More than one hundred years ago Charles Sanders Peirce wrote a series of articles for the *Monist* which today are known as his Metaphysical Series. They came near the end of a long and very productive intellectual career which began with a series of articles about the nature of inquiry and belief which introduce an idea for which he is today famous—the pragmatic maxim. Peirce was critical of his original formulation of the maxim and changed the name he had given his idea from pragmatism to pragmaticism. The general idea that interests us in this paper is how a small handful of Peirce’s writings introduce the broad outlines of a fundamental theme in cognitive science, that of functionalism, and then how important developments in his thought culminating in the *Monist* series both anticipate some of the most significant junctures in the brief history of functionalism’s theoretical development and propose a novel and far-reaching solution to contemporary debates about its nature and viability as a general theoretical model with application in the social sciences.

We will begin by sketching the philosophical context in which Peirce’s famous early papers were written and the conclusions he came to, specifically, the pragmatic maxim and his doctrine of thought-signs. We will show how this mirrors to a remarkable degree the philosophical debates that surrounded contemporary functionalism’s birth as well as its basic theoretical tenets. We then move to an
examination of the impact functionalism as had in semantic theory, specifically with the theory of functional-role semantics, perhaps the most popularly held position among the majority of cognitive scientists and workers in artificial intelligence. After a discussion of some fundamental problems this position faces we will return to Peirce and his critique of the original formulation of the pragmatic maxim. We believe that the pragmaticism that he then expounded responds to the difficulties just alluded to. Specifically, we will claim that functional-role semantics must widen its ontological commitments beyond the currently fashionable physicalism (which implies the nominalism that Peirce so despised) in order to incorporate the counterfactual quality of mental states which in the contemporary debates which functional-role semantics faces has found it necessary to admit.

Peirce and the Pragmatic Maxim

Peirce began his philosophical career not in the armchair but rather in the laboratory. Almost all of his writings reflect in some way or other his experience in scientific research and attempt to elaborate in piecemeal fashion an answer to the question: “How should we go about investigating the world?” In Peirce’s day the logic of scientific discovery was understood in terms of deduction and induction. Scientific work accumulated observations which induction generalized into laws or general principles that could be used to deductively determine individual cases that fell under them. But Peirce insisted that an important element was lacking in this picture, that hypothesis, or what he called abduction, is an indispensable part of the process of inquiry. We confront a phenomenon that does not fit into our current explanatory scheme. We propose a hypothesis to explain it, deduce the consequences of the hypothesis, and test them inductively in our experience.

It was in the context of this framework that Peirce introduced his now famous pragmatic maxim:

Consider what effects that might conceivably have practical bearings we conceive the object of our conception
He proposed it as a logical tool for the determination of the admissibility of hypotheses. It does not tell us which hypothesis is correct, but only in what it consists and how it can be distinguished from others. The functionalist character of the maxim is expressed quite clearly in a comment Peirce makes in an earlier article in the context of a discussion of George Berkeley. He says,

A better rule for avoiding the deceptions of language is this: Do things fulfill the same function practically? Then let them be signified by the same word. Do they not? Then let them be distinguished. (CP 8.33)

Although for Peirce the maxim exercised a limited and very specific role in his vast architectonic, it constituted a major challenge to the modern epistemological paradigm by completely breaking open the mysterious regions of the Cartesian mind in which our ideas were held in some occult way to reside, giving them an accessibility and publicly observable quality until then occluded by competing epistemological theories grounded in what Peirce argued to be the mistaken Cartesian assumptions of the nature of mind and reality. Instead of asking what an idea is, Peirce asked what it does, which, as we will see, anticipates in a clear way the functionalist doctrines of the late twentieth century.

But the maxim is only part of a larger critique that Peirce makes of the modern epistemological paradigm. Descartes’ substantialist metaphysics generated the problem of how the res cogitans is to relate to res extensa in the formation of judgments of which we can be certain. Descartes’ answer was that this certainty was attained in acts of intuition, utilizing of course the criteria of clearness and distinctness. By demonstrating the practical inefficacy of these criteria Peirce profoundly problematized the intuitionist model and proposed instead an inferential one. Here, mental activity is framed in terms of inferences from external facts. There is no substance, or first premise, which grounds a series of inferences. Rather, the self itself is an inference created as a hypothesis to explain external facts and is embedded in

---

1CP refers to the six-volume Collected Papers of Charles Sanders Peirce. The numerical citation refers to the volume number, followed by a period, followed by the paragraph number within that volume.
a series of such inferential chains in a community of inquirers. This model leads Peirce, in “Some Consequences of Four Incapacities,” to a decisive break with modern forms of dualism. One of the consequences of our incapacity to intuit or introspect in an immediate manner is that the consequent inferential nature of cognitive activity equates man with sign processes. He says,

[T]here is no element whatever of man’s consciousness which has not something corresponding to it in the word; and the reason is obvious. It is that the word or sign which man uses is the man himself. For, as the fact that every thought is a sign, taken in conjunction with the fact that life is a train of thought, proves that man is a sign; so, that every thought is an external sign, proves that man is an external sign. That is to say, the man and the external sign are identical, in the same sense in which the words homo and man are identical. Thus my language is the sum total of myself; for the man is the thought. (CP 5.314)

And,

The psychologists undertake to locate various mental powers in the brain; and above all consider it as quite certain that the faculty of language resides in a certain lobe; but I believe it comes decidedly nearer the truth (though not really true) that language resides in the tongue. In my opinion, it is much more true that the thoughts of a living writer are in any printed copy of his book than that they are in his brain. (CP 7.364)

In the Discourse on Method, Descartes takes pains to distinguish the human from the animal and mechanical orders. An automaton could be built that mimicked human behavior, but lacking consciousness it would be a mere machine because it “could never use words or other signs, composing them as we do to declare our thoughts to others.”

In this view, signs are part of the world of extended things that the *cogitans* uses to express its ideas. By conflating man with signs Peirce strikes to the heart of the Cartesian philosophy, and with the later elaboration of his pansemioticism problematizes the entire approach to epistemological and cognitive issues that conceives of man and consciousness in substantialist terms, or in terms of any of the variety of its conceptual cognates.

Reading the above quote, one cannot help thinking that had computers existed in Peirce's day, the field of artificial intelligence would surely have begun to take shape a full century before, for here one finds the radical rupture with modern metaphysics that makes such thinking possible. Curiously, although most consider Alan Turing the grandfather of computing and artificial intelligence, his famous test for intelligence, understood as an operationalist maxim, could easily be seen as only another formulation of the pragmatic maxim. The important point, however, is that in a few seminal papers Peirce anticipated the basic functionalist idea that many things, such as human cognition and social systems, can be understood and differentiated in terms of the practical effects or functions which these things bring about. His later critique of the popular interpretation of the maxim will provide us a way of reforming the problematic ontological commitments of functionalism, but first let us examine the context of functionalism's birth and its development in contemporary semantic theories.

*From Functionalism to Functional-Role Semantics*

Contemporary functionalism grew out of circumstances similar to those of the pragmatic maxim. As Peirce found dissatisfactory the operational ambiguity of the Cartesian criteria, so philosophy of mind, in the 1950s, found itself with an attractive but problematic proposal today known as type-identity theory. It was attractive because it naturalized consciousness, making mental phenomena supervene on physical brain states (a position that most contemporary philosophers of mind and cognitive scientists hold), but problematic because it held each *type* of mental state (e.g. seeing red, or feeling pain in general) to be identical with or reducible to a particular type of physiological state of the brain. For obvious reasons, this dissatisfied people interested
in the possibility of artificial intelligence. But for purely philosophical reasons it appeared untenable because it seemed to make the red or the pain that other species see and feel something qualitatively distinct. And this just seemed counterintuitive. The clear alternative would have to be that mental states are multiply realizable, and this is just what Hilary Putnam’s proposal, which he called functionalism, did. The ontological commitment continues being physicalist but it restricts it to tokens of mental states, not types. Thus token-identity theory allows that a type of mental state can be realized in different types of physiological brain states.\textsuperscript{33}

But the novelty of the functionalist proposal was not its contribution to ontology but rather to semantics. It gave to researchers in psychology, cognitive science, and artificial intelligence an operational handle, much as did Peirce’s maxim, in the sense of offering an experimentally testable procedure for distinguishing between mental states in terms of the functions or roles that these states involved, thus providing a path through the Scylla of the crude experimental parameters of behaviorism and the Charbydis of the metaphysical abstractions of the philosophy of mind.

The general movement toward naturalization in theories of cognition has naturally manifested itself as well in the field of semantics. Among the leading theories in the naturalization camp one finds functional-role semantics, which seems to be the preference of many in the cognitive science community. In an incisively argued study, Botterill and Carruthers make the case for a functionalist notion of meaning by evaluating, in turn, the deficiencies of informational and teleological semantics.

Informational semantics holds meaning to consist in particular causal relationships between mind and world, so that “for a mental term ‘S’ to mean S, is for tokenings of ‘S’ to causally co-vary with Ss.”\textsuperscript{44} The idea is, to use their example, that the term ‘mouse’ means mouse because a tokening of ‘mouse’ in belief is reliably caused by the presence of mice, and only by the presence of mice. The problem informational semantics faces, appears to be similar to that which Frege tried to

\textsuperscript{3}Incidentally, the type/token distinction was one that Peirce originally introduced.

\textsuperscript{4}George Botterill and Peter Carruthers, \textit{The Philosophy of Psychology} (Cambridge: Cambridge University Press, 1999), p.163.
resolve in “On Sense and Reference.” If meaning is a function of a causal relationship with its referent, then we have the problem of explaining the nature of non-tautologous identity statements as he famously did with the morning star and the evening star. Frege responded by distinguishing between sense and reference. Informational semantics has responded by simply including in the referent all possible states of affairs in the world that could token a given mental term. But this is a cumbersome way out that provides no principled criteria for delimiting the referential extension of a concept in any given case.

A proposal that has responded to this problem has been teleosemantics. It comes quite close to the functionalist model by supposing that the mind, like the body, has evolved under certain conditions, such that given elements of the mind have given functions which correspond to them. The idea is that certain mental states, concepts, or ideas have in the past proven successful, which in this theory implies that the meaning of the mental state in question, or its function, is simply the way of acting that has proved successful.

On the surface this does indeed provide a principled way for determining mental content, but Botterill and Carruthers illustrate the problem it faces with the example of the snapping reflex of a frog. We would normally say that this reflex is triggered by the perceptual representation of a fly, such that, for the frog, the representation of flies is the function or “meaning” of the frog’s percept. But they ask if this can be vindicated in terms of function. Why flies, and not small black things or shot-gun pellets (for these also cause the snapping reflex)? It could equally be said that the function of the frog’s percept is to represent small black things, “since that state has evolved because of its effect in causing sniffings which have led (given the fact that all small black things are flies) to enhanced survival.”

The problem of delimitation returns to haunt teleosemantics. As these authors note, Fodor has argued that the notion of function can be used to make distinctions between flies and shotgun pellets only in counterfactual terms, and not in terms of the actual selectional history. That is, had the frog lived in an environment where all small moving black things were shot-gun pellets, the perceptual state in question would not have been evolved. “Then if the function of a state is those

of its effects which insure its existence in actual *and counterfactual* circumstances, we can say that the function of the frog's movement detector is to represent *flies* and *not* either *small black things* or *flies-or-shot-gun pellets*."\(^6\)

Given the problems with a historicist functionalism, functional-role semantics has been proposed in resolutely synchronic terms, wherein the content of a mental state is determined in terms of "its functional, or causal, role within the system—where to characterize the causal role of a state is just to describe the characteristic pattern of causes and effects which it normally has within the system..." (p. 176).

The notion of 'effects' recalls Peirce's pragmatic maxim wherein mental content is cashed (to use William James' unfortunate phrase) in terms of the conceivable practical effects of the object of our idea. Peirce came to see the unattractive consequences that his maxim implied and worked to distance himself from them in the formulation of his pragmatism. As we will see, the unattractive consequences that the functional-role model will face is the arbitrary or observer-relative determination of function. In order to avoid it, it will have to take recourse to a position that we consider very similar to Peirce's pragmatism. But before discussing this let us turn to an examination of Peirce's critique of his maxim and the solution to it he proposes.

*From Pragmatism to Pragmaticism*

In the same article where the maxim is introduced Peirce applies it to a particular conception, that of hardness. He says, "what we mean by calling a thing hard" is evidently that it will not be scratched by many other substances. The whole conception of this quality... lies in its conceived effects. There is absolutely no difference between a hard thing and a soft thing so long as they are not brought to the test. (5.403)

What Peirce is describing here is an actual test such that the meaning of hardness would be all those consequences experienced through the course of testing our expectations of something. It seems that he is conflating meaning with the sensible effects borne of some testing procedure.

Peirce further probes the consequences of the maxim by

\(^6\)Ibid.
considering a situation where a diamond is crystallized in the midst of a soft cushion of cotton, and remains there until it is finally burned up. He asks, what would prevent us from saying that the diamond was soft, or from saying that all hard bodies remain perfectly soft until they are touched, "when their hardness increases with the pressure until they are scratched... there would be no falsity in such modes of speech." In saying this, Peirce brings himself dangerously close to a nominalist position, which would have ended up undermining his realist convictions. But realizing his error he embarked on a decades-long endeavor to frame the maxim in the context of the metaphysical, logical, and semiotic components of his architectonic. He renamed it, as we know, pragmaticism, and he offered an example that more clearly illustrates the parameters he intended for the maxim. "The inkstand upon the table is heavy [by which] we only mean, that if its support be removed it will fall to the ground. This may perhaps never happen at all — and yet we say that it is really heavy all the time." (CP 7.341) To refer to the diamond example, it is to say that we do not conceive its beginning to be hard only when we try to scratch it with something. Peirce means to illustrate that when we employ the maxim, meaning is determined not by some actual test, but by a real law manifested by the effects that would result if we were to test it. Employment of the maxim is a thought experiment and turns on its use of counterfactual conditionals, not actual tests, in explicating meaning. The practical effects that we observe are particular events in space and time, what Peirce would call seconds, but the intellectual concepts which correspond to them are not reducible to them. As Peirce says,

Intellectual concepts... essentially carry some implication concerning general behavior... and so convey more, not merely than any feeling, but more, too, than any existent fact, namely, the 'would-acts,' 'would-dos' of habitual behavior; and no agglomeration of actual happenings can ever completely fill the meaning of a 'would-be.' (CP 5.467)

Meaning is thus something general, not particular, and is revealed, in our use of the maxim, not as a function of any particular act but of the habits that guide those acts.
We can return now to the functionalist approach to semantics. We ended our discussion with the historicist problem generated in teleological semantics, but if we are to reject the evolutionary nature of the notion of function this returns us to the problem of objective criteria for the delimitation of content. But we saw that Fodor’s remarks concerning counter-factuality seem to point a way out of the problem, and in fact these considerations form the point of departure for functional-role semantics.

To recall, the function of a state is no longer that which has been historically selected for under evolutionary pressure, but simply its causal role within the system, that is, the characteristic pattern of causes and effects which it normally has within the system. But as Butterill and Carruthers ask, “...precisely which of a state’s normal causes and effects are to be used to individuate it? All of them? And only the actual causes and effects within a particular thinker?” The authors respond in the negative, that functionalism as applied to mental states claims to individuate such states in terms of their potential causal interactions with bodily stimuli. If this were not so, then it would seem to be the case, for example, that we would have to claim that my pain must be a distinct kind of mental state from yours, merely because I happen to have a desire to appear brave whereas you do not. On the contrary, functionalists will insist that our states are the same, provided that they would have the same effects if all our mental states were similar.8

The set of potential connections, rather than actual ones, is what in the end provides a principled way of distinguishing states that avoids the problems associated with informational and teleological semantics as well as averting the problem of observer-relativity and arbitrariness in the determination of content.

Functionalism and Ontology

This is a very brief review of the case these authors make for functional-role semantics, but it affords us a very clear opening to discuss the ontological presuppositions of traditional (naturalized)

---

7Ibid., p. 177.
8Ibid.
cognitive science. In *Mindware*, Andy Clark quite artfully sums up these presuppositions: "Mindfulness is just matter, nicely orchestrated." Peirce, quite to the contrary, affirms: "Matter is effete mind." (CP 6.25) In another context the conjunction of these two statements might seem to constitute a glaring, incommensurable gap between two competing worldviews, realism and idealism. But in the present context it seems rather profoundly suggestive, given the extent to which Peirce’s pragmaticism, articulated in the framework of a synthesis of realism and idealism, foreshadows the very conclusions to which functional-role semantics has come.

Functionalism was never offered as an ontological theory of mind. The ontology of twentieth century cognitive science has always been brain stuff, physical matter (nicely orchestrated, of course). Its introduction was necessary, as we saw, in order to avoid the crude consequences of type-identity theory. If mental states, as functionalism claims, are multiply realizable, the crass identity between mind and brain is no longer a sufficient criterion of distinction. Rather a mental state is to be identified with the effects normally associated with it. Thus, functionalism provides an operational criterion, or maxim, quite similar to Peirce’s pragmatic maxim. Now it would seem that our discussion of semantic theories in this paper, in order to address the question of the ontology of mental states in terms of function, is in some degree separate from a discussion of the nature of mind or mental states as such. Is it not the case that we can ask what a mental state is, and then ask what it means? In a functionalist approach the answer is no because the two are conflated. We are to speak of mental states in terms of their function, and any possible meaning the state could have must be related to that function, to the series of effects consequent upon it. As Peirce said, "[O]ur conception of those effects is the whole of our conception of the object." Relating meaning to anything else, whether brain stuff or Platonic forms, returns us to all the problems functionalism was meant to resolve.

It is for this reason that we esteem the conclusions of functional-role semantics as decisive for a discussion of the ontology of function in theories of mind. The recourse that functionalist semantics takes

---


**BUDHI** 1—2007
to counterfactual conditionals, the insistence on the notion of the potential, or in Peirce's language the "would-be", as characterizing the distinguishing marks of a mental state, problematizes the physicalist ontology at its base. In what way? It seems to us that the very use of the word function, to say that such and such mental state has such and such function, goes beyond the permissible limits of any general empiricist approach. It sounds like something general, metaphysical. In strict terms, a functionalist should speak only of the actual effects which issue from the state, leaving behind any inductive generalizations we might make about those effects and to which we would give the name 'function,' generalizations which in a Humean empiricist sense would be strictly impermissible.

But we cannot do this, according to functional-role semantics, without falling back into the very problems it hoped to resolve. The conclusion would seem to be that functionalism can be a viable theory of mind only if it expands its ontological commitment such as to accord to mental states a reality not wholly reducible to some particular phenomenon, whether individual neurons or isolated effects. If we are to speak of mental states in terms of function, then function cannot be reduced to a series of actual effects (what Peirce would understand in terms of secondness) but rather is of the nature of a general (Peirce's thirdness). A general is something that governs individual events in the future.

Although Peirce affirms quite clearly that matter is effete mind, this does not mean that he considers matter as some sort of epiphenomenal product of thought, as an extreme idealism would hold. Peirce is certainly a realist, but a realist of a particular "stripe." He writes:

[I]t follows that since no cognition of ours is absolutely determinate, generals must have a real existence. Now this scholastic realism is usually set down as a belief in metaphysical fictions. But, in fact, a realist is simply one who knows no more recondite reality than that which is represented in a true representation. Since, therefore, the word 'man' is true of something, that which 'man' means is real. The nominalist must admit that man is truly applicable to something; but he believes that there is beneath this a thing in itself, an incognizable reality. His is
the metaphysical figment. (CP 5.311-312)

Peirce quite deftly turns the tables here on the nominalist and reveals the pernicious metaphysics lurking in his worldview. But as we see from the quote, this does not commit Peirce to some ether-breathing Platonism. Reality is simply that which is 'represented in a true representation.' It is opposed, not to the non-existent (nominalism), but rather to the fictitious. Things are either real or fictitious, not in themselves, but only in relation to other things. The logic of relations that informs his ontological view implies hierarchies of emergent order and being that dispels the easy dichotomies of modern Cartesian philosophy. It is not a question of either/or, either material brain stuff or spiritual mental stuff. What is needed, and what Peirce's thought in my opinion brings quite clearly to the discussion, is a metaphysical turn away from the substance philosophy, that still lurks among us, to a sort of "functionally differentiated monism" articulated in semiotic terms. The basis for this differentiation, as for all of Peirce's philosophy, is his categorial structure of Firstness, Secondness, and Thirdness. Very briefly, philosophical categories attempt to account, ontologically and epistemologically, for all phenomena in the universe. Peirce's categories are fundamentally relational in nature. Secondness is the category of existence, two things in a brute dyadic relationship with each other. In our context Secondness would correspond to the empirical physicalism of traditional cognitive science. Thirdness is the category of mediation, where one thing mediates between two others thus bringing about a law-like regularity. Here, Thirdness would correspond to the non-physical mentalism characteristic of traditional Cartesian-inspired philosophy of mind. The missing link, or category, in the cognitive sciences, that which could organically unite these two camps, is Peirce's firstness, to which would correspond not the mental nor the empirical but rather the affective. Study of the latter constitutes as of late a burgeoning field of interest in many disciplines. 10 Hopefully a philosophical synthesis of these diverse studies will be able to offer

in the near future a conceptual framework capable resolving the difficulties I have tried to articulate in this paper.11

7A colleague and I are currently working on a book with the provisional title, Sentient Semiotics: Steps toward the Incarnation of the Ineffable, in which we attempt to trace the general outlines of such a synthesis.