, and that gives me pretty direct access to them. But in terms of seeing, touching, smelling, etc., it's only via sensory organs.) Only God (*sive Natura*) would have such access.

Hill: Professor Bringsjord notes that "[Rapaport's] professional life to this point has passed through the evolution of the computational sciences over a period of decades, during which time a lot has happened." Indeed. He invites you to reflect on the history, an invitation that I will accept for you in asking these questions: How has your view of the philosophy of computer science matured in the time that you have been thinking about it? Compared to your mature appreciation, was your view limited at the start? Which aspects are stale and which deserve more thought? Are there some issues still under contention that are analogous to determining how many angels can dance on the head of a pin? Can you predict which inquiries may rise to prominence in the future?

Rapaport: When I created my course, I didn't know what the philosophy of computer science was, nor could I find anything about it, so I created my list of questions principally as an organizing rubric for the course:

- What is CS?
    - Is it a science? (What is science?) – Engineering? (What is engineering?)
    - Both?
    - Something else?
- What does it study?
    - Computers? (What is a computer?) – Computation?
        - What is computation?
            - What is an algorithm?
            - What is the Church-Turing Computability Thesis?
            - What is hypercomputation?
        - What is a computer program?
            - What is the relation of a program to the world?
            - Are programs scientific theories?
            - If a program implements an algorithm, what is implementation?
            - What is software, and how is it related to hardware?
            - Can or should programs be copyrighted or patented?
            - Can programs be verified?
- What is AI?
    - What is the relation of computation to cognition?
    - Can computers think?
    - What is the Turing Test and the Chinese Room Argument?
- Computer Ethics
    - Should we trust decisions made by computers? – Should we build intelligent computers?

My views have matured in the sense that I have some firmer beliefs about where I stand on some of these issues than I did when I first began looking at them (and my views continue to mature as I revise my textbook). And the two ethical issues that I chose to focus on have become much more central to the philosophical conversation than they were thirteen years ago when I created the course, what with the advent of autonomous vehicles and advances in "deep learning" AI.

I think hypercomputation might be an "angels on a pin" question. As for which questions might become more prominent, perhaps the issue of how computer programs relate to the real world that they model is one of them, or the issue that you discuss and that I formulate as "Is the form of an algorithm A to accomplish goal G merely 'Do A' or is it 'To accomplish G, do A'?" (These may be the same issue.)

***

Since 2005, Dr. Rapaport has been developing, publishing, and exercising teaching materials for the philosophy of computer science, asking and answering myriad questions connected with the exploration and establishment of that subject in the academy. In Professor Rapaport's outline of the philosophy of computer science, we see examination of every aspect of computer science. We look forward to more disciplined elaboration, complete with encyclopedic references, of this interesting subject that deploys a perspective from the humanities to reflect on the roots of technology.

NOTES

1. William J. Rapaport, "Philosophy of Computer Science," https://cse.buffalo.edu/~rapaport/Papers/phics.pdf.
2. Selmer Bringsjord, "Logicist Remarks on Rapaport on Philosophy of Computer Science+."
3. Daniel Dennett, *Darwin's Dangerous Idea*, 48, n. 6.
4. Robin K. Hill, "What an Algorithm Is."
5. Rapaport, "Philosophy of Computer Science," §12.4.4.1.2.2.
6. Daniel Dennett, "Examining the Work and Its Later Impact: Daniel Dennett Is Inspired by Turing's 'Strange Inversion of Reasoning'."
7. William J. Rapaport, "Semantics as Syntax;" "Comments on Bringsjord's 'Logicist Remarks'."
8. Drew McDermott, "Artificial Intelligence Meets Natural Stupidity."

REFERENCES

Bringsjord, Selmer. "Logicist Remarks on Rapaport on Philosophy of Computer Science+." *APA Newsletter on Philosophy and Computers* 18, no. 1 (2018): 28–31.

Dennett, Daniel. "Examining the Work and Its Later Impact: Daniel Dennett Is Inspired by Turing's 'Strange Inversion of Reasoning'." In *Alan Turing: His Work and Impact*, edited by S. Barry Cooper & Jan van Leeuwen. Amsterdam: Elsevier, 2013.

Dennett, Daniel. *Darwin's Dangerous Idea*. USA: Simon and Schuster, 1995.

Hill, Robin K., "What an Algorithm Is." *Philosophy & Technology* 29, no. 1 (March 1, 2016): 35–59. http://dx.doi.org/10.1007/s13347-014-0184-5.

McDermott, Drew. "Artificial Intelligence Meets Natural Stupidity." *SIGART Newsletter* 57 (April 1976): 4–9. http://dx.doi.org/10.1145/1045339.1045340.

Rapaport, William J. "Semantics as Syntax." *APA Newsletter on Philosophy and Computers* 17, no. 1 (2017): 2–11. Available at https://cdn.ymaws.com/www.apaonline.org/resource/collection/EADE8D52-8D02-4136-9A2A-729368501E43/ComputersV17n1.pdf.

Rapaport, William J. "On the Relation of Computing to the World." In *Philosophy and Computing*, edited by Tom Powers. 2017. Available at https://cse.buffalo.edu/~rapaport/Papers/rapaport4IACAP.pdf.

Rapaport, William J. "Philosophy of Computer Science." Latest version June 2018. Available at https://cse.buffalo.edu/~rapaport/Papers/phics.pdf.

Rapaport, William J. "Comments on Bringsjord's 'Logicist Remarks'." APA Newsletter on Philosophy and Computers 18, no. 1 (2018): 32–34.

TEACHING PHILOSOPHY ONLINE

Synchronous Online Philosophy Courses: An Experiment in Progress

Fritz J. McDonald
OAKLAND UNIVERSITY

There are a number of reasons why professors teach online. Some have decided on their own to do so. Some might have been pressured by their colleges or universities to teach online. As Peter Boltuc has stressed in this newsletter, as more and more online courses are offered, it is vital for the survival of philosophy as a profession that philosophy is taught online.[1] If philosophers do not take the opportunity to teach such courses, we risk the possibility of losing students completely who might otherwise take philosophy classes.

In my own case, I decided to teach online for practical reasons. I live in Ann Arbor, Michigan, and my university is fifty-five miles away in Metro Detroit. Saving myself a two-hour-a-day commute is quite useful for personal and financial reasons. I do love teaching, and my hope was that I could teach online while not leaving behind the elements of teaching that I love the most.

Many of my colleagues teach online, and the most common approach, I have found, is the asynchronous course. In an asynchronous online course, students can do work at any time, generally at their own pace, without any required class sessions. This is a good thing for students with complex life and work schedules. In my experiments in online teaching, I have taken a synchronous approach instead. In a synchronous online course, there are required meeting times where students and professors gather together in an online environment.

It is not my purpose here to argue that synchronous courses are superior to or even inferior to asynchronous online courses. To really study this issue, one would have to ensure that the asynchronous and synchronous courses being studied were similar in many relevant respects except insofar as they are asynchronous or synchronous. This would be difficult. It certainly seems possible that a good asynchronous course could be much better than a poor synchronous course and that a good synchronous course is better than a poor asynchronous course. How do we compare courses when the quality of instruction can differ in any class? College courses are complex things, and a lot has to be taken into consideration to make any good comparisons. Blanket judgments of relative quality between broad categories like synchronous and asynchronous courses, or face-to-face and online courses, are likely to be somewhat dubious.

As a result, I view my own work in teaching online as something of an experiment in progress. At Oakland University in Rochester, Michigan, where I am an associate professor of philosophy, I have taught several online-only courses using synchronous elements. These courses include Introduction to Philosophy, Introduction to Ethics, Ancient Greek Philosophy, and Early Modern Philosophy. All of these courses involve the teaching of historical materials, particularly because our introductory philosophy and ethics courses at Oakland University are part of the Western civilization component of our general education program. To get my students to learn important aspects of Western civilization and the history of philosophy, I require them to do extensive reading of historical texts and a good deal of writing.

At my university, we use the open-source course management software Moodle. Moodle contains resources for assignment collection, journals, forums, and a gradebook, among many other features. We also have access to the videoconferencing software WebEx from Cisco. There is other software, such as Elluminate, that can serve similar purposes. Faculty have also used popular software such as Skype or Google Hangouts for videoconferencing. E-Learning and Instructional Services (E-LIS) at Oakland University linked WebEx to Moodle, so videoconferencing sessions and recordings of these sessions can be accessed directly through the course management software.

VIDEOCONFERENCING AND THE ONLINE CLASSROOM

My aim in the courses I teach is to use the technology available to create online courses that do not lack any of the elements that might be helpful in good face-to-face instruction. To do this, I have leaned fairly heavily on the videoconferencing software. In all of my online courses, I have put in the university schedule of classes language indicating that there will be required, regularly scheduled, online course sessions on specific dates and times. One of the challenges I have faced is that students do not always read the schedule of classes completely and closely; so I have had to explain the requirement of attendance at videoconferencing sessions on several occasions to students, sometimes after they had registered, very often after the course start date.

Software such as Cisco WebEx provides many resources that are highly useful for online class sessions. I can use my computer's webcam and microphone to broadcast video and audio of myself. WebEx also allows for the broadcasting of whatever is on the screen of my laptop. This can include a presentation, whether made in Microsoft PowerPoint, Apple Keynote, Google Slides, or any other presentation software. In fact, anything whatsoever that can be displayed on a computer, including video, can be displayed live to students who are viewing the WebEx presentation.

WebEx also allows my students to broadcast audio and video of themselves to the class. One of my favorite features of the software is that the window that contains video of the instructor automatically switches to show video and broadcast audio of anyone who happens to raise their voice during the class session. This is seamless when it works well. (As noted below, it does not always work well). So if I have finished making a point, and one of my students wants to make a point herself, as soon as she starts talking, audio and video of her is broadcast to the entire class.

It is also possible to allow students or other participants in the class to capture the images on their screens and broadcast them to the class. So, if you would like your students to give in-class presentations, it is possible to have your students use the resources of WebEx to present PowerPoints or any other sort of presentation from their own computer screen to the entire class. Whatever your student puts on her desktop can be broadcast to the class.

Students do not need to use audio or video in WebEx to communicate with me and with each other. The software also contains a chat window, where the instructor and students can send messages either to the entire class or to each other. In my experience, this sometimes brings about more class discussion than is common in my face-to-face classes: my students have tended to be very comfortable with communicating through text.

In case you do not want your students to broadcast audio or video, or to participate in chat, there is a feature in WebEx that allows you to deny permission to any student in the class to utilize such features.

In addition to the ability to broadcast presentations to one's students, one can type or draw using the virtual whiteboard contained in the software. This whiteboard has its limits, and it does not allow for the kind of quick drawing one might like to use to create, say, logic diagrams. To make up for the lack of a chalkboard or whiteboard in the videoconferencing environment, I purchased a USB document camera. I use an IPEVO Ziggi-HD High Definition camera. When attached to my laptop, I can use the document camera's software to put images on my screen of whatever I happen to be drawing at a given time. This allows for the real-time writing of anything whatsoever. This seems like it would be especially useful for teaching logic—one could write out a truth-table or truth-tree live on a piece of paper for students. The WebEx feature that allows me to broadcast whatever is on my screen allows me to broadcast the live video I am capturing with my document camera.

A further useful feature of broadcasting with the document camera is the ability to display a textbook. I can show the students specific passages from the textbook to guide them to particular words or a diagram. I like to do a lot of close readings of texts in my classes, so the ability to display the text comes in quite handy.

While I require my students to attend the online class sessions, it is, of course, unavoidable that some students might miss class for emergencies. Fortunately, WebEx allows for the recording of online class sessions. These recordings contain all of the elements noted above—the video and audio of myself, the video of whatever is on my screen at a given time, audio and video of my students, and the text from the online chat sessions.

Were this sort of online environment to work perfectly, I cannot think of any element of a face-to-face class that would be missing from this kind of online environment. This might help overcome objections from those who think philosophy can only be done best in the face-to-face environment. Lecturing, one obvious element of the traditional face-to-face class, is easily made possible. Perhaps more importantly, discussion can occur as well. The professor can use presentation software to convey information, and so can the students. A document camera allows one to use a writing surface to present text, diagrams, or pictures. Videoconferencing even allows for additional elements that are not part of the typical face-to-face classroom environment, such as the chat feature, which allows professors or students to send text to the entire classroom or to each other. As I mentioned above, this chat feature proves really useful for spurring discussion. As Frank McCluskey has noted, "It is well known that introverts prefer writing to speaking. So here is medium where they are able to take chances in ways they might not in the bricks and mortar classroom."[2] In my experience, this is true of live online chat writing.

Videoconferencing is not the only online resource I have found helpful. The many elements of the Moodle course management software we use at Oakland University come in quite handy. For example, as many professors have encountered, it can be difficult to get students to do the reading. So I use the journal feature on Moodle. I require students to write about each reading before we discuss them in the online class session. These journal entries are graded, and they are a quite useful way in any class to get students to do the reading.

The kind of synchronous class I offer hence has the potential to contain all of the relevant features of face-to-face classes, and there is no reason not to also use all of the useful elements that have been part of asynchronous online classes as well. Nothing prevents the teacher of the synchronous class from also including recorded lectures, using an online forum for discussion, or using a software program to provide guided instruction. Anything that can be done in an asynchronous class can obviously be done in a synchronous class as well. So it would seem that synchronous online classes have their own advantages while not missing any of the advantages of asynchronous online instruction or face-to-face instruction.

LIMITS OF THE SYNCHRONOUS CLASS

While, in the ideal, there should be nothing lost in a synchronous online course, I have faced a number of difficulties over the past five years of teaching online.

I face the challenge of conveying the importance of synchronous class sessions to students who are accustomed to taking asynchronous online courses. I have had students balk at the requirement to show up regularly for online classes. One of my students told me early in a semester that my requirement that students show up at a regular time for online classes was the stupidest thing he had ever heard. There have been complaints about my synchronous approach to teaching courses, including in my teaching evaluations. Even students who have spent an entire semester registered in my class without dropping have complained about the synchronous element in their student evaluations. It seems to me that a number of students have had the impression that it is almost by definition that an online class is an asynchronous online class, where one can work on one's own schedule without a required class time.

I am sympathetic to the needs and concerns of students who would be best served by asynchronous classes. Some students have work or life schedules that prevent them from being available regularly at a given time. I have had students in the military whose work prevents them from involvement in such a class. All that being said, I do indicate to my students, both in our schedule of classes and immediately at the beginning of the class, that there are mandatory online videoconferencing sessions. This sometimes leads to students dropping my section of the class. Even at a university with a robust online program such as Oakland University, this might possibly lead to lower student demand for online courses with required synchronous elements. This is a concern. As noted above, Peter Boltuc has rightly stressed the importance for philosophy as a profession of offering online courses that appeal to students. The administrators at my university are certainly interested in all of my courses having a heavy enrollment!

In my experiences teaching online, I have not been able to achieve the ideal environment I had hoped for. As a result, I do think, until student preparation and technology advances, there are some significant limits to the use of synchronous online teaching sessions in courses.

I base the following points only on my own experiences. Others might have had different experiences, more positive or negative. This is a place where greater study and collection of data across the profession might be useful for improving online instruction.

I find that students are hesitant to use the audio and video capturing features available from their own computers. I often have to tell students to turn on their webcams or microphones to make video and audio of themselves available. WebEx does not automatically turn on webcams and microphones. Even after insisting on the students turning their equipment on, there are still a significant number of students in my courses who never turn on their microphones or webcams. I suspect I have had a few students who do not even have webcams and/or microphones on their computers.

When the relevant equipment is turned on, the students are not always in the best environments for philosophical discussion. Unlike the controlled environment of a classroom, a virtual classroom composed of video and audio captured from a number of students' computers can contain many distractions. Friends and family can often be heard in the background. Crying babies are not uncommon. Noises of all sorts interfere with the discussion.

Joshua Kim has noted issues of this sort in *Inside Higher Education*:

> The other audio problem that I see is feedback. Students forget to mute their computer audio, causing terrible echoes. Or background noise, the barking dog or busy office, makes it difficult for everyone to hear. Online meetings are often loud and distracting events. I'm not sure how to achieve a quieter experience.[3]

To best allow for really good discussion, all students should be in a relatively quiet environment with minimal distractions. Getting actual, real-world students to do this during synchronous online class sessions is a difficult task.

A number of students will not have excellent presentation skills. We are not all accustomed to broadcasting ourselves to others. I have encountered some difficulties with poorly placed webcams and microphones. A number of students do a poor job of using the technology on hand to communicate. This is not a skill that all students have practiced or learned.

There is still a serious concern of a digital divide. Not every student will have the computer, webcam, microphone, and high-speed internet access required for videoconferencing. There are also issues of disability and accommodation. Elements of online learning that might best serve disabled students might not serve other populations of disabled students well. A student who might be best served by a text-only online course might be at a disadvantage in a text, video, and audio class such as mine.

An instructor might also, by teaching a course using videoconferencing technology, lose some of the historical advantages of text-based online education. Ron Barnette has noted the ways in which an online environment might be useful insofar as it removes any sort of influence factors like race, gender, or age might have upon listeners to an open discussion.[4] If students are using webcams and audio, some of the effects of factors such as this might reoccur.

In the ideal environment, the technology currently available ought to allow for online courses to have all of the positive features of traditional face-to-face classes. Terry Weldin-Frisch has expressed a similar opinion of synchronous online instruction in this newsletter:

> There is simply no good reason to think that this particular educational delivery method cannot be perfected in such a way that long held predispositions regarding the second-rate quality of online philosophical education need continue to dominate a twenty-first century educational milieu.[5]

Perhaps in the near future this perfection will be achieved. Even in the present day, where technology is such a key part of our lives, a cultural shift is required, with students being more prepared, both personally and with their technology, to bring to the table what is needed for a good synchronous online class.

It is also worth noting that not all universities and colleges have put the considerable resources that my own university has put into online education. I have been provided with some of the best hardware and software available for teaching online, and I recognize that I am fortunate to have it.

THE FUTURE OF ONLINE PHILOSOPHY EDUCATION

It is exciting to think of what technology might make possible. More and more, computing is done not on desktop or laptop computers but on smartphones, tablets, and wearable devices. If the kind of technology I am using now is able to be ported well to such devices in the future, synchronous online videoconferencing classes might be even more accessible in a wide range of environments, including for students who have cut the cord on broadband in favor of cellular-only data.

With many challenges having been noted, I have had several good experiences teaching a number of students online, and I have found the experience of teaching using videoconferencing software to be often as rewarding as the experience of teaching in the traditional face-to-face environment. At times, it has even been better. When videoconferencing works, it makes possible in online classes what has henceforth only been possible within the physical classroom.

NOTES

1. Peter Boltuc, "e-Teaching Philosophy," *APA Newsletter on Philosophy and Computers* 5, no. 1 (2005): 3–5.
2. Frank McCluskey, "Reflections from Teaching Online," *APA Newsletter on Philosophy and Computers* 11, no. 1 (2011): 23.
3. Joshua Kim, "Running Synchronous Online Classes," *Inside Higher Education*, August 28, 2013.
4. Ron Barnette, "Reflecting Back Twenty Years," *APA Newsletter on Philosophy and Computers* 11, no. 1: 20–22.
5. Terry Weldin-Frisch, "A Comparison of Four Distance Education Models," *APA Newsletter on Philosophy and Computers* 11, no. 1 (2011): 25.

The Paradox of Online Learning

Adrienne Anderson
UNIVERSITY OF ILLINOIS, SPRINGFIELD

Marshall McLuhan observed that "we shape our tools and afterwards our tools shape us." The truth of this statement has never been more apparent to me than after reflecting on my e-learning experience. The time spent in the digital classroom studying philosophy has affected me, my way of being, and my worldview, which is truly valued. However, in contemplating my experience, it is clear that for every pro, there is also a con. The strengths of online learning seem to be, at this moment in time, also its weaknesses. The purpose of this paper is to shed light on the advantages and disadvantages of e-learning in pursuit of progress for prospective students of cyberlearning and facilitators within the field.

Similar to the online gaming world where you are free to be and act according to your desires, digital learning provides students the liberty to explore ideas and fully express them in an environment where physicality is completely removed. Engagement occurs strictly within the infosphere where inhibitions that accompany age, race, class, and/or physical threats are significantly lessened. Students participate in the safety and convenience of their own home. Consequently, thoughts and ideas can be expressed openly and without the physical fear or anxiety that is sometimes experienced in traditional classroom settings.

At the same time, cyberlearning completely ignores the physical body. Experiential learning, which promotes a unique type of knowledge, is nonexistent. Signals that the body receives from others in a classroom environment, which can serve to inform the mind, is just not available online. Passion of experience, which is retained and reflected in the body, is excluded from the online experience. The physical aspect of learning is a missing piece within the digital space.

Another feature of traditional learning that is absent in online learning is the time-honored role of a professor who lectures and disseminates information into a bunch of empty heads. Rather, instructors utilize the online forum to post critical questions that promote guided discussions. Students are free to read and take in information at their own pace. Responses are then shared openly in an online forum wherein each student initiates a discussion, as opposed to following one main discussion led by an all-knowing teacher. This allows for dynamic discussions to take place among students and facilitator. Contributions to different discussions occur during the week within multiple threads, allowing time for reflection and integration of new information. Learning is enhanced from the perspective of others, not just a single Mind who expects information to simply be parroted back.

As a result of an exorbitant amount of time reading and writing online, those skills are remarkably enhanced. As a philosophy student, the ability to present a clear, concise argument and to engage others online becomes second nature. Writing online provides immediate feedback due to the ability to gauge your work and the work of others' and allows for self-assessment and redirection if necessary. This approach, at first glance, is liberating and revolutionary. However, as my education progressed, it became clear that there are significant drawbacks to this approach as well.

I was shocked to find out that I had difficulty expressing into words the material I had so passionately studied online. In natural conversations in everyday life with family, friends, and coworkers, I found that the philosophical concepts and ideas I had read and written about were sometimes swirling above my head. It was difficult to filter them down into coherent sentences. Outside my experience as a student, my entire adult career has been as a court reporter. I listen to sounds and carry those sounds through my hands into the written word, completely bypassing my mouth. The online college experience is no different. It is extremely reading and writing intensive; however, it omits speech entirely in the process. Verbal expression of ideas is entirely lacking. My mind quickly turns to a story from the book *Mirrors* wherein it states, "Hunger, which kills silently, kills the silent. Experts speak for them, poorologists who tell us what the poor do not work at, what they do not eat, what they do not weigh, what height they do not reach, what they do not have, what they do not think, what parties they do not vote for, what they do not believe in."[1] The poor, the marginalized, the weak, they do not have a voice. In order to be heard, it is necessary to be able to speak. With this revelation, it is clear that an oral pathway must be formed and grounded in reality if students want to be fully empowered in the world. Therefore, it is absolutely necessary that verbalization is more explored in the e-learning community.

If it were not for e-learning, however, I would not have been afforded the opportunity to expand my education and, consequently, worldview. Its accessibility, flexibility, and affordability cannot be found in brick-and-mortar schools. As a full-time working, single mother, I was able to manage my time and availability in accordance with my hand-picked class schedule. Adhering to a set schedule dictated by an institution would not have been viable. Although online learning is a solitary and sometimes lonely endeavor, it is also extremely rewarding and gratifying. It opened me up to worlds I never would have imagined existed. The ideas I have been exposed to are now navigating my life in a new direction that seeks to augment reality with the help of the digital world, which, ironically, is the whole point of online learning.

So is e-learning perfect? No, but neither is traditional learning. They both have their pros and cons. However, online learning is a convenient entry point into the educational system for those who cannot otherwise access it. Although online learning lacks the physical aspects of engagement such as the essential ability to orally express ideas, it nevertheless provides a pedagogy that provides a new depth of interaction that involves the ability to follow and explore multiple threads of thought and ideas from a variety of perspectives. As reality transforms due to technological advances, e-learning serves as a gateway to the masses to navigate in this new world. As we transition from a separate offline life to a fully integrated digital life, the advantages and disadvantages will, no doubt, become a moot point.

NOTES

1. Eduardo Galeano, Mirrors: Stories of Almost Everyone (Nation Books, 2009), 125.

Sustaining Success in an Increasingly Competitive Online Landscape

Jeff Harmon
SOUTHEAST MISSOURI STATE UNIVERSITY

The line between for-profit and not-for-profit universities is being blurred everywhere as state or publicly funded universities are receiving less funding from their respective governments. While the online education market is global, the US market is more mature, and can serve as a good example for European and other universities to learn from. But even as these differences become more profound, some aspects of this industry segment have not changed at the rate that many of us expected.

Recently, I was searching for new insights in online education marketing. To my surprise, when I searched for those terms on Google, the first result below the ads was a link to my presentation from the 2004 University of Wisconsin distance education conference. Much of what I discussed included many of the options available today such as search engine optimization, pay per click ads, and content marketing. For this article however, we will start the discussion at a higher level—the institutional brand.

THE BRAND

Your institution brand is not just a logo. It is comprised of so much more. The brand is the complete experience. And that experience is created and delivered by the staff and faculty of your institution. In the brand training sessions I lead, I try to instill the thought that each of us has a responsibility to treat our students (customers) professionally and consistently. Yes, the students are customers who can purchase their education elsewhere. And online, the process of choice is very convenient. Think about your favorite retail store or restaurant and how it makes you feel special every time you frequent the establishment. That is what we should be striving for as an experience for our students. So how do we deliver on a brand promise in the online environment? For starters, let's reply to student emails promptly and also ensure our technology support team does the same. There is nothing more frustrating for students than technology that is not working at the time they need it most.

Communication style is another way by which we convey the brand. At Southeast Missouri State, we communicate our brand by using particular personality terms that we believe bring our brand to life. Your institution probably has personality words as well. But it is not just the words that matter. The tone, particularly in personal communication, can have an impact on your students' perception of your institution. When good students transfer and share the negative experience in social media platforms, the results are certainly negative for the brand.

LET'S TALK MARKET RESEARCH

The importance of market research cannot be overstated. Understanding your prospective and current customers is invaluable, and too often dismissed due to time, budget, or expertise. Not only should we be researching how to best reach our prospective students, but we also must prove that there is demand for any new degree programs prior to developing plans to launch.

A recent study regarding the online student market by the Learning House and Aslanian Marketing Research[1] revealed a few important statistics:

- 65 percent of online students choose an institution within one hundred miles
- Direct mail is very effective
- Affordability and accreditation are the top messages
- Follow-up with prospective students within twenty-four to forty-eight hours is imperative

I was surprised that direct mail is considered effective in this environment. After a few conversations with colleagues around the country, we are all experiencing a resurgence of effectiveness of direct mail. It seems there is less noise in this channel and thus, it is working quite well.

LANDING PAGES AND TRACKING

Maximizing return on investment from marketing tactics requires special attention and effort. Many things must be in place. First, always consider creating special landing pages for campaigns. If the landing page is similar to an ad a prospective student clicked on or watched, they will know they are at the right place. The chance of them taking an action is dramatically increased because of this. If you cannot create a special page, please ensure that your online degree pages are consistent and succinct, and include phone, chat, a web form, etc. Make it as easy as possible for prospective students to connect with you.

You also need to ensure you are maximizing the potential of your everyday online degree website. The home page and all of the degree pages must be well optimized for the search engines. While great search results are not easily obtained, you will be surprised at how much more visible your pages can be with a little effort. Let me give you an example of why this is so important.

Generating leads with pay-per-click ads can cost hundreds of dollars per lead, especially if you do not have the expertise to properly target the ads and set up campaigns in Google or Facebook. Look for a Google Partner to assist you at https://www.google.com/partners/. Listing your degree programs on lead aggregator sites is an easier alternative. Consider sites such as GetEducated.com or US News Degree Finder. These leads will cost $50 to $150 each.

But why is search engine optimization (SEO) so important? SEO is important because it will generate organic leads from prospective students who are genuinely interested in your degree program and institution, and will convert at a rate much higher than leads from ads or other websites. Organic leads are naturally more familiar with your brand and want to connect. Secondly, they cost practically nothing. While at University of Illinois Online (UIOL), we optimized our home page for the phrase "online degrees." We stayed in the top five results on Google for three years. That effort generated nearly 1,200 leads per month for free. It could cost more than $120,000 per month to generate or purchase that number of leads.

The first step was to create an inquiry form accessible from each degree program description page. Then, work to optimize every single page for the search engines. Optimization includes the following basic tasks:

- Check Google Trends (https://www.google.com/trends/explore#) to ensure you are optimizing for the best terms
- Include the degree and institution name (which should include the top terms) in the page title.
- Ensure that the best terms are also used at least three times in the page content. The more they are used the better.
- Create alt tags for images and use the same top terms.
- Create a news section on the homepage so that Google sees fresh content appear. If your content gets stagnant, Google knows and will lower the page rank.[2]

You should see improvement in the search results within a couple of weeks. Keep tweaking and updating content for continual improvement.

TRACKING SUCCESS

In order to get the most out of every advertising dollar (or lack thereof) you need to track return on investment (ROI). There are several ways to accomplish this. For every web ad (Google, Facebook, YouTube, etc.), you will be asked for a click destination URL when you place the ad. Let's say you want visitors to go to http://semo.edu/admissions. The URL you submit will need to be more complicated than that in order to provide tracking data. You need to add variables describing the source, content, and campaign, etc. And you need to set up your customer relationship management (CRM) system to receive the data and save it with each prospective student record. A tracking URL for the page listed above could be http://www.semo.edu/admissions**?utm_source=facebook&utm_medium=web&utm_content=LowTuition&utm_campaign=SEOnlineFall2016**.

Notice that everything in bold is for tracking. If you only have one variable, and that variable is "source," then the additional text you add at the end could be simply. ?utm_source=facebook. You can add as many variables as you wish. By doing this, you will be able to see what campaigns, ads, and even specific messaging produces the highest ROI and thus the best results.

In order to calculate marketing ROI on a campaign, you need to establish the average value of a student. So let's say your financial office has established that the average value of an online student is $10,000. Calculate the campaign ROI as follows:

Campaign cost = $100,000

New Students Enrolled = 30

Revenue Value of 30 Students = 30 x $10,000 = $300,000

$$\frac{\text{Value of Students - Marketing Cost}}{\text{Marketing Cost}} \qquad \frac{\$300{,}000 - \$100{,}000}{\$100{,}000}$$

ROI = 200%

MARKETING CHANNELS TO CONSIDER

There are many ways to reach prospective online students, and not all of them are online. Naturally, web ads are a great way to start, but an integrated approach will always work best. For example, direct mail is great for driving traffic to a website. If an admissions representative is going to attend an education fair at a large company in a particular city, it would be a good idea to place web ads in that geographical location a week or two prior to the event. The ads would be targeted to the audience the representative will interact with at the fair in order to increase awareness and interest in advance. The result will be much more traffic to the representative's table.

Recommended channels for consideration include the following:

- Google ad network (includes desktop and mobile)
- Facebook ads and post boosts
- You Tube commercials
- Twitter
- Instagram
- Pinterest
- Pandora
- Direct Mail
- Radio
- Television
- Billboards
- US News Degree Finder
- Get Educated.com

You might be wondering why Spotify is missing. Spotify requires a minimum expenditure that is quite high and would not make sense for most institutions. Google, Facebook, and You Tube should be part of any marketing mix. You can easily target and maintain tight control on your budget. Tracking is easily managed as well. If you have a slightly larger budget, consider adding Pandora audio and digital ads, Twitter, and Instagram. Larger brands such as Apple are really starting to view Instagram as a powerful brand builder. Pinterest may seem an outlier, but many brands are seeing it as a requirement. Direct mail is still very effective but can be expensive. And TV and radio, while effective, require a very large budget in order to be effective. Billboards are great for awareness, but there is no way to track effectiveness unless your team is very good at asking each new student how they found out about your programs.

INFRASTRUCTURE

Every single thing that has been discussed in this article should not be implemented if you do not have the proper infrastructure in place. And by that I mean software, people, and processes. Institutions waste hundreds of thousands of marketing dollars due to lack of infrastructure.

First, institutional administration must all be supportive and clearly understand the complex nature of offering online degrees and the effects on capacity of traditional face-to-face students who wish to take an online course. In fact, many traditional on-campus students take at least one online course per year, and the trend is growing. About 5.8 million students were enrolled in at least one distance learning course in fall 2014—up 3.9 percent from the previous fall, according to "Online Report Card: Tracking Online Education in the United States," an annual report by the Babson Survey Research Group.[3] This trend causes many problems for enrollment managers for online degree programs. Colorado State University attempted to make it easier to manage the online degree enterprise by launching Colorado State University's CSU Global Campus. It is a completely separate entity, free from restraints of managing seats for on campus and online students. This is just one example, and I believe the level of success of this venture cannot yet be measured. The University of Illinois failed when it launched the U of I Global Campus. There could be many reasons for the success or failure of these entities.

Other issues that may arise when attempting to grow online programs tied to the main institution include developing a good strategy for using adjunct faculty, and establishing course content that can be used without strings attached by strict agreements with faculty. In other words, efficiency is key. There is no need to market programs if an institution does not have a solid plan addressing all of these areas that will constrain growth.

If all of these issues regarding capacity (product availability) are taken care of, congratulations, you are ready to put a team together and give them the tools to succeed.

First, identify if you have a Customer Relationship Management (CRM) system in place through your admissions team that you can leverage. If not, they probably need one, and you should consider Salesforce.com. Salesforce has plugins for many marketing tools that will allow your marketing team to be incredibly efficient. Salesforce also has a foundation that allows nonprofits to use up to ten licenses for free. You will need to do some customization to give the system a higher education nomenclature, but it can certainly be done. There are also third-party companies who have already customized Salesforce for higher education, but their versions come with a hefty price tag and annual fees.

You probably have great advisors that can get applicants and current students where they need to go to be successful, but practically no nonprofit understands the realities regarding "selling" online degree programs. The market has become very competitive; just complete a lead form at University of Florida Online, Arizona State University

Online, or try a for-profit such as Grand Canyon University. Once they have your mobile number and email address, I think you will fully understand the competitive nature of the online education industry. Once these institutions receive a lead from a prospective student, they begin calling within just a couple of hours. You may even get a call within fifteen minutes. Once, I secret shopped University of Phoenix and received eight calls in twenty-four hours. While you do not need to be this aggressive, you do need to understand what the competition is doing.

So how do you compete? You either need to hire enough admissions reps (sales reps) in-house or consider outsourcing. A couple of trusted firms to keep in mind are Perdia Education and Barker Education Services Team. Perdia just launched an enrollment management mobile app that ensures prospective students move through the enrollment process efficiently and easily. Because of this, the conversion rate increases dramatically. Pearson Education has been a large player in this area as well as all other aspects of higher education for some time. They may also be a viable solution.

Lastly, I want to share some numbers. If you plan to grow, you need to clearly understand how many admissions reps you need to get you to your goal. One rep can handle 150 leads per month—converting to ten to fifteen enrolled students per rep on average.

Five hundred to 1,000 older leads can be managed per rep while working new leads. These numbers have been vetted by discussions with higher education call center professionals.

So, after reading this article, are you ready to take your institution's online degree program enrollment to the next level? I think you have much to think about.

NOTES

1. David L. Clinefelter and Carol B. Aslanian, *Online College Students 2016: Comprehensive Data on Demands and Preferences*. Rep. Louisville: Learning House and Education Dynamics, 2016.
2. Christopher Ratcliff, "Search Engine Watch," Guide to Google Ranking Signals Part 4 Content Freshness Comments. *Search Engine Watch*, September 26, 2016.
3. I. Elaine Allen, and Jeff Seaman, Online Report Card Tracking Online Education in the United States. Rep. Online Learning Consortium, February 2016, http://onlinelearningconsortium.org/read/online-report-card-tracking-online-education-united-states-2015/.

CALL FOR PAPERS

It is our pleasure to invite all potential authors to submit to the *APA Newsletter on Philosophy and Computers*. Committee members have priority since this is the newsletter of the committee, but anyone is encouraged to submit. We publish papers that tie in philosophy and computer science or some aspect of "computers"; hence, we do not publish articles in other sub-disciplines of philosophy. All papers will be reviewed, but only a small group can be published.

The area of philosophy and computers lies among a number of professional disciplines (such as philosophy, cognitive science, computer science). We try not to impose writing guidelines of one discipline, but consistency of references is required for publication and should follow the *Chicago Manual of Style*. Inquiries should be addressed to the editor, Dr. Peter Boltuc, at pboltu@sgh.waw.pl.