BRIAN CANTWELL SMITH

# THE SEMANTICS OF CLOCKS

*The inexorable ticking of the clock may have had more to do with the weakening of God's supremacy than all the treatises produced by the philosophers of the Enlightenment . . . . Perhaps Moses should have included another Commandment: Thou shalt not make mechanical representations of time.*

— Neil Postman [1985, pp. 11—12]

## 1. INTRODUCTION

Clocks?

Yes, because they participate in their subject matter, and participation — at least so I will argue — is an important semantical phenomenon.

To start with, clocks are about time; they represent it.[1] Not only that, clocks themselves are temporal, as anyone knows who, wondering whether a watch is still working, has paused for a second or two, to see whether the second hand moves. In some sense everything is temporal, from the price of gold to the most passive rock, manifesting such properties as fluctuating wildly or being inert. But the temporal nature of clocks is essential to their semantic interpretation, more than for other time representations, such as calendars. The point is just the obvious one. As time goes by, we require a certain strict coördination. The time that a clock represents, at any given moment, is supposed to be the time that it is, at that moment. A clock should indicate 12 o'clock just in case it *is* 12 o'clock.

But that's not all. The time that a clock represents, at a given moment, is also a function of that moment, the very moment it is meant to represent. I.e., suppose that a clock does indicate 12 o'clock at noon. The time that it indicates a moment later will differ by an amount that is not only proportional to, but also dependent on, the intervening passage of time. It doesn't take God or angels to keep the clock coördinated; it does it on its own. This is where participation takes hold.



*James H. Fetzer (ed.), Aspects of Artificial Intelligence, 3—31.*

As well as representing the current time, clocks have to identify its "location" in the complex but familiar cycle of hours, minutes, etc. They have to measure it, that is, in terms of a predetermined set of temporal units, and they measure it by participating in it. And yet the connection between their participation and their content isn't absolute — clocks, after all, can be wrong. How it is that clocks can participate and still be wrong is something we will have to explain.

For clocks, participation involves being dynamic: constantly changing state, in virtue of internal temporal properties, in order to maintain the right semantic stance. This dynamic aspect is a substantial, additional, constraint. A passive disk inscribed with 'NOW' would have both temporal properties mentioned above (being about time, and having the time of interpretation relevant to content) and would even maintain perfect coördination. A rendering of this word in blinking lights, mounted on an chrome pedestal, might even deserve a place on California's Venice Boardwalk. But even though it would be the first time piece in history to be absolutely accurate, such a contraption wouldn't count as a genuine chronometer.

We humans participate in the subject matter of our thoughts, too, when we think about where to look for our glasses, notice that we're repeating ourselves, or pause to ask why a conversant is reacting strangely. Why? What is this participation? It's hard to say exactly, especially because we can't get outside it, but a sidelong glance suggests a thick and constant interaction between the contents of our thoughts, on the one hand, and both prior and subsequent non-representational activity, on the other, such as walking around, shutting up, or pouring a drink.

Take the glasses example. Suppose, after first noticing their absence, I get up and look on my dresser, asking myself "Are they here?" My asking the question will be a consequence of my wonder, but so will my (non-representational) standing in front of the dresser. Furthermore, the two are related; the word 'here' will depend for its interpretation on where I am standing. And who knows, to drive the example backwards in time, what caused the initial wonder — eye strain, perhaps, or maybe an explicit comment. The point is that the representational and non-representational states of participatory systems are inexorably intertwined — they even rest on the same physical substrate. We can put it even more strongly: the physical states that realise our thoughts are caused by non-representational conditions, and engender non-representational consequences, in ways that must be coördinated with the contents of the very representational states they realise. Participation is something like that.

AI and general computational systems also participate — more and more, in fact, as they emerge from the laboratory and take up residence with us in life itself: landing airplanes, teaching children, launching nuclear weapons. Far from being abstract, computers are part of the world, use energy, affect the social fabric. This participation makes them quite a lot like us, quite unlike the abstract mathematical expression types on which familiar semantical techniques have been developed.

My real reason for studying clocks, therefore, can be spelled out as follows. First, issues of semantics, and of the relationship between semantics and mechanism, are crucial for AI and cognitive science (this much I take for granted). Second, it is terrifically important to recognise that computational systems participate in the world along with us. That's why they're useful. Third, as I hope this paper will show, participation has major consequences for semantical analyses: it forces us to develop new notions and new vocabulary in terms of which to understand interpretation and behaviour. Clocks are an extremely simple case, with very modest participation. Nonetheless, their simplicity makes them a good foil in terms of which to start the new development.

So they're really not such an unlikely subject matter, after all.

## 2. INFERENCE AND TIME-KEEPING

Let's start by reviewing the current state of the semantical art. Consider a familiar, paradigmatic case: a theorem-prover built according to the dictates of traditional mathematical logic. As suggested in Figure 1, two relatively independent aspects will be coördinated in such a system. First, there is activity or behaviour — what the system does — indicated as $\Psi$ (for psychology). All systems, from car engines to biological mechanisms of photosynthesis, of course do something; what distinguishes theorem provers is the fact that their $\Psi$ implements (some subset of) the proof-theoretic inference relation ($\vdash$). Second, there is the denotation or interpretation relation, indicated as $\Phi$ (for philosophy), which maps sentences or formulae onto model-theoretic structures of some sort, in terms of which the truth-values of the formulae are determined. In a computer system designed to prove theorems in abstract algebra, for example, the interpretation function would map states of the machine (or states of its language) onto groups, rings, or numbers — the subject matter of the algebraic axioms.

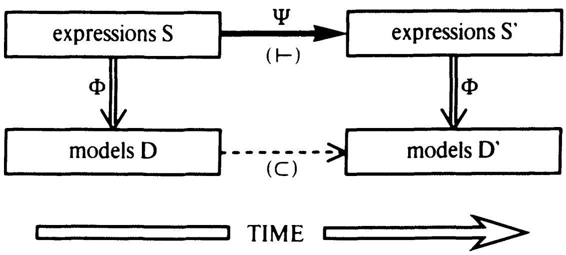



Fig. 1. Activity and semantics for a theorem prover

Four things about this situation are important. First, although proof theory's putative formality suggests that $\Psi$ must be definable independent of $\Phi$, you couldn't claim to have a proof-theoretic or inference relation except with reference to some underlying notion of semantic interpretation. Conceptually, at the very least, $\Psi$ and $\Phi$ are inextricably linked (salesmen for inference systems without semantics should be reported to the Better Business Bureau). Furthermore, the two relations are coördinated in the well-known way, using notions of soundness and completeness: inferences ($\Psi$) should lead from one set of sentences to another only if the latter are true just in case the former are true ($\vdash$ should honour $\models$). And truth, as we've already said, is defined as in terms of $\Phi$: the semantic relation to the world.

Second, even though the proof-theoretic derivability relation ($\vdash$) can be modelled as an abstract set-theoretic relation among sentences, I will view inference itself ($\Psi$) as fundamentally temporal — as an activity. 'Inference' is a derived noun; 'infer' first and foremost a verb, with an inherent asymmetry corresponding directly to the asymmetry of time itself. It might be possible to realise the provability relation non-temporally, for example by writing consequences of sentences down on a page, but you could hardly claim that the resulting piece of paper was doing inference.

Third, when its dynamic nature is recognised, inference is (quite properly) viewed as a temporal relation between sentences or states of the machine's memory, not as a function from times onto those corresponding sentences or states. Mathematically this may not seem like much of a difference, but conceptually it matters a lot. Thus, taking $\sigma$ to range over interpretable states of the system, and $t$ over times, $\Psi$ is of type $\sigma \rightarrow \sigma$, not $t \rightarrow \sigma$. Of course it will be possible to define a temporal state function of the latter type, which I will call $\Sigma$; the point is that it is $\Psi$, not $\Sigma$, that we call inference. Details will come later, but the relation between the two is roughly as follows: if $t'$ is one temporal unit past $t$, and $\Sigma(t) = \sigma$, then $\Sigma(t') = \Psi(\sigma)$. Inference, that is, has more to do with changes in state than with states themselves. To study inference is to study the dynamics of representational systems.

Fourth, of all the relations in Figure 1, only $\Psi$ need be effective; *neither $\Phi$ nor $\Sigma$ can be directly implemented or realised*, in the strong sense that there cannot be a procedure that uses these functions' inputs as a way of producing their outputs (the real reason to distinguish $\Psi$ and $\Sigma$). This claim is obviously true for $\Phi$. If I use the name 'Beantown' to refer to Boston, then the relation between my utterance and the town itself is established by all sorts of conventional and structural facts about me, about English, about the situation of my utterance, and so forth. The town itself, however, isn't the output of any mechanisable procedure, realised in me, in you, or in anyone else (fortunately — as it would be awfully heavy). It might require inference to understand my utterance, but that would only put you in some state $\sigma$ with the same referent as my utterance, or state. You don't compute the referent of an utterance you hear, that is, in the sense of producing that referent as an output of a procedure. Nor is the reference relation directly mediated, at least in any immediate sense, by the physical substrate of the world. Not even the NSA could fabricate a sensor, to be deployed on route 128, that could detect Boston's participation as a referent in an inferential act.[2]

That $\Sigma$ isn't computed is equally obvious, once you see what it means. The point is a strong metaphysical one: times themselves — metaphysical moments, slices through the *flux quo* — aren't causally efficacious constituents of activity; they don't have causal force. If they were, clocks wouldn't have been so hard to develop.[3] As it is, mechanisms, like all physical entities, manifest whatever temporal behaviour they do in virtue of momentum, forces acting on them, energy expended, etc., all of which operate in time, but don't convert time, compare it to anything else, or react with it. The only thing that's available, as a determiner of how a system is going to be, is how it was a moment before, plus any forces impinging on it. That, fundamentally, is why inference is of type $\sigma \rightarrow \sigma$, not $t \rightarrow \sigma$. *It could not be otherwise.* The inertness of gold, and the indifference of a neutrino, are nothing as compared with the imperturbability of a passing moment.

Given these properties of theorem provers, what can we say about clocks? Well, to start with, their situation certainly resembles that of Figure 1. As in the inference case, a clock's being in some state $\sigma$ represents ($\Phi$) it's being noon, or 7 : 15, or whatever; the interpretation function is what matters. Similarly, clocks, like theorem provers, change state ($\Psi$) in a simple but important way. Not only that; state change is what the clock designer has to work with; no mortal machinist, unfortunately, could build a device that would directly implement $\Sigma$. Furthermore, as in the case of the theorem prover, the change in state of the clock face is important *only because of its relation to its content.* Forget the Better Business Bureau; no one would buy a clock without a clue as to how its states represented time. Once again, systematic coördination between activity and interpretation is what matters.

But despite these similarities, there is a difference between clocks and theorem provers — suggested by the fact that many people (including me) would be reluctant to say that a clock was doing *inference.* To get at the difference, note that we haven't yet said what inference's coördinated pattern of events is for (on the face of it, going from truths to truths sounds a little boring). But the answer isn't hard to find: given a set of sentences or axioms that stand in (or enable you to stand in) a given semantical or informational relation to a subject matter, proofs or inference lead you to a new informational relation to the same, unchanged, subject matter. For example, the famous puzzle of Mr. S and Mr. P[4] focuses your attention on a pair of numbers under a peculiar description; a considerable amount of inference is required in order to give you access to those same numbers under a more traditional description (or give you access to other more familiar properties of numbers — there are many ways to discharge the ontological facts). The numbers themselves, however, and their possession of all the relevant properties, are expected to stay put during the inferential process. None of this implies, of course, that the subject matter of inference cannot itself be temporal, as the situation calculus and temporal logics illustrate. The point is that the temporality of the inference process and the temporality of the subject domain aren't expected to interact.

The situation for clocks, on the other hand, is almost exactly the opposite. What changes, across the time slice mediated by $\Psi$, isn't the stance or attitude or property structure that clocks get at. What changes, rather, is the subject matter itself. Clocks never have a moment's rest; no sooner have they achieved the desired relationship to the current time than time slips out from under their fingers, as if God were constantly saying "It's later than you think!" Clocks should perhaps be viewed as the world's first truth maintenance systems: they do what they do merely in order to retain the validity of their single semantic claim. Like any other meter or measuring instrument, they must track the world.

We can summarise:

> *Inference, at least as traditionally construed, is a technique that enables you to change your relation to a fixed subject matter. Clocks, in contrast, maintain a fixed relationship to a changing subject matter.*

If reconstructing time-pieces were really my subject matter, rather than simply being a foil, I might stop here. But my real interest is in developing a single semantical framework so that we can not only handle both of these cases (mathematical inference and real-time clocks), but also locate everything in between. So let's spend a minute to see how clocks fit into the general case.

## 3. SEMANTICALLY COHERENT ACTIVITY

I will use the term 'representational system' to cover anything whose behaviour fits within the broad space of semantically constrained activity. To be a representational system, in other words, is to be an element of the natural order that acts in a semantically coherent way. Of all possible kinds of representational activity, inference will be analysed as a particular type. The representational space is large, of course, and certainly includes all of computation (more about that in a moment), but it's still a substantive notion: not everything is in it. Planets, for example, are excluded, because planets don't represent their orbits; they just have them. Clocks, on the other hand, do represent the time, just as I can represent to myself how the sunrise looked this morning, as I drove down from the mountains.

Clocks do however fall outside most traditional models of computation, including the "formal symbol manipulation" model so familiar in cognitive science.[5] First, clocks (their faces, and the clockworks that run them) are fully concrete, physical objects, part of the natural order; nothing abstract here. Furthermore, this concreteness is crucial to our understanding of them; for some purposes one might treat clocks at a level of description that abstracted away from their physical being, including their temporal being, but since our purpose is to show how participation in their subject matter influences their design, to do so would be to miss what matters most. Second, at least some clocks (especially electrical ones operating on alternating current) are analog, even though more and more recent ones are digital. Third, to the extent that clocks have representational ingredients, there is no obvious decoupling to be made between a set of structures that represent and an independent process that inspects and manipulates them according to the shapes it sees. In other words, whereas Fodor's characterisation of a computer's "standing in relation" to representational ingredients suggests a modular division between symbols and processor, no such division is to be found in the chronological case. Fourth, there is another separation that can't be maintained in the case of clocks: that between "internal" and "external" properties. Time (rather like neutrinos) permeates everything equally, being as much an influence on internal workings as on surrounding context. And of course it is one and the same time, inside and out — clock design depends on this. Fifth, clocks, especially analog clocks, aren't usually "programmed" in any sense; they are designed, but they aren't universal computers specialised by physical encodings of time-keeping instructions. Like so many other properties of clocks, this is important, and leads to the sixth salient difference. Even on the view that Turing machines are concrete, physical objects (of which abstract mathematical quadruples are merely set-theoretic models), there is still no guarantee, given a particular universal one, that *any set of instructions could make it be, or even simulate, an accurate time keeper* — because there need be no consistency or regularity as to how long its state changes take. Turing machines, *qua* Turing machines, don't really participate.

I have come to believe, however, that not one of these six properties — being abstract, being digital, exhibiting a process/structure dichotomy, having a clear boundary between inside and outside, being programmable, or being necessarily equivalent to any Turing machine — is essential to the notion of computation on which the economy of Silicon Valley is based, or to the notion that underlies AI's hunch that the mind is computational. Quite the contrary. In (Smith, forthcoming) I argue for a much stronger conclusion: that the only regularity essential to computation has to do with computation being a physically embodied representational process — an active system or process whose behaviour represents some part or aspect of the embedding world in which it participates. This has the consequence, needless to say, of defining computation squarely in terms of undischarged semantical predicates. My position on theoretical cartography is therefore the inverse of Newell's (1980): whereas he thinks that computer science has answered the question of what it is to be a symbol, I believe in contrast that the integrity of computation as a notion rests full-square on semantics. So we have lots of homework, but it's homework for another day.

In the meantime, clocks are a good test case for comprehensive semantical frameworks. They lack many important properties of more general computers: they don't act, for example, or have sensors. But since every semantical property they do exhibit is one that computers can exhibit too — including participation — they are a useful design study.

## 4. THREE POINTS ON TWO FACTORS

In the previous section I distinguished two aspects or factors of any representational system: its behaviour, activity, or causal connection with the world (which I'll call the *first factor*) and its interpretation, content, or relation to its subject matter (the *second factor*). I have previously used this two-factor framework to reconstruct the semantics of Lisp, the programming *lingua franca* of AI, and argued for its general utility in analysing knowledge representation systems (Smith, 1982, 1984, 1986). And I will use it here, to analyse clocks. But three points must be made clear.

First, the ordering of the two factors may seem odd. There is no doubt that having interpretation or content — standing in semantic relation to a subject matter — is what particularly distinguishes the systems we are interested in. Given this pride of place, it might seem that content should be called first. But this is a mistake. We theoreticians typically treat semantics as primary when we analyse both natural and artifactual languages (such as the predicate calculus). We typically define semantics over rather abstract entities — sentence types, for example — and then understandably define the other dimension (proof theory, inference) over the same domain. But this overall strategy, especially in conjunction with the formal-symbol manipulation view of computation, gives a very abstract feel to inference, leading such people as Searle to wonder how, or even whether, such a system could ever possess genuine semantical powers. In contrast, by calling activity the first factor I want to recognise that computational systems are, first and foremost, systems in the world. *Everything* has what I am calling a first factor; that's what gives a system the ability to participate. The second factor of representation or content, which enables a system (a thinker, a clock) to stand in relation to what isn't immediately accessible or discriminable, is a subsequent, more sophisticated capacity. It is the second factor, furthermore, that distinguishes the representational or interpretable systems from other natural systems, but it distinguishes them as a sub-type, not as a distinct class. First factor participation in the world ("being there", roughly) is always available — which is fortunate, since it is only with respect to the first factor that second factor content can ever be grounded. In sum, recognising the metaphysical primacy of the first factor is an important ingredient in the defense of naturalism.

Second, there is a natural (almost algebraic) tendency to think that, in accepting a two-factor stance, one is committed to thinking that the two factors, in any given system, will in some important sense be *independent*. This tendency is amplified by the fact that in standard first-order logic an almost total independence of factors is achieved — this is one of the many meanings of the ambiguous claim that first-order logic is *formal*. Truth, content, and interpretation in logic are thought to be relatively independent of proof-theoretic role, and provability or inferential manipulation analogously independent of content or interpretation. In fact it is only because of this conceptual independence that proofs of soundness and completeness, even the very notions of soundness and completeness, are conceptually coherent. In computer systems, however — and minds, and clocks — there is no reason to expect this total degree of disconnection or independence. We should expect something more like the relationship between the mass and velocity of a physical object, on the one hand, and the center of gravity or resonance of the system of which it is a part, on the other: a web of constraints and conditions tying the two factors together, piece-wise, incrementally, thereby giving rise to a comprehensive whole. The situation of a complete proof system defined on an abstract set of mathematical expression types is extreme: a global but locally unmediated coherence, with no part of the proof or inferential system touching the semantic interpretation or content, except in the final analysis, when an outside theorist's proof grandly ties the whole thing together. For computers, and for us, it seems much more plausible to take a step or two apart from our subject matter, and then check in with it, to stay in "sync" — by taking a look, for example, or (following AT&T's recommendation) by reaching out and touching it. Participation is a resource, not a complication.

Third, as both the first two points make clear, it's a little hard to justify calling the two factors *semantical*, especially when the first is shared with every other participant in the natural order. It's not just that the first should be viewed as syntax, the second as semantics (as application of this more general framework to the predicate calculus would suggest). Rather, it's not clear what, if anything, the terms "syntax" and "semantics" should mean in a context where the coupling between factors is so much richer and more complex than in the traditional idealised case — if indeed they mean anything at all. Clockworks are mechanisms that enable first-factor behaviour — that much seems innocuous enough; calling the momentum of a clock's pendulum *semantic* is more difficult. First and second factors aren't distinct objects that somehow coöperate in engendering semantical activity; rather, one and the same causal constituents of a semantic system play both first and second factor roles.

This whole question is complicated by the use of the word 'semantics' (especially in AI) to describe inferential and structural relations among ingredients within a computational system. In (Smith, 1986) I attempt to resolve some of these issues, but instead of reconstructing that argument here I'll simply use the two-factor terminology without prejudice as to what does and doesn't have legitimate claim to the overloaded term.

## 5. THEORETIC MACHINERY AND ASSUMPTIONS

Let's look, then, at how clocks represent time, starting with some basic assumptions. As suggested in Figure 2, *qua* theorists we need accounts of four things:

1. States of the clock itself, including the face ($\sigma$);

2. The time or passage of time that the clock represents ($\tau$);
3. The first factor movement or state change between clock states ($\Psi$); and
4. The second factor representation relation ($\Phi$) between clock states and times.

All four of these are shared with standard semantical analysis: the first two would be the syntactic and semantic domains; the third, inference or proof theory; the fourth, semantics or interpretation.

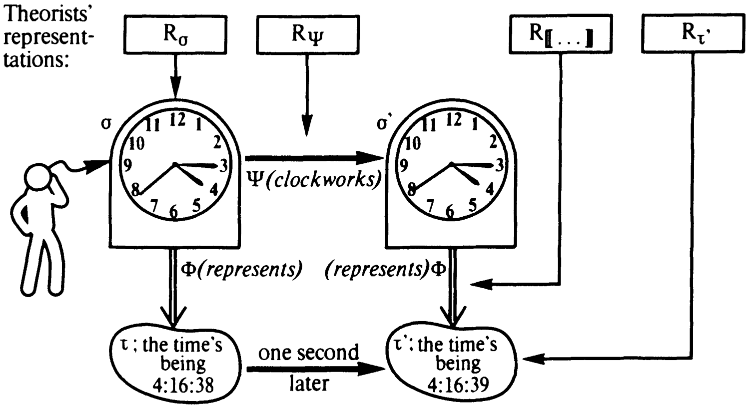



Fig. 2. The typology of clock semantics

I will adopt what I will call a *direct* rather than model-theoretic approach to these analytic tasks. Typically, when doing semantics, instead of talking directly about clock faces, orientations of hands, etc., you model them. For example, the state of a three-hand analog clock might be modelled as a triple, consisting of the orientations of the hour-hand, minute-hand, and second-hand, respectively, measured clock-wise from the vertical, in degrees. Thus the clock face shown in Figure 2 would be modelled as follows:

(S1) $M_\sigma$: $\langle 128.31666\ldots, 99.8, 228 \rangle$

The problem with this technique, however, as suggested in Figure 3, is that a model $M$ of a situation $S$ is itself a representation of $S$, since modelling is a particular species of representation ($M_\sigma$, for example, represents the clock face; it isn't the clock face, since for example it has

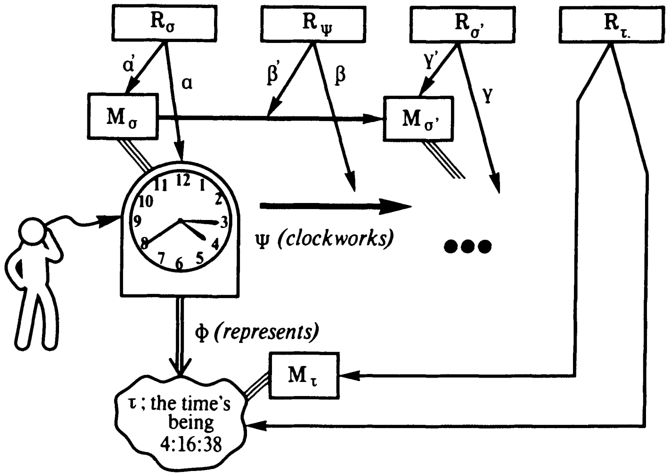



Fig. 3. The model-theoretic approach

a length of 3). The general character and complexity of the model—clock relation $M_\sigma$—$\sigma$, therefore, is the same as that between the clock and the time it represents ($\sigma$—$\tau$). It is therefore very hard to know whether what is crucial about $\sigma$—$\tau$ will be revealed or hidden in its $M_\sigma$—$M_\tau$ form. For example, using simple numbers to represent the orientations of hands presumes an absolute accuracy on the clock face, counter to fact. When studying something like natural language, which makes use of a much more complex representation relation than a model, the problems of indiscriminate theoretic modelling may be minor, or (more likely) go unnoticed. In our case, however, the representation relation we are studying, between clock faces and periodic times, is essentially an isomorphism. In this situation indiscriminate modelling would be much more theoretically distracting.

This direct semantical stance will have consequences, of two main sorts. First, we will need some machinery for talking precisely about the world without modelling it; for this I will use an informal "pocket situation theory", based unapologetically on Barwise and Perry (Barwise and Perry, 1983; Barwise, 1986a). Second, in the analog case it will be tempting to use some elementary calculus, which is problematic because a situation-theoretic reconstruction of continuity hasn't been developed yet. On the other hand, since the continuities underlying the integrity of the calculus presumably derive, ultimately, from the fundamental continuity of the physical phenomena that the mathematics was developed to describe, and since exactly those continuous phenomena will be our subject matter here, I'll take the liberty of applying its insights anyway. We're not really going to do any mathematics, so we won't get into trouble.

The direct semantical stance also highlights a question: how as theorists are we going to describe or *register* the phenomena we are going to study — i.e., in terms of what concepts, categories, and constraints are we going to explicate its regularity? When giving semantical analyses of linguistic or syntactic objects (sentences, expression types, etc.), tradition provides standard registrations in terms of constituent terms, predicate letters, etc. Similarly, purely abstract objects are typically categorised in advance in terms of a defining set of properties or relations. Clocks, on the other hand, are neither traditional nor abstract, so the question remains.

My metaphysical bias is to treat the world as infinitely rich, not only in the sense that there is more to everything than anything we can say, but also in that there is both more uniformity and structure, and more heterogeneity and individual difference, than theory or language can ever encompass. So I will say that clock faces, being actual, have enough structure so that one can be wrong about them, but still don't come labelled in advance by God, like plant slips at a nursery identified with a white plastic tag. Since every clock face, furthermore, exemplifies an infinite number of properties and relations (such as the property of being the subject matter of this paragraph), even after a basic registration scheme has been settled on, we have considerable latitude in making our choice.

None of this is intended to be problematic, or new; it's worth mentioning only because we need to make room for there being a difference between how we theorists do it, and how clocks do it, for themselves or (more likely, in the case of clocks) for their users. The problem is particularly acute for time itself, especially the periodic cycle of hours, minutes, and seconds that I keep referring to without explanation. If this were a paper on the semantics of time, not just of clocks, that explanation would have to be given, which would raise the incestuous fact that clocks themselves are probably largely responsible for the temporal registration (hours, mintues, seconds, etc.) of the times they represent, as argued for example by Lewis Mumford (1934). In this paper, however, I will merely adopt the periodic cycle without analysis, taking its explanation as a debt that should ultimately be paid.

Given these preliminaries, we should set out the ontological type structure, as summarised in Figure 4. Variables ranging over objects will be spelled with lower-case italic letters; over properties and relations, in lower-case Greek; over functions, in upper-case Greek. Thus $c$ and $c'$ will range over clocks; $t$, $t'$, etc., over full-blooded times, which are taken to be instantaneous slices through the metaphysical flux. Times are meant to include the time Kennedy was shot, the referent of "now", the point when the ship passed out of sight behind the island — that sort of thing. Intervals — intuitively, temporal durations between times — will be indicated by $\Delta t$, $\Delta t'$, etc. I will extend the use of '+' to allow adding intervals to times (overloading '+', as computer scientists say); thus $t + \Delta t$ will be of type $t$.

**Objects and Properties**

| | |
|---|---|
| $c, c', \ldots$ | — clocks |
| $t, t', \ldots$ | — times (instantaneous moments: slices through the flux quo) |
| $\Delta t, \Delta t', \ldots$ | — temporal intervals |
| $\tau, \tau', \ldots$ | — o'clock properties: being midnight, being 4 : 01 : 23,... |
| | $\tau_t$ — the o'clock property that holds of time $t$ |
| $\sigma, \sigma', \ldots$ | — states of clock faces (both hands pointing upwards,...) |
| | $\sigma_{c,t}$ — the state of clock $c$ at time $t$ |

**Primary Theoretic Functions**

| | |
|---|---|
| $\Psi : \sigma, \Delta t \rightarrow \sigma$ | — clockworks (from clock states and intervals onto clock states) |
| $\Sigma : c, t \rightarrow \sigma$ | — state function (from clocks and times onto clock states) |
| $[\ldots] : \sigma \rightarrow \tau$ | — content function (from clock states onto o'clock properties) |

**Overloaded Addition**

| | |
|---|---|
| $t + \Delta t : t$ | — times plus intervals are times |
| $\tau + \Delta t : \tau$ | — o'clock properties plus intervals are o'clock properties |

Fig. 4. Theoretic type structure

As opposed to times themselves, I will assume that times are located on the periodic cycle by what I will call the *o'clock properties*, such as that of "being 4 : 01 : 23", "being midnight", etc. The idea is not so much to license a continuum of distinct properties, but rather to assume that they arise out of a continuous relation between times and the abstract locations on the periodic time cycle to which they correspond

("4 : 00", etc.). Various possible explanations of this relation are possible, but since the intent of this paper is not to present an independently justified metaphysical account of time, but only to relate clocks to such a thing, I will employ a notation that simply picks up o'clock properties, whatever they are, from times that have them. Thus I will use $\tau_t$ to refer to the particular o'clock property that actually holds of time $t$. Also, I'll take differences between o'clock properties to be intervals (e.g., the difference between 5 : 00 and 3 : 00 will be two hours). Thus the sentence $\tau_t(t')$ says of time $t'$ that it has o'clock property $\tau_t$ — i.e., that it has whatever o'clock property $t$ has. The term $\tau_t - \tau_{t'}$ denotes an interval, of type $\Delta t$.[6]

In an analogous way, $\sigma$, $\sigma'$, etc. will range over a continuous (in the analog case) set of states of clock faces. For traditional circular analog clocks, a $\sigma$ representing 4 : 30 might be "having the hour hand at 135°, the minute hand at 180°, and the second hand at 0°, all measured clockwise from the '**XII**'."

Given this framework we can type the various semantical functions already encountered. As suggested in the previous section, $\Sigma$ will be a (non-computed!) function of type $t \rightarrow \sigma$, from times onto clock states; $\Psi$, a function of type $\sigma, \Delta t \rightarrow \sigma$, from clock states and temporal intervals onto clock states; and $\Phi$, a function of type $\sigma \rightarrow \tau$, from clock states onto o'clock properties. The important typological point for general semantic analysis is that both factors ($\Psi$ and $\Phi$) are defined as functions between states objects can be in, not between objects that are in them. This is as you would expect for scientific laws.

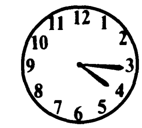
Two more theoretical points, before we take up the analysis itself. First, as just mentioned, I claimed in Section 2 that times $t$ weren't causal agents — that they couldn't be in the domain of a strongly effective realisable function. It is probably more important to the life of clock designers that the o'clock properties ($\tau$) are equally impotent. Even if it's 4 : 00 all around you, there's nothing that it's being 4 : 00 can cause to happen — like serving tea and crumpets. With respect to engendering behaviour, a moment's being midnight is more like Boston's being a referent than it is like ice-cream's being sticky: it just isn't the sort of thing that a sensor could detect. So functions of the form $\tau \rightarrow x$ are as unrealisable (in the strong sense discussed earlier) as those of type $t \rightarrow x$, for arbitrary $x$. Such is life.

Second, I mentioned earlier that using numbers to represent the orientations of the hands of clocks presumes an accuracy that outstrips physical plausibility. Even if quantum physics would theoretically support there being a fact of the matter as to where a hand points within $\pm 10^{-50}$ degrees, say (which it won't), there are also pragmatic realities of producing a macroscopically observable clock subject to the forces of gravity, anomalies of manufacture, etc. Furthermore, if the hour-hand were anything like this accurate, then at least for theoretical purposes the minute and second hands would be redundant: a perfect observer could gaze at a clock and read off a time of, say, 4 : 15 : 38 : 17.[7] One might object, of course, that human users wouldn't be able to register the hour-hand more accurately than, say, $\pm$ 1° or 2°, and therefore, even with internal calculation, wouldn't be able to determine the time on a single-handed clock more accurately than to within about 5 minutes, no matter how much more accurately than that the time was actually signified. In fact casual observation suggests that hour hands on modern analog clocks are much more accurately positioned than necessary merely to determine which hour the minute hand signifies time with respect to.

These issues again raise the question of the relation between how we as theorists register clock faces and the times they represent, and how clock faces themselves register those represented times.[8] But I won't answer this question here, since we will primarily be dealing with semantic constraints on clock and time registrations, rather than with individual registrations themselves.

## 6. TEMPORAL REPRESENTATION: THE SECOND FACTOR

Given these premises and caveats, let's look at how times are represented. Intuitively, we are aiming for something like the following:

(S2) [  ] = the property of being 4 : 16



To do this, we start with $\Phi$, of type $\sigma \rightarrow \tau$ from (representing) states of clock faces onto (represented) states of times — i.e., onto o'clock properties. Instead of the name '$\Phi$', however, I will use so-called semantic brackets ('[. . .]'), in the following way: [. . . $\sigma_{c,t}$ . . .] will be the o'clock property signified by the state '. . . $\sigma_{c,t}$ . . .', assuming that $\sigma_{c,t}$ is the state $\sigma$ of clock $c$ at time $t$. For example, the sentence $[\sigma_{c,t'}](t)$

claims of time $t$ that it has the o'clock property that clock $c$ indicates at time $t'$; $[\Psi(\sigma_{c,t'}, \Delta t)](t)$ claims of time $t$ that is has the o'clock property that clock $c$ would indicate $\Delta t$ later than time $t'$, since $\Psi(\sigma_{c,t'}, \Delta t)$ is the state it would then be in.

Using this terminology, we can say that clock $c$ is chronologically *correct* at time $t$ just in case $t$ is of the type that the clock then indicates:

$$\text{(S3)} \qquad \text{Correct}(c, t) \equiv_{\text{df}} [\sigma_{c,t}](t)$$

So far, of course, this is a constraint on possible interpretation functions [. . .], since we haven't defined any specific instances. Longer-term notions of correctness (over extended intervals, for example) could be defined by quantifying over times; similarly, approximate degrees of correctness could be characterised in terms of the difference between what time it actually was and what time was indicated.

## 7. CLOCKWORKS: THE FIRST FACTOR

With respect to operation, the basic point is this: if at time $t$ a clock is so-and-so ($\sigma$), then at some point $\Delta t$ later it will be such-and-such ($\sigma'$), where $\sigma'$ is $\Psi(\sigma, \Delta t)$. The function $\Psi$, which takes a clock into the future in this way, must be realised by the underlying physical machine — must be implemented, that is, by the clockworks. The important constraint on this relation, which I will call the *realisability* constraint, is that $\Psi(\sigma, \Delta t)$ can depend on $\sigma$ and on $\Delta t$, but not on the time $t$ that is "happening" when the clock is in state $\sigma$.

In symbol manipulation or semantical contexts, where time and symbols are both digital, we often view $\Psi$ as a state-transition function (such as for a Turing machine controller). In such cases $\Delta t$ drops out, being assumed to be a single time "click". For example, suppose $\mathbb{S}$ is a (discrete) function from states to states ($\sigma \rightarrow \sigma$). The equation for a single state change, of the sort one would expect in a digital world, would be something like $\sigma' = \mathbb{S}(\sigma)$, or generalised to $\Delta t$'s of $n$ tick's duration, $\sigma' = \mathbb{S}^n(\sigma)$. In the continuous world of physical mechanics, on the other hand, $\Psi$ is merely "what the world does", explained in terms of velocities, accelerations, etc. From this perspective, the calculus can be viewed as a theoretical vehicle with which to explain first factor futures for continuous systems, where the state $\sigma$ of some system in an amount of time $\Delta t$ after it is in a starting state $\sigma_0$, assumed to depend on the continuity of the underlying phenomena, can be expressed in the familiar equation:

$$\text{(S4)} \qquad \sigma = \sigma_0 + \frac{\mathrm{d}\sigma}{\mathrm{d}t}\,\Delta t + \frac{1}{2}\,\frac{\mathrm{d}^2\sigma}{\mathrm{d}t_2}\,\Delta t^2 + \ldots$$

My aim isn't to contrast the discrete and continuous cases (I want to develop results applicable to both analog and digital clocks), but rather to highlight the common focus on state change, represented computationally by state transition functions, and physically by temporal derivatives. There is, however, this apparent difference: the theoretic notions employed in physics (force, acceleration, etc.) are essentially relative; they describe how the new state will differ from the old one. The real identity of the new state — what state the system will actually arrive in — is obtained, as if it were conceptually subsidiary, by altering the previous state in the prescribed manner. State transition tables, in contrast, are typically absolute. They still describe state *change*, of course — they aren't temporal state functions like $\Sigma$. The point is that the new state is specific *de novo*, so to speak, not as a modification of the old one, though of course the extent to which the new state differs from the old can be calculated as a difference between the two.

This difference in theoretic stance, however, is superficial, since in actual use (in describing programs, operations on memory, etc.) state transition functions are defined with explicit reference to how the new state differs from the old. In giving environment transition functions, for example, showing the consequence of binding a variable, the requisite function from total environments onto total environments is defined as modifying the value of the given variable in question, and *otherwise being just like the prior one.* Practice suggests, in other words, that in the computational case, as in the physical case, state change is conceptually prior, new total state conceptually dependent. Thus there is general support for our specific focus on $\Psi$.

Intuitively, a proper $\Psi$ for a clock will specify that it runs at the right speed. It is easy enough to calculate, in the case of circular analog clocks, that this amounts to having the hour hand, minute hand, and second hand rotate at 0.008333°/sec, 0.1°/sec, and 6°/sec, respectively. But to *characterise* correctness this way is exactly like characterising the correctness of a proof procedure by pointing to the syntactic inference rules. It may indeed be true that, if this condition is met, the clock will be running at the correct speed, but that doesn't mean that this condition expresses what it is to be running correctly. Rather, we want to say that if at time $t$ (say, 12 : 00) a clock designates o'clock property $\tau_{t'}$ (say, 3 : 11), then at time $t + \Delta t$ (12 : 01, for a one minute $\Delta t$) it should indicate the property of being $\Delta t$ later, i.e., $\tau_{t'+\Delta t}$ (3 : 12). We can do this as follows:

(S5) $\quad \text{Right-speed}(c, t, \Delta t) \equiv_{\text{df}} [\sigma_{c,\, t+\Delta t}] = [\sigma_{c,\, t}] + \Delta t$

which has the consequence, given the definition of $\Psi$, that

(S6) $\quad [\Psi(\sigma_{c,\, t}, \Delta t)] = [\sigma_{c,\, t}] + \Delta t$

Properly, we should state something stronger: that a clock runs correctly throughout the interval from $t$ to $t + \Delta t$ if and only if it advances at the right speed for the whole time (note that the following is neutral as to whether this is a continuous or discrete interval — i.e., as to whether $\forall$ is a discrete or continuous quantifier):

(S7) $\quad \text{Right-speed}(c, t, \Delta t) \equiv_{\text{df}} \forall \Delta t' \mid 0 \leqslant \Delta t' \leqslant \Delta t$
$\qquad\qquad [\sigma_{c,\, t+\Delta t'}] = [\sigma_{c,\, t}] + \Delta t'$

again directly yielding

(S8) $\quad \forall \Delta t' \mid 0 \leqslant \Delta t' \leqslant \Delta t\ [\Psi(\sigma_{c,\, t}, \Delta t')] = [\sigma_{c,\, t}] + \Delta t'$

These equations involve a property identity, but I defer any questions on that issue to situation theory. Note also that in each version the two instances of '+' are of different types: the first takes a time and an interval onto a time, the second an o'clock property and an interval onto an o'clock property. No problem.

Given (S3) and (S7), the temporal analogues of soundness and completeness can be proved: if a clock is correct at time $t$, and runs at the right speed during the interval from $t$ to $t'$, then it will be correct during that interval, and conversely if it is correct throughout the interval it must be running at the right speed. But it is more fun to do this in the continuous case, so let's turn to that.

Very simply, we want to talk of an analog clock's running at the right speed *instantaneously*, which means, intuitively, that we should differentiate the temporal state function $\Sigma$ — or, what is equivalent, take the limit of $\Sigma$ as $\Delta t$ approaches 0, in the standard way:

(S9) $$\underset{\Delta t \to 0}{\text{limit}} \frac{(([\sigma_{c,\, t}] + \Delta t) - [\sigma_{c,\, t}])}{\Delta t} = \underset{\Delta t \to 0}{\text{limit}} \frac{([\sigma_{c,\, t+\Delta t}] - [\sigma_{c,\, t}])}{\Delta t}$$

Since, as we've already said, differences between o'clock properties are intervals, the left side of this reduces to $\text{limit}_{\Delta t \to 0}\ (\Delta t/\Delta t)$, which is identically 1, yielding:

$$\text{(S10)} \quad 1 = \lim_{\Delta t \to 0} \frac{([\sigma_{c,\, t+\Delta t}] - [\sigma_{c,\, t}])}{\Delta t}$$

The right hand side, however, is merely the derivative, with respect to time, of the interpretation of the state. We can't differentiate $\sigma$ directly, its not being a function of time (in fact it's not a function at all), but we can rewrite (S10) in terms of $\Sigma$:

$$\text{(S11)} \quad 1 = \lim_{\Delta t \to 0} \frac{([\Sigma(c, t + \Delta t)] - [\Sigma(c, t)])}{\Delta t}$$

This enables us to take the limit ($\Sigma$ is continuous by assumption), since the right hand side is the derivative of a function that is essentially the composition of the second and first factors ($[\ldots] \circ \Sigma$).[9] I will abbreviate this $[\Sigma]$, giving us:

$$\text{(S12)} \quad \text{Right-speed}_{\text{analog}}(c, t) \equiv_{\text{df}} \frac{\text{d}}{\text{d}t}[\Sigma] = 1$$

If the derivative (with respect to time) of a function is unity, of course, it follows that the function is of the form $\lambda t \cdot t + k$ — or rather, in our case, $\lambda t \cdot \tau_t + k$, as dictated by our type constraints — where $k$ is a constant of type $\Delta t$. This is exactly what we would expect: the constant represents the error in the clock's setting — the difference between the actual and indicated times. The equation, predictably, says that if a clock is running at the right speed the error will (instantaneously) remain constant. Furthermore, since (S3) implies that

$$\text{(S13)} \quad \text{Correct}(c, t) \Leftrightarrow [\Sigma(c, t)](t)$$

it follows that the constant would be 0 for a correctly set clock, as expected.

We can summarise these results as follows:

$$\text{(S14)} \quad \text{Correct}(c, t) \equiv_{\text{df}} [\Sigma(c, t)](t)$$

$$\begin{array}{lll} \text{(S15)} & \text{Right-speed}(c, t, t') \equiv_{\text{df}} & \forall \Delta t \,|\, 0 \leqslant \Delta t \leqslant (t' - t) \\ & & [\sigma_{c,\, t+\Delta t}] = [\sigma_{c,\, t}] + \Delta t \\ & \quad \textit{implying that} & \forall \Delta t \,|\, 0 \leqslant \Delta t \leqslant (t' - t) \\ & & [\Psi(\sigma_{c,\, t}, \Delta t)] = [\sigma_{c,\, t}] + \Delta t \\ & \quad \textit{implying that} & \forall \Delta t \,|\, 0 \leqslant \Delta t \leqslant (t' - t) \\ & & [\Sigma(c, t + \Delta t)] = [\Sigma(c, t)] + \Delta t \end{array}$$

(S16) $\text{Right-speed}_{\text{analog}}(c, t) \equiv_{\text{df}} \frac{\mathrm{d}}{\mathrm{d}t}[\Sigma] = 1$

and in their terms define what it is for a clock to be "working" properly from time $t$ to $t + \Delta t$:

(S17) $\text{Working}(c, t, \Delta t) \equiv_{\text{df}} \text{Correct}(c, t) \wedge \text{Right-speed}(c, t, \Delta t)$

(S18) $\text{Working}_{\text{analog}}(c, t) \equiv_{\text{df}} \text{Correct}(c, t) \wedge \text{Right-speed}_{\text{analog}}(c, t)$

For either version, the constraint can be shown to be satisfied (over the interval, or instantaneously, depending) in exactly the following condition:

(S19) $[\Sigma(c, t)] = \lambda t \cdot \tau_t$

Given the abbreviation adopted above, we can state this even more simply:

(S20) $[\Sigma] = \lambda t \cdot \tau_t$

I would be the first to admit that (S20) is obvious — at least retroactively, in the sense that, once stated, it is hard to imagine thinking anything else. In English, it says that the state function and the interpretation function should be proportional inverses; given a clock that (so to speak) maps time onto some sort of complex motion, the appropriate interpretation function is merely that function that maps that motion back onto the o'clock properties of the linear progression of time that was started with. So the putative clock of Figure 5, for example, with a million-mile pendulum and a 24 hour period, would have a pointer position ($\sigma$) proportional to $\sin(t)$, and an interpretation function analogously proportional to $\sin^{-1}(\sigma)$.[10]

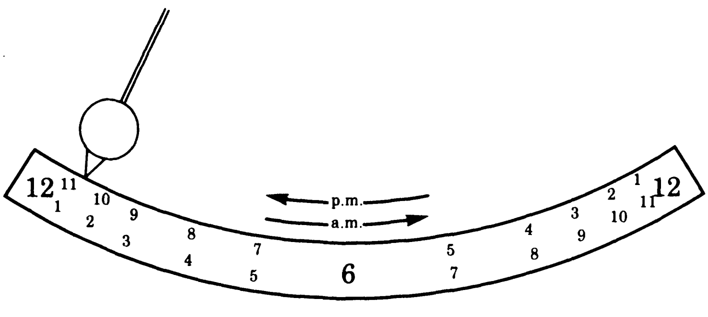



Fig. 5. The million mile clock

Still, (S20) isn't trivial, for a reason that shows exactly why clocks were hard to build. It says that working clocks map all times onto their o'clock properties. The problem for clockmakers is that $\Sigma$ isn't directly computable, since, to repeat, neither times nor o'clock properties enter into causally efficacious behaviour. What can be implemented is $\Psi$, not $\Sigma$, and $\Psi$ is essentially the temporal derivative of $\Sigma$.

In sum, we have determined the function of clockworks: to integrate the derivative of time. When you set the hands on the clock's face, you are supplying the integration constant.

## 8. MORALS AND CONCLUSIONS

What have we learned? Four things, other than some fun facts to tell our friends.

The first has to do with the interaction among notions of participation, realisation, and formality. Clocks' participation in their subject matter (being temporal, as a way of measuring time), which depends on their physical realisation, might seem to violate the formality constraint that is claimed to hold of computational systems more generally. In fact, however, clocks' temporality doesn't relieve them of much of the structure that characterises more traditional systems: separable $\Psi$ and $\Phi$, the possibility of being wrong, etc. This similarity of clocks to symbol manipulation systems arises from the fact that the particular aspect of times that clocks represent — the o'clock properties — aren't within immediate causal reach of a clockwork mechanism (or of much else, for that matter). In (Smith, forthcoming) I argue that this is a manifestation of a deep truth: the limitations of causal reach are the real constraints on representational systems. Formality, as a notion, is merely a cloudy and approximate projection of these limitations into a particular construal of the symbolic realm.

The second moral has to do with the impact, for theoretical analysis, of the relation between $\Psi$ and $\Phi$. The function $\Psi$, realised in clockwork, is what the engineers must implement; without an analysis of it, effective clocks couldn't be designed. But theories of clocks must go much further. Our characterisation of what it was for a clock to work properly, for example, had to reach beyond the immediate or causally accessible aspects of the underlying clockwork mechanism. Whatever one might think about more complex cases, methodological solipsism doesn't work in this particular instance.

Third, the similarity between the state transition functions of computer science and the temporal derivatives of mechanics, both of which focus not on time itself but on temporal change, suggest the possibility of a more unified treatment of representational dynamics in general. So far most of what we have to say deals with specific cases. So, for example, in Section 2 we characterised inference as a particular species of representational activity, having to do with changing content relations to a fixed subject matter. Inference was contrasted with clock's maintenance of a fixed content relation to a changing subject matter. Remembering what is perceived, to take quite a third sort of representational behaviour, is a form of retaining a fixed relation to a fixed subject matter in ways that make it immune to changes in the agent's circumstances. It doesn't seem impossible that a common framework could be uncovered.

Fourth, and finally, by occupying a place very different from that of either Turing machines or traditional theorem provers, clocks help illuminate the fundamental constraints governing computers and representational systems in general. As Figure 6 suggests, there are two basic kinds of constraint — causal relations and content relations — that a representational system must coördinate as it moves through the world.

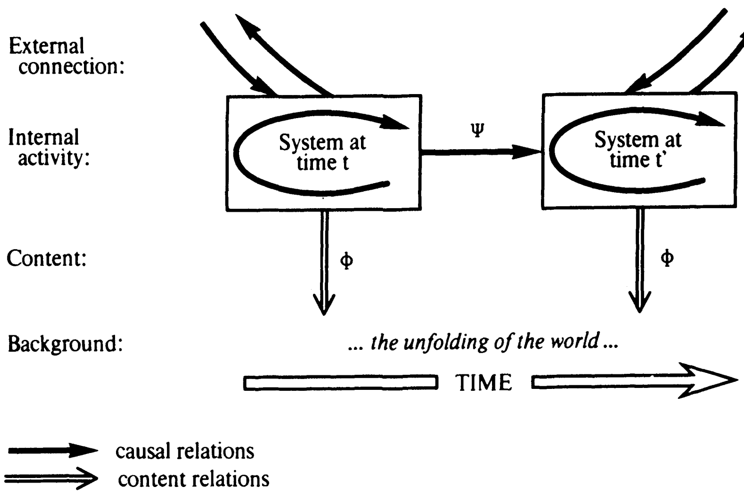



Fig. 6. C5: Coordinated constraints on content and causal connection.

Both kinds, in general, will be complex — much more so than we have seen in the case of clocks. Two aspects of content we haven't encountered, for example, are its "situational" dependence on surrounding circumstances, as discussed for example in (Barwise, 1986b; Perry, 1986), and the three-way semantic interactions among language, mind, and world that arise in cases of communication. Causal connections are similarly complex, and can be broken down into three main groups:

1. *Internal activity or behaviour*: the relation between a system at some time and the same system shortly thereafter. This is what we've called $\Psi$.
2. *External connection*: actions the system takes that affect the world, and effects on the system of the world around it — the results, that is, of sensors and effectors. (Clocks have none of this, but other systems are clearly not so limited.)
3. *Background dynamics*: the progress or flow of the surrounding situation. The passage of time would be counted as one instance, as would one's conversant's behaviour, or the passing visual scene.

In the traditional case of pure mathematical inference, there is no connection (action or sensation), and the background situation, as we saw, is presumed to stay fixed. Barwise's construal of "formal inference" (the "non-situated" reading), (Barwise, 1986b, p. 331) strengthens this constraint by assuming that the content relation is also independent of surrounding situation. The clock example gives us a different point in the space: again no connection, an essentially unchanging (and relatively situation-independent) content relation, but an evolving background situation, mirrored in the internal activity or behaviour. Finally, semantic theories of action, involving everything from intentionally eating supper to making a promise, must deal with cases where the connection aspect makes a contribution. They must therefore deal with cases where the surrounding situation is affected not only by its own background dynamics, but as a result of internal activity on the part of the representational agent. But simpler systems will require an analysis of external connection, as well: computerised (ABS) brakes on late model cars, for example, are directly connected (even vulnerable) to the content of their representations, in a way that seems to free them from the need to have their representational states externally interpreted.

In the end, however, the similarity among these systems is far more important than the variance. We can put it this way. Causal participation in the world is ultimately a two-edged sword. On the one hand, it is absolutely enabling. Not only could a system not exist without it, but in a certain sense it's total: everything the system is and does arises out of its causally supported existence. There are no angels. On the other hand, causal connection on its own — unless further structured — limits a system's total participation in the world to those things within immediate causal reach.

Representation, on this view, is a mechanism that honours the limits of causal participation, but at the same time stands a system in a content relation to aspects of the world beyond its causal reach. The trick that the system must solve is to live within the limits (and exploit the freedoms!) of the causal laws in just such a way as to preserve its representational stance to what is distal. This much is in common between an inference system and a clock.

## ACKNOWLEDGEMENTS

This paper grew out of a bet made with Richard Weyhrauch during a discussion late one night in a bar in Alghero, Sardinia about what was involved in reading one's watch. Specifically, I promised to develop a semantical analysis of the familiar behaviour cited in the third paragraph of the paper: waiting a second to see whether a watch moves before reading the time. This paper is part one of the answer; interpreting a clock will come later. My thanks to him and other members of the Cost 13 Workshop on Reflection and Meta-Level Architectures, especially including Jim des Rivières and John Batali. Thanks also to Jon Barwise, John Etchemendy, David Israel, John Perry, Susan Stucky, and the other members of the situation theory and situation semantics (STASS) group at CSLI, to Pat Hayes for discussions of measurement, and to John Lamping for his help on celestial mechanics. The research was supported by Xerox Corporation and the System Development Foundation, through their mutual support of the Center for the Study of Language and Information.

## NOTES

[1] Clocks represent time for us, as it happens, not for themselves, but that will count, at least here. I'm sympathetic to the distinction between original and derivative semantics (in fact I'm interested in participation for just such reasons), but I am very much against relativising representation to an observer at the outset, especially to a human observer

(Winograd and Flores, 1986). To do that would be to abandon any hope of explaining how the human mind might itself be representational, my ultimate goal. See (Smith, forthcoming).

[2] In computer science the claim that reference isn't computed is viewed suspiciously, for a very interesting reason. To see it, consider why the claim is true. Suppose in a room of 100 people some person $A$ is the average height. Then suppose a new person enters the room. Suddenly, *and without any computation*, a different person $B$ will be the average height. No work needs to be done to lift the property from $A$ and settle it on $B$; no energy expended, no symbols massaged. The new state just *comes to be*, automatically, in virtue of the maze of conditions and constraints that hold. Reference, I take it, is something like that; conditions and constraints hold so that, when a word is uttered or a thought entertained, some object becomes the referent. (Nor is it possible to reply "Well, the room computed it"; on that recourse everything that happens would be computed, which would make the word 'compute' vacuous.)

How could computer scientists object to this? For the following reason. Note that the way that $B$ becomes the person of average height is by participating in the situation at hand: he enters the room. Participation, in other words, is what enables relationship to exist. Computers, on the other hand, are traditionally viewed in purely abstract terms, and abstractions, whatever they are, presumably don't participate. The closest an abstraction comes to the property of average height — or indeed to anything at all — is by designating it. And so, because of this abstract conception of computers, one gets lulled into thinking that *everything* has to come into being in this disconnected, putatively "computational" way.

Needless to say, I don't believe the abstract conception of computers is right. More strongly, I am arguing that participation — the opposite of abstraction — is exactly what allows you to connect to the world in other ways than through explicit symbol manipulation. See Section 8, and (Smith, forthcoming).

[3] For accurately measuring distances on roads, one attaches a "fifth wheel" to a car and reads off the passing miles. Maybe, if time had been causally efficacious, we could have built clocks the same way, running a wheel against time and reading off the passing seconds.

[4] There are two numbers between 1 and 100. Mr. P knows their product, and Mr. S their sum. They have the following conversation:

Mr. P. *I don't know the numbers.*
Mr. S. *I knew you didn't. Neither do I.*
Mr. P. *Now I do.*
Mr. S. *Now I do too.*

What are the numbers?

The earliest publication of this problem I am aware of is by H. Freudenthal in the Dutch periodical *Nieuw Archief Voor Wiskunde*, series 3, **17**, 1969, p. 152 (a solution by J. Boersma appears in the same series, **18**, 1970, pp. 102–106). It was subsequently submitted by David J. Sprows to *Mathematics Magazine* **49**(2), March 1976, p. 96 (solution in **50**(5), Nov. 1977, p. 268). Perhaps the most widely read version appears in Martin Gardner's 'Mathematical Games' column in *Scientific American* **241**(6), Dec. 1979, pp. 22–30, with subsequent discussions and slight variations in 1980: **242**(3), March, p. 38; **242**(5), May, pp. 24–28; and **242**(6), June, p. 32.

[5] The two other primary models, conceptually distinct from the formal symbol manipulation idea, are the automata-theoretic notion of a digital or discrete system and the related idea of a machine whose behaviour is equivalent to that of some Turing machine. Although the formal symbol manipulation view seems to go virtually unchallenged in cognitive science, the other two have much more currency in modern computer science. See (Smith, forthcoming).

[6] A more detached theoretic viewpoint should point out that o'clock properties $\tau_t$ are in fact two-place relations between times and places (a time that is midnight in London will be 7 : 00 p.m. in New York). More generally, whereas I assume throughout that activity ($\Psi$) and interpretation ($\Phi$) are functions, they should properly be viewed as more complex relations between agents and their embedding circumstances.

[7] "**third**, n. ... 5. The sixtieth part of a second of time or arc." — *Webster's New International Dictionary*, Second Edition. New York: G. & C. Merriam, Co. 1934.

[8] Clock faces, and representations in general, don't need to register themselves, in order to represent.

[9] Strictly speaking this isn't quite accurate, since both $[\ldots]$ and $\Sigma$ should depend on $c$ and $t$: the function we are differentiating should really be $\lambda c, t. \ [\Sigma(c, t)]$. But being strict would add only complexity, not insight.

[10] This clock would be even harder to build than you might suppose. At first blush, it might seem as if the equation of motion for a pendulum would imply that a very large bob, swinging in an arc at the surface of the earth (an arc, say, 100 feet in length), whose mass completely dominated the mass of a long string by which it was suspended from a geosynchronous point 1 150 000 miles above the surface of the earth, would have a period of 24 hours. Unfortunately, however, such a device would have a period of slightly less than an hour and a half. Why this is so, and how to modify the design appropriately are left as an exercise for the reader (hint: the result would be difficult to read).

## REFERENCES

Barwise, Jon, and Perry, John: 1983, *Situations and Attitudes*, Bradford Books, Cambridge, MA.

Barwise, Jon: 1986a, 'The Situation in Logic — III: Situations, Sets and the Axiom of Foundation', in Alex Wilkie (ed.) *Logic Colloquium 84*, North Holland, Amsterdam. Also available as CSLI Technical Report CSLI-85-26 from the Center for the Study of Language and Information, Stanford University, 1985.

Barwise, Jon: 1986b, 'Information and Circumstance', *Notre Dame Journal of Formal Logic*, **27**(3) 324—338.

Brachman, Ronald J., and Levesque, Hector J. (eds.): 1985, *Readings in Knowledge Representation*, Morgan Kaufmann, Los Altos, CA.

Fodor, Jerry: 1975, *The Language of Thought*, Thomas Y. Crowell Co., New York. Paperback version, Harvard University Press Cambridge, MA, 1979.

Fodor, Jerry: 1980, 'Methodological Solipsism Considered as a Research Strategy in Cognitive Psychology', *The Behavioral and Brain Sciences*, **3**(1), pp. 63—73. Reprinted in Fodor, Jerry, *Representations*, Bradford, Cambridge, MA, 1981.

Mumford, Lewis: 1934, *Technics and Civilization*, Harcourt, Brace & Co., New York. Reprinted 1943.

Newell, Allen: 1980, 'Physical Symbol Systems', *Cognitive Science* **4**, 135—183.

Perry, John: 1986, 'Circumstantial Attitudes and Benevolent Cognition', in J. Butterfield (ed.), *Language, Mind and Logic*, pp. 123—134, Cambridge University Press Cambridge.

Postman, Neil: 1985, *Amusing Ourselves to Death: Public Discourse in the Age of Show Business*, Penguin Books, New York.

Smith, Brian C.: 1982, *Reflection and Semantics in a Procedural Language*, Technical Report MIT/LCS/TR-272, M.I.T., Cambridge, MA, 495 pp. Prologue reprinted in (Brachman and Levesque, 1985), pp. 31—39.

Smith, Brian C.: 1984, 'Reflection and Semantics in Lisp', Conference Record of 11th Principles of Programming Languages Conference, pp. 23—35, Salt Lake City, Utah. Also available as Xerox PARC Intelligent Systems Laboratory Technical Report ISL-5, Palo Alto, California, 1984.

Smith, Brian C.: 1986, 'The Correspondence Continuum', appeared with the Proceedings of the Sixth Canadian AI Conference, Montreal, Canada, May 21—23. Available as CSLI Technical Report CSLI-87-71 from the Center for the Study of Language and Information, Stanford University, 1987. To appear in *Artificial Intelligence* (forthcoming).

Smith, Brian C.: *Is Computation Formal*? MIT Press/A Bradford Book Cambridge, MA. (Forthcoming.)

Winograd, Terry, and Flores, Fernando: 1986, *Understanding Computers and Cognition: A New Foundation for Design*, Ablex, Norwood, New Jersey.

*Intelligent Systems Laboratory, and*
*Center for the Study of Language and Information*
*Xerox Palo Alto Research Center*
*Palo Alto, CA 94304, U.S.A.*