

## KRIPKE ON FUNCTIONALISM\*

### KRIPKE SOBRE EL FUNCIONALISMO

**Abstract** Saul Kripke's still largely unpublished views on functionalism in the philosophy of mind are expounded on the basis of a transcript of a 1984 lecture of his on the topic, and some unresolved questions identified.

**Resumen** Las opiniones de Saul Kripke acerca del funcionalismo en la filosofía de la mente, que aún permanecen en gran parte sin publicarse, son expuestas con base en una transcripción de una plática suya de 1984 sobre este tema, y algunas preguntas sin resolver son identificadas

**Keywords** *finite automata, idealization, phenomenal character, teleology, Turing machines*

**Palabras clave** *autómatas finitos, carácter fenomenal, idealización, máquinas de Turing, teleología*

For half a century functionalism has been a prominent, and for much of that period the dominant, position on the mind-body problem among philosophers, not to mention cognitive scientists. The view has also been subject to many criticisms, but one of the most telling, the critique due to Saul Kripke — addressed to what are generally considered functionalism's strong points rather than to its weak points — has remained largely unpublished and has not played a conspicuous role in debate so far. What follows here will be a concise account of Kripke's view, followed by some discussion of two questions it leaves open.

#### The Ambiguity of *Function*

While French has the fabled *Académie Française* and Spanish the *Asociación de Academias de la Lengua Española*, English has no formal highest linguistic authority. Yet there does exist an informal highest authority, a very British institution whose preeminence even North American academics must recognize, in the form of the *Oxford English Dictionary* (OED). This work has earned its authoritative reputation through the thoroughness of its documentation of usage and the incisiveness of its definitions.

In approaching the subject of functionalism in the philosophy of mind and Saul Kripke's critique thereof, it will be well to turn first to this authority and look at its definition of *function*. For the definition turns out to have a disjunctive form, exposing a crucial ambiguity in the word, an ambiguity highly relevant to Kripke's criticisms. The OED definition of *function* reads as follows: *an activity or purpose natural to or intended for a person or thing*. The ambiguity it makes evident is between the non-teleological notion of *function as activity* and the teleological notion of *function as purpose*.

How these differ will come out if we ask: *Is it the function of the human heart to pump blood around the body?* That is certainly the heart's natural activity, but is it the heart's intended purpose? It is indeed the intended purpose of the Jarvik artificial heart, a human invention designed by its creator precisely to circulate blood around the body. But what of the heart we are born with? If one thinks in the traditional way of the human body as being designed by a god, then though one can hardly claim to know, one can still at least plausibly conjecture that one of the creator's intended purposes for that organ, perhaps among others more mysterious, was to pump blood. But what if the heart is merely a product of evolution by natural selection? Then, obfuscatory metaphors aside, there is no agent, no design, no intention, no purpose. It is merely that animals that have some efficient means of circulating fluids around their bodies tend to survive and reproduce better than those that do not. Aristotelian-Scholastic teleology, banished from physics in the seventeenth century, seems with Darwin to be belatedly banished from biology as well, two hundred years or so later. Darwinians can say that moving blood around the body is what the heart *does do*, but not that it is what the heart is *for doing*.

Now one cannot read far in the typical elementary exposition of functionalism in philosophy of mind without finding the expositor, though supposedly discussing natural activities of the human mind or brain, turning to analogies with artifacts designed by agents with intended purposes for them: most often computers, but sometimes pumps, and in one case even mousetraps (Anderson 2003). Such a step from function in the sense of *activity* to function in the sense of *purpose* presents a danger of leading us into fallacies of equivocation. Kripke's remarks on functionalism may be viewed as one long

commentary on this observation.

### Sources for Kripke's Views

Kripke's thoughts on functionalism have so far been generally available only in the form of one long footnote, about a page and a half of fine print, in his Wittgenstein book (Kripke 1982 pp.35-37). Now controversial questions of Wittgenstein exegesis are the last thing I want to get involved in here — or anywhere, for that matter — but fortunately the reason the material on functionalism is relegated to a footnote in the Wittgenstein book is that it is recognized to be rather tangential to the general theme of that work, and therewith to exegetical controversies. The key footnote does begin by citing a work of Wittgenstein, but the work cited is not the text under discussion in the bulk of the book, the *Philosophical Investigations*; it is, rather, the *Remarks on Foundations of Mathematics*, and specifically certain passages on ideal versus real machines. About halfway through the long footnote the word *functionalism* first appears. It does so in what traditional rhetoric calls an *apophasis*, a rhetorical device consisting in mentioning something by saying that one is not going to mention it. Kripke's suggestion is that, though he is not going to discuss functionalism, what he has just been saying about machines up to that point should apply *mutatis mutandis* to functionalism. The note ends with an expression of a hope to elaborate elsewhere.

Elaborations did occur within a couple of years, in public lectures in the spring and summer of 1984, at Duke University and then at a Wittgenstein conference in Kirchberg-am-Wechsel, Austria. An audiotape transcript of the latter lecture is being edited at the Saul Kripke Center and the City University of New York for eventual publication, and I will be using it as my main source for the views of Kripke that I will be expounding and examining. One account of Kripke's views using this same lecture material has already been published as (Buechner 2011). The Austrian lecture is an easier source to use than the book, since Kripke emphasizes that in it, unlike the Wittgenstein book, he is speaking for himself, expounding views he himself endorses. A few brief references to Wittgenstein are inevitable

given the venue of the talk, but they do not take us far into any contentious issues.

Kripke offers no precise characterization of what he understands *functionalism* to be. He describes it as the dominant position in philosophy of mind in his own country at the time of his talk, but it is when a creed becomes dominant that it is most likely to experience fission into sects that need to be distinguished. Kripke's omission of anything much in the way of characterizing functionalism and distinguishing varieties may be disappointing, but it is not surprising, given his complaint that up to the time at which he was speaking professed functionalists, apart from David Lewis, had been unclear about many important aspects of their various versions of the doctrine.

### Machine-State Functionalism

It is now customary for accounts of functionalism in standard references, such as (Levin 2013), to distinguish three main varieties: *machine-state functionalism* and two forms of *causal-role functionalism*: *aposteriori psycho-functionalism* and *apriori analytic functionalism* (for the *cognoscenti*, in its *role functionalist* rather than *realizer functionalist* version). Kripke's remarks appear to be intended to apply quite broadly, presumably therefore to all three versions of the doctrine, but let me begin with the machine-state view, which is the version he addresses most directly and explicitly, and which Buechner, too, emphasizes, as is clear from the very title of his account.

Machine-state or computational functionalism traces back to work of Hilary Putnam beginning with (Putnam 1967), though as Kripke mentions, by the time of the Austrian lecture Putnam had abandoned the view. Indeed, he eventually propounded several arguments against it: (Buechner 2007) counts four, none coinciding with Kripke's argument, though the Putnamian argument from *trivialization*, claiming that any computational description can be attributed to any physical system, has points of contact. So does a sequel to Searle's notorious *Chinese Room* or *syntax does not determine semantics* argument, namely, his *physics does not determine syntax* argument in (Searle 1990). But comparisons of the anti-functionalisms of Searle, Putnam, and Kripke are best postponed until more of

Kripke's work is out.

To return to the Putnam of the period when he was still a functionalist, the idea behind his machine-state functionalism is by now very well known, and in any case very simple. Consider a piece of software, a computer program, say for definiteness an algorithm for addition. To execute the program one uses a piece of hardware, a physical computing machine, but the same program may be implemented on different kinds of machines, for instance an old-fashioned mechanical adder, or a new-fangled electronic adder. The same step in the program, say carrying one at a certain point, will be executed in different ways on the different machines, say by turning a gear in the one case, and by transmitting an electric pulse in the other. In consequence, there are two different ways in which or levels at which one may describe what the machine is doing at a given stage. The two may be called the *functional* way and *physical* way. One may describe the machine functionally as carrying one, or one may describe it physically as doing something mechanical or electrical, turning a gear or transmitting an impulse, depending on the kind of device at work.

An account in functional terms will generally be more perspicuous, though it is inevitably incomplete in comparison to a physical account, by leaving out details about *how* the machine does what it does. Indeed, it achieves perspicuity through this very incompleteness, by omission of what are arguably irrelevant details about implementation. Nonetheless, the most satisfying account may consist of the functional account together with information, for each basic kind of step in the program, as to how that kind of step is physically realized in the given machine. From such materials a step-by-step physical account of the whole computation presumably could be deduced if one were wanted. But what matters here about all this is that the functional process, *carrying one*, clearly cannot be identified with any particular physical process such as *gear-turning* or *impulse-transmitting*. Still less is it a *spiritual* process taking place in some kind of ectoplasm or mental substance. The idea behind machine-state functionalism is that if one takes mental processes and states to be functional processes and states, like software processes and states, then one can avoid both physicalism or materialism and equally idealism

or dualism, in the philosophy of mind. That such avoidance is desirable seems to have been more or less taken for granted by Putnam and others.

And so we are led to an idea that became immensely popular, that of comparing the brain to a computer. Since the brain is hooked up on the one side to eyes and ears and so forth and on the other side to muscles, we presumably should think of the computer involved as one that is hooked up on one side to cameras and microphones and so forth and on the other side to motors. We are then thinking of a computer that is the central processing unit of a robot. About this idea Kripke has makes three kinds of remarks, which I will take in order of increasing importance.

### The Limitations of Finite Automata

Let us focus on one type of mental activity in which we sometimes engage: doing multiplication problems in our heads, without paper and pencil. Here in each instance we are given a sensory input or stimulus, say hearing the question, *What is thirteen times fourteen?* And in each instance we give a motor output or response, say moving the tongue and vocal cords to utter the answer *One hundred eighty-two*. The mental process going on inbetween the publicly observable stimulus and response is what I want to focus on. According to machine-state functionalism, this type of process should correspond to, or rather, should be constituted by the execution of, a type of subroutine in an overall program being run on the brain. The difficulty for functionalism comes in trying to pin down what type of subroutine or subprogram this could be. What could make a subprogram to be such that executing *it* would constitute the type of mental activity that we call *multiplying*?

A first difficulty arises if we try to answer this question by saying, *What makes this type of subprogram multiplying is just that when it is applied to any two inputs, it gives their product*. About this reply Kripke observes that a realistic comparison of our brains to computers would have to be a comparison to computers like so-called *finite automata*, with a bounded storage capacity. But in that case, no subprogram is such that executing it will give the product in all cases. Kripke's lecture is

entitled *Lessons on Functionalism and Automata*, and he devotes considerable time to beginning lessons in the theory of finite automata, explaining how such automata can add, and explaining why such automata cannot multiply.

### The Imperfection of Physical Realizations

A second difficulty arises if we try to avoid the difficulty just described by taking as our model, not finite automata, but Turing machines, which have potentially infinite capacity. About this reply, Kripke emphasizes that we must take note of the distinction with which machine-state functionalists begin, between abstract software and concrete hardware for executing or implementing it. What is conventionally but misleadingly called a Turing *machine*, he remarks, really should be called a Turing *program*. It is not a piece of hardware, a mechanical contraption or electronic gadget, but something abstract that can be concretely executed or implemented by such hardware — and according to the machine-state functionalist also by the so-called *wetware* of the brain. The difficulty here is that all physical realizations of a Turing program are imperfect. And if the brain is regarded as such a realization it certainly must be so, precisely because the Turing machine has a potentially unlimited storage capacity and our brains do not.

Any concrete realization of an abstract ideal machine will occasionally malfunction temporarily, and eventually break down permanently. This means that whatever algorithm for multiplication may be included as a subroutine in the Turing program, implementation thereof is *not* going to give the product as output for *every* pair of numbers that may be given as input. Occasionally there will be a malfunction and an output given other than the genuine product; eventually there will be a breakdown and thereafter no more outputs at all. Besides the multiplication function there will be other computable functions that produce the same outputs for all the inputs that anyone ever has encountered or ever will encounter, but differ for other inputs. If *multiplication* is one such function — for definiteness let us say the one that gives the product for inputs less than  $N$  and the output zero otherwise, where  $N$  is some enormous

number, some power of 10 safely greater than the number of elementary particles in the visible universe — we may ask: *What is it that makes what I am doing to be an implementation of an algorithm for multiplication rather than multiplication?* One can ask the same question about a mechanical contraption or electronic gadget: What makes what it is doing when it is doing what we call *multiplying* to be an implementation of an algorithm for multiplication rather than for some other function?

### The Relativity of Counterfactual Idealization

A third difficulty arises if we try to answer this question by appealing to idealization, to counterfactuals about what would happen under ideal conditions, by saying, *The product of the inputs is what I or the contraption or gadget would give as output if there were no malfunction or breakdown.* About this response Kripke in effect observes — and this is the single most crucial point in his whole discussion — that though the brain, or the mechanical or electronic device, is a physical object, the one-place predicate “*x* is malfunctioning” does not express a physical property of such an object. For the physical object *x* in question is malfunctioning if and only if it is not doing what it is supposed to be doing. This makes malfunctioning a teleological notion. And hence it is not a physical notion, since teleology was excluded from physics at the time of scientific revolution four hundred years ago, when one ceased saying things like *Heavy objects fall towards the center of the earth because they are trying to get to where they are supposed to be.*

Really what one has to do with here is a *two*-place predicate, “*x* is malfunctioning if considered as an implementation of Turing program *y*.” Or rather, since there are different manners in which a given physical object might be construed as an implementation a given Turing machine, we would need a third variable *z* for manners of construal; but let me for simplicity pretend that two will do. The two-place predicate, by taking various specific Turing programs *T*, *T'*, and so on for *y*, gives us a series of one-place predicates “*x* is malfunctioning if considered as an implementation of Turing program *T*”, “*x* is malfunctioning if considered as an implementation of Turing program *T'*”, and so on. These arguably

*do* express physical properties, but that is of no help until we can answer the question: Of which program, T or T' or whatever, is *x* to be considered an implementation? Here T might be an abstract algorithm for *multiplying* and T' an abstract algorithm for *nultipliyng*. You can't ask a machine what it is doing, multiplying or nultipliyng. You can ask me, and if you do, I will of course say that what I am doing is what I call *multiplying*. But what makes it to be the case that when I say *multiplying* I really mean multiplication and not nultipliyng? That is the form in which Kripke confronts this kind of difficulty in the Wittgenstein book.

We may fail to notice the fatal relativity here because ordinarily when we are speaking about a computing machine we are not speaking of some physical object that, as Kripke puts it, *just fell from the sky*. Rather, we are speaking of a contraption or gadget that was designed by human agents with definite intentions as to what algorithms they are supposed to be implementing. Design by an agent cannot be assumed for a brain, however. Obviously the human brain was not designed by some human agent, some Doctor Jarvik or Doctor Frankenstein, since such an agent needs to have a brain already to design anything with. As Kripke notes, functionalists generally are the sort of people who would not want to assume that the brain is the invention of some god. And anyhow, just as for Descartes what mattered was not whether there is a god, but whether there is a god who is not a deceiver, so for us what matters is not whether the brain was devised by a god, but whether it was devised by a god who has revealed to us his intentions. In the absence of any such a revelation — I haven't had one, and I expect the reader hasn't, either — the brain for us is a piece of meat which may be regarded from different points of view as an implementation of various different kinds of programs. As to which if any it was intended to implement, all we can say is, as the expression goes, *God only knows*.

### Lewis and Analytic Functionalism

Let me take the issue of machine-state functionalism no further than this, referring the reader to Buechner for further discussion, and turn to consider causal-role functionalism. Kripke says that one

seldom finds any version of this doctrine spelled out in detail, though Lewis takes things further than anyone else. I would add here a mention of quite recent discussion of Lewis and functionalism by Wolfgang Schwarz. For (Schwarz 2015) gives all in one place the theory that developed in stages through various works of Lewis over the years. I will consider here only the variety of causal role functionalism that has been most thoroughly spelled out, this Ludovician analytic functionalism as per Schwarz.

The underlying idea of the view is that we all implicitly subscribe to some kind of psychological theory. The theory is supposed to be about two kinds of things: on the one hand, there are external stimuli and external responses, presumably describable in terms of things in principle objectively observable by third parties; on the other hand, there are internal mental states. The theory consists of a lot of generalizations, typically guarded by words like *typically*, about various relations, typically various cause-effect relations, of internal mental states to stimuli and responses, and to each other. Nothing, however, is said about the intrinsic nature of these internal mental states, but only about their causal, or as is said, *functional* relations to each other and external, objectively observable matters.

One of the mental states mentioned would presumably be that of *being in pain*. About this it might be said, among many other things, that cutting the body tends to cause one to come to be in this state; that being in this state tends to cause a desire to get out of it, which in turn causes various forms of care-seeking behavior. Here cutting the body is an example of a causal relation of an objectively observable stimulus to an internal mental state, while being in pain producing a desire of a certain kind is an example of a causal relation of one internal mental state to another, and the desire causing care-seeking behavior is an example of a causal relation of an internal mental state to an objectively observable response.

And the functionalist view is that subject  $x$ 's being in pain at time  $t$ , for instance, just is a matter of two things being the case: First, that there is some unique collection of states that stand in the network of causal relations to each other and to stimuli and responses that the theory attributes to the various

internal mental states. Second, that at time  $t$  subject  $x$  is in the state in that collection that plays the role the theory attributes to the state it calls that of *being in pain*. Subtle refinements of doctrine are needed to deal with the case where there is no unique collection of states of the kind called for by the theory, but I must pass over, referring the reader to Schwarz. What I have said so far applies to both types of causal role theory. Lewis's analytic functionalism contrasts with psycho-functionalism in that the various assertions in the theory should, for Lewis, involve no concepts not generally familiar to the lay public, and make no assertions not generally accepted by the lay public as something like platitudes. That is what Lewis means when he says that the theory belongs to *folk psychology* rather than *scientific psychology*.

On this Ludovician view, as Schwarz makes clear, folk psychology contains an analytic and a synthetic component. What is supposed to be synthetic is the assumption that there is a unique system of states playing the roles attributed by folk psychology to the various internal mental states with which it is concerned. What is supposed to be analytic is the conditional assertion that if there is indeed some such system of states, then the state of being in pain is — and the phrase *being in pain* refers to — the state in the collection that plays the role the theory attributes to the state it calls *being in pain*. Thus *being in pain* in effect just means *being in a state the plays the following role...* followed by the specification of the role folk psychology attributes to being in pain.

#### Non-Kripkean Criticism of Analytic Functionalism: The Case of Pain

Kripke says this sort of view is utterly implausible in application to sensory states like being in pain. He compares trying to get a sensory state out of a functional role to trying to get water out of a stone. He does not elaborate, considering the point familiar — he calls it the *most popular* objection to functionalism — but let me elaborate a bit here as if on his behalf. First note that just as some people are born blind and are unable to see and some people are born deaf and are unable to hear, so some people are born with a condition called *congenital analgesia*, and are unable to feel pain. This is a dangerous

condition, especially for children, as one can well imagine. One might doubt whether anyone could survive it, but there is a cluster of cases in one small village in the far north of Sweden, and isolated cases are sporadically reported elsewhere, as a web search on the search phrase *congenital analgesia* will soon reveal. What does the occurrence of this rare condition have to do with our present concerns? Just this: The analytic functionalist's account of the meaning of *being in pain* seems plausible as an account of all that an individual incapable of feeling pain could mean by the phrase. Congenital analgesiacs, who do not react to being cut unless they happen to notice blood dripping, may say of pain: *It is some internal state which in people not like us is caused by being cut...* and so on. For them, being in pain is some kind of theoretical posit needed to explain the observable behavior of other people.

But the analytic functionalist's account of the meaning of *being in pain*, offered as an account of what ordinary people, in contrast to congenital analgesiacs, mean by the phrase, is wholly and wildly implausible, to the point of seeming like a bad joke. Surely we mean a feeling, with which we are acquainted from painful experience, and whose name we know by having been told it in childhood — long before we ever learned the whole of so-called *folk psychology*, unless we are being asked to believe that this whole elaborate theory is innate — while experiencing the state. The functionalist account leaves out the most important element of the meaning of being in pain, which we all learn when quite young, namely, that it feels like *this* [where in oral presentation I would at this point slap my face by way of demonstration].

#### Non-Kripkean Criticism of Analytic Functionalism: The Case of Belief

No subtle argument of indirect Wittgensteinian inspiration, concerned with issues in automata theory, seems needed before we dismiss analytic functionalism as an account of what it means to speak of having a sensation such as pain. The use Kripke sees for his argument in connection with analytic functionalism would be in the case of its account of internal mental states that, unlike sensations such as pain, introspection seems to show to be lacking anything much in the way of a so-called *phenomenal*

*character*. Nowadays many advocates of analytic functionalism anyhow tend to emphasize such examples, rather than sensations like pain; which is to say that nowadays many are only partial advocates of analytic functionalism, as an account plausible for a certain range of internal mental states but not others. This seems to be the view of Schwarz, for instance, sympathetic though he is to Lewis.

The types of mental states Kripke mentions in these connections, his examples of states without much in the way of phenomenal character are *belief* and *meaning*. Kripke's main example of meaning, carried over from the Wittgenstein book, is meaning addition by *plus* and multiplication by *times*. Following Kripke's practice in *Naming and Necessity* of using controversial political figures (such as Richard Nixon) as examples, let me take as my main example of belief the belief that Donald Trump will not be the next president of the U.S.A. [At the time of this writing, which happens to coincide with Trump's visit to Mexico, this can only be called a *belief*; but the hope is that before this essay sees print it will be possible to call it *knowledge*.]

Now it certainly cannot be unambiguously claimed that there is something it feels like to believe that Donald Trump will not be voted into the office for which he is standing as candidate. But there does seem to be something it feels like to be, as the expression goes, *saying to oneself* — that's an idiomatic expression, meaning *thinking to oneself while not saying aloud* — that the Donald will not win the upcoming election. There must be something that it is like to be Bill Clinton or Chris Christie saying to himself, *Trump is not going to be moving into the White House in January*. Clinton may be saying this to himself with delight, Christie perhaps with disappointment, or more likely with secret relief; but despite the difference in accompanying affect there is something in common to their experiences, and to the experiences of neutral observers who say the same to themselves.

In jargon, I am claiming that what is called *occurrent belief* has a phenomenal character, and so cannot be accounted for functionalistically, since an account of its functional relations will be leaving something out, how it feels to be in the state in question. But, of course, not all belief is occurrent belief, with a feeling to it. We may attribute to either of the Clintons a belief that they will be moving back into

the White House even when Hillary and Bill are both deep in dreamless sleep and feeling nothing. This is a case of what is called *dispositional* belief, because in such a case there typically is at least a disposition to have an occurrent belief, when one is awake and the pertinent question is raised, even if there is no occurrent belief right this moment.

I have said that a functionalist account, mentioning only causal relations among internal mental states and between them and external, objectively observable stimuli and responses, involving no mention of phenomenal character, cannot be adequate as an account of occurrent belief. But there seems to be enough of a causal or so-called *functional* relation between occurrent belief and dispositional belief, that a theory that is inadequate as an account of the former must *ipso facto* be inadequate as an account of the latter as well. A crucial part of its characterization of the internal mental state it calls *dispositional belief* will be that it has certain functional relations to another internal mental state that it calls *occurrent belief*; but the account of the latter will be insufficient, by leaving out its phenomenal character, and this will make the account of the former inadequate, too.

Kripke in his Wittgenstein book argues from introspection that meaning does not itself have a phenomenal character. But he does seem to allow that it has some relation to internal states that do have a phenomenal character. At any rate, that is what I would claim. For instance, it is surely part of folk psychology that when people are asked to fill out a long bureaucratic form — a customs declaration, let us say — in a language whose meaning they do not know, they tend to become puzzled or perplexed, and perhaps even distressed, in some way. Here is a connection between meaning, a state that I am granting does not itself have a phenomenal character, and distress, an affectual or emotional state that most assuredly does have one. But if there is such a connection, then an account solely in terms of external stimuli and external responses, plus causal or functional relations, with no mention of phenomenal character, cannot be any more adequate for meaning than for distress. Kripke emphasizes that functionalism is supposed to be a so-called *holistic* view, according to which one gives an account of any one internal mental state by and only by giving an account of every other internal mental state as

well, and indeed Kripke cites this holism as another implausible feature of the view. But what is important for me at this point is that holism implies that, on their own showing, if functionalist accounts are unsatisfactory as accounts of one class of internal mental states, those with an introspectable phenomenal character, then they must be unsatisfactory as accounts of all internal mental states, including such as are without any introspectable phenomenal character.

## Two Questions

As Levin's survey makes clear, there are other important criticisms of analytic functionalism. As for the other variety of functionalism, which Kripke does not discuss at all, and whose existence he barely acknowledges, one can say the following, that just as Putnam, the founder of computational functionalism, became one of its most important critics, so did Ned Block, a founder of psycho-functionalism, come, even before Kripke's book or lecture, to abandon and criticize the view in (Block 1978). But then, Kripke never claimed or believed that his criticism was the only one available, or the strongest, for either form of causal-role functionalism. Still, according to Kripke the various forms of functionalism have enough points of similarity that his putative refutation of machine-state functionalism should carry over *mutatis mutandis* to analytic functionalism. I cannot, however, myself see how the analogy is supposed to go, and so would ask: *Just how are considerations about the gap between abstract ideals and concrete implementations supposed to bear against causal-role functionalism?*

A second question is simply *How would Kripke respond to a defense of functionalism based on the Platonic notion of carving reality at the joints as developed in work of Lewis and followers, mainly in the period after Kripke's lecture?* This line of thought has been developed by Ludovicians partly in response to the skeptical problem of Kripke's Wittgenstein book. Kripke briefly discusses the notion that I mean multiplication and not nultiplication by *multiplication* because multiplication rather than nultiplication is the simplest function approximated by the answers I give to what I call *multiplication*

problems, as far as these go. He remarks, among other things, that the relevant measure of simplicity has not been explicitly worked out, and that there is a danger of circularity in appeal to any human measure of simplicity, since that would itself require interpretation that might be subject to skeptical problems. But the Ludovician idea does not involve human measures of simplicity, but rather distinctions in the inhuman universe as it is in itself, independently of and indifferent to us and our concerns, distinctions in nature that for want of a better term we may call *natural*. Lewis's conception is compared in (Hawthorne 2006, p.viii) to *veins in the marble of reality*. Another comparison might be to inherent planes of potential cleavage in a crystal.

The anti-skeptical thought would then be that though our feeble tapping on a small portion of the surface of this crystal can never trace a distinction across its whole infinite extent, or even trace one over a small finite portion in a completely reliably reproducible way, nonetheless, like the tapping of a diamond-cutter's hammer, what we do suffices to make the crystal split at one of its planes of cleavage. The distinction that cleavage marks thereby becomes the one that we may be said to have been marking with our tapping, and it provides the criterion for the rightness or wrongness of any further taps. No fact purely about us determines rightness and wrongness here, but certain facts about us together with certain facts about ultimate metaphysical reality independent of us. I myself can only respond to such suggestions with the *incredulous stare* that Lewis used to mention as something his metaphysical speculations would often encounter. Can Kripke give a more substantive response?

Perhaps answers to such questions as the two I have sketched are to be found in materials not yet scheduled for release from the vaults of the Saul Kripke Center. In the meantime, it is to be hoped that the projected publication of the Austrian lecture will occur speedily, and that the work, though appearing very belatedly, will be given some of the attention it deserves.

## REFERENCES

- Anderson, David L., 2003, "Introduction to Functionalism,"  
«[www.mind.ilstu.edu/curriculum/functionality\\_intro/functionality\\_intro.php](http://www.mind.ilstu.edu/curriculum/functionality_intro/functionality_intro.php)»
- Block, Ned, 1978, "Troubles with Functionalism," *Minnesota Studies in the Philosophy of Science* 9:261-325.
- Buechner, Jeff, 2007, *Gödel, Putnam, and Functionalism: A New Reading of "Representation and Reality,"* MIT Press.
- Buechner, Jeff, 2011, "Not Even Computing Machines Can Follow Rules: Kripke's Critique of Functionalism", in Berger, Alan (ed.), *Saul Kripke*, Cambridge University Press, 343-368.
- Hawthorne, John, 2006, *Metaphysical Essays*, Clarendon Press.
- Kripke, Saul, 1982, *Wittgenstein on Rules and Private Language: An Elementary Exposition*, Harvard University Press.
- Levin, Janet, 2013, "Functionalism," *Stanford Encyclopedia of Philosophy*, «[plato.stanford.edu/archives/fall2013/entries/functionality/](http://plato.stanford.edu/archives/fall2013/entries/functionality/)»
- Putnam, Hilary, 1967, "Psychological Predicates," (retitled "On the Nature of Mental States" in its many later reprintings), in Capitan, W. H. and Merrill, D. D. (eds.), *Art, Mind, and Religion*, Pittsburg University Press, 37-48.
- Schwarz, Wolfgang, 2015, "Analytic Functionalism", in Loewer, Barry and Schaffer, Jonathan, *A Companion to David Lewis*, Wiley, 504-518.
- Searle, John R., 1990, "Is the Brain a Digital Computer?" *Proceedings and Addresses of the American Philosophical Association* 64:21-37.

## NOTE

\* The present paper is a slightly revised version of a talk, “Kripke on the Follies of Functionalism,” presented at the event “Colors, Sounds, Numbers, Names, Inferences, Minds: A Conference on Topics in Contemporary Philosophy,” held at the Instituto de Investigaciones Filosóficas of the Universidad Nacional Autónoma de México, 29-30 August, 2016. I am grateful to Saul Kripke for speaking as commentator and to Margarita Valdés for serving as moderator at the session at which I presented the paper. I am especially grateful to Mario Gómez-Torrente for the invitation to speak on Kripke’s anti-functional views, and to him and the other organizers, and more generally the staff and students at U.N.A.M., for making the whole conference an enlightening and enjoyable event.

.John P. Burgess  
Department of Philosophy  
Princeton University  
Princeton, NJ 08544-1006, US  
jburgess@princeton.edu