Interaction and Representation

Mark H. Bickhard

Mark H. Bickhard
Department of Philosophy
15 University Drive
Lehigh University
Bethlehem, PA 18015
610-967-6770
mhb0@lehigh.edu

Interaction and Representation

Mark H. Bickhard

Abstract

There is a form of representation that is naturally emergent in the organization of interactive systems. *Interactive representation* has claims to be the fundamental form of representation, from which all others are derivative.

In particular, it naturally satisfies a meta-epistemological criterion that is not addressed by standard approaches in contemporary literature, and is arguably impossible to satisfy within any version those standard approaches. Furthermore, the interactive approach naturally avoids other multiple aporias that bedevil standard approaches.

Much effort has been devoted in recent literature to attempts to satisfy a critical meta-epistemological criterion: representation must be capable of being in error. The criterion that I will apply is a strengthening of this one: representation must be capable of being in error in such a way that that condition of being in error is detectable by the agent or system that is doing the representing — the meta-epistemological criterion of *system detectable error*, for short. Whatever the status may be of current attempts to satisfy the criterion of error, the criterion of *system detectable* error is not even addressed.

The interactive approach to representation has strong affinities with the general pragmatist approach and with the Heideggerian skill intentionality version of it. The model that I will outline, however, will be my own approach to these shared pragmatist insights.

The basic insight is that interactions can possess truth conditions without explicitly representing those truth conditions, and that the course of an interaction can detect failures of those truth conditions. This, I argue, is the fundamental form from which all representation is derived.

Interaction and Representation

Mark H. Bickhard

There is a form of representation that is naturally emergent in the organization of interactive systems. *Interactive representation* has claims to be the fundamental form of representation, from which all others are derivative.

In particular, it naturally satisfies a meta-epistemological criterion that is not addressed by standard approaches in contemporary literature, and is arguably impossible to satisfy within any version of those standard approaches. Furthermore, the interactive approach naturally avoids other multiple aporias that bedevil standard approaches.

Much effort has been devoted in recent literature to attempts to satisfy a critical meta-epistemological criterion: representation must be capable of being in error (Dretske, 1981, 1988; Fodor, 1987, 1990a, 1990b; Hanson, 1990; Loewer & Rey, 1991; Millikan, 1984, 1993). The criterion that I will apply is a strengthening of this one: representation must be capable of being in error in such a way that that condition of being in error is detectable by the agent or system that is doing the representing — the meta-epistemological criterion of *system detectable error*, for short. Whatever the status may be of current attempts to satisfy the criterion of the possibility of error, the criterion of *system detectable* error is not even addressed.

The interactive approach to representation has strong affinities with the general pragmatist approach (Hoopes, 1991; Rosenthal, 1983; J. E. Smith, 1987) and with the Heideggerian skill intentionality version of it (Heidegger, 1962; Dreyfus, 1967, 1982, 1991; Dreyfus & Haugeland, 1978; Guignon, 1983; Okrent, 1988). The model that I will outline, however, will be my own approach to these shared pragmatist insights.

The basic insight is that interactions can possess — can presuppose — truth conditions without explicitly representing those truth conditions, and that the course of an interaction can detect failures of those truth conditions. This, I argue, is the fundamental form from which all representation is derived. That is:

- actions and interactions can involve presuppositions about the environment in which those actions and interactions take place;
- those presuppositions can be false;
- the failure of an interaction is an indication that at least of one of those presuppositions was in fact false; and
- such interaction failure is detectable in and by the interactive system itself.

The connection of these points to representation is that:

- only that which can be in error for the system can be “not in error” for the system;
- only that which can be “not in error” for the system can be representational for the system.

Representation must be capable of truth value, and something is representational only for that for which it is capable of truth value. If we want representation for a system itself, not derivative representation from the perspective of some observer or user or designer of the system — if we want original representation, naturalized representation — then representation must be capable of truth value for the system itself. Interactive representation *is* capable of truth value for the interactive system itself.

There are additional relevances of the criterion of system detectable error to representation. In particular, only if representational error is at least potentially *detectable* by a system is it possible for system activity to be *guided* by such error. Error guided system activity would include various kinds of goal directed or servomechanism activity — *and* it would include representational learning. Only with system detectable representational error can, in the general case, representation be learned: any "learning" that occurs without such possibility of error detection is pre-formed or pre-designed (Bickhard & Terveen, 1995). One of the central arguments of the skeptics is the inability to detect representational error (Burnyeat, 1983; Clay & Lehrer, 1989; Rescher, 1980) — any such representational check is a circular check of representations against themselves[1] — yet, if representational error is not detectable, then neither representational goal directedness nor representational learning are possible, whether in animals or in machines. System detectable error, thus, is a fundamental criterion for an acceptable model of representation.

Contemporary Approaches to Representation

Covariational Approaches. System detectable error is a criterion that can *not* be satisfied by standard approaches to representation in terms of informational covariances — correspondences — between representing states and that which is to be represented (Dretske, 1981; Fodor, 1990b; Hanson, 1990). In fact, one problem that emerges for standard approaches is that of the possibility of representational error *at all*, setting aside any issues of the *system detectability* of representational error. If the representationally constitutive correspondence exists, then the representation exists, and it is correct. If the constitutive correspondence does not exist, however, then the representation does not exist, and so it cannot be incorrect.

There are a number of proposals in the literature attempting to deal with this narrower problem of the possibility of error. One is the asymmetric dependency proposal (Fodor, 1987, 1990a; Loewer & Rey, 1991). The core intuition here is that the possibility of mistaken representations is in some sense *dependent* on the possibility of correct representations; they are parasitic. Explicating that sense of parasitic-ness is the aim of asymmetric dependency. If we consider a representation COW that is supposed to represent cows, then the error problem shows up if COW is evoked in conditions of, say, a horse on a dark night. We want to say that the horse induced evocation is in error. But there *is* a correspondence there, even with the horse, and, if we take it as constitutive of representation at all, then what is to block the conclusion that COW actually represents the disjunction "cow *or* horse on a dark night"? This version of the error problem is called the disjunction problem.

The asymmetric dependency proposal points out that if COW is a representation of cows, then the possibility of any horse evocations of it will be dependent on the possibility of cow evocations of it, and will be dependent in a non-reciprocated way. That is, cows will evoke COW even if horses never do, but horses, even on dark nights, will not evoke COW unless cows do. Evocations by cows, then, is privileged in that all errorful evocations are dependent on them, but they are not reciprocally dependent on the errorful evocations. The proposal is that such asymmetric dependence provides a (partial) criterion for correct versus incorrect representation.

---

[1] Or against other representations, which simply spreads the circularity out a little.

I will not attend to details of asymmetric dependency and its vicissitudes (e.g., Loewer & Rey, 1991), nor to other approaches to the error problem, because, even if they were agreed upon as succeeding in their aims, they at best establish a notion of error for some observer of the scene — an observer that already has representational access to the COW tokens, to cows in the environment, and to horses in the environment. It is only some such observer that could evaluate asymmetric dependency, for example, to determine whether some particular token of COW was in error or not in error — could determine whether or not the COW token had been evoked in an asymmetrically dependent manner. The system in which COW was evoked is not in such a position.

Asymmetric dependence, in its own terms, does not address the criterion of system detectable error. Therefore, it does not address the problem of original representation for the system itself. In fact, asymmetric dependence does not even distinguish between representational error and functional error, and so it *can* not address system detectable representational error.

Two counterexamples help bring home these points: Consider a transmitter molecule docking on a receptor in a cell surface and triggering internal activities in the cell, internal functional activities that are the products of evolutionary adaptation. Here we have an evocation of internal states that is in full covariational information correspondence with external conditions — whatever conditions yielded the release of the transmitter. But there is only a functional story to be told here, not a representational story: the internal activities of the cell do not represent *for the cell* any of those external conditions that might in fact be being covariationally tracked.

Consider next a poison molecule that mimics that transmitter molecule. It docks on the receptor and triggers internal activity. Furthermore, the poison molecule's ability to trigger that activity is dependent upon the ability of the transmitter molecule being able to trigger that activity, and that dependence is asymmetric. But there is still only a functional story to be told here.

There are a host of additional problems for standard informational or covariational approaches to representation: too many correspondences; wide and narrow content; how could representation be emergent, either in evolution or cosmology; and so on. Elsewhere I argue that they are all red herrings: they exist only because of the attempt to render representation in terms of correspondences (Bickhard, 1991b, 1993; Bickhard & Terveen, 1995).

Functional Approaches to Representation. The same general failure to address system detectable representational error holds for rival proposals (Dretske, 1988; Loewer & Rey, 1991; Millikan, 1984, 1993). In these models, X is supposed to represent Y if it is the function of X — or of X's — to represent Y, or Y's, and *function*, in turn, is modeled in terms of the learning and evolutionary histories of X's and the types of systems in which they occur. In these cases, the required assessments for the determination of error are of, among other things, various learning and evolutionary histories of the systems and their purported representations, rather than of asymmetric dependencies or the lack thereof. But those assessments of histories too are not possible for a system itself. Does a frog — could a frog — know anything at all about the evolutionary history of its internal representations of flies? Could a frog compare such a history of frog-fly representations with a current evocation of a fly representation by a pebble to determine that the current evocation is in error? Insofar as representational error, and thus representation, is dependent on such comparisons, then these models similarly yield the conclusion that representation is not possible for a system itself.

System Detectable Error is Not Possible in Contemporary Approaches to Representation. No system, animal or machine, can compare what an occurrent internal correspondence representation is *supposed* to be representing with what it is *actually* representing to determine if it is in error. No system, animal or machine, can assess the asymmetric dependency, or lack thereof, of an occurrent representation to determine whether or not it is in error. No system, animal or machine, can compare its own learning or evolutionary history to an occurrent representation to determine whether or not the occurrent representation is in error.

The required assessments and comparisons are question begging in that they require the system to compare what the “representation” is occurrently *actually* “representing” with what it is *supposed* to be representing — whether in terms of dependencies or histories — in order to detect error. But determining what is being occurrently *actually* represented is precisely the problem to be addressed: if there is no possibility of determining what is actually being represented independently of the relevant dependencies and histories, then there is no possibility of engaging in the required comparisons with those dependencies and histories (setting aside the impossible requirement for the system to have considerable knowledge of such dependencies and histories, even for frog representations). If it is not possible for the system, animal or machine, to engage in those comparisons, then there is no possibility of the system detecting error. In such models, therefore, there is no possibility of original representation — representation for the system itself.

These requirements in such models for representations that are independent of — external to — the system being modeled are obscured by the models being framed within the perspective of an external observer on both the system and its environment. In such a case that observer *does* have such an independent epistemic perspective on the occurrent representations of the system *and* on what in the environment is actually being represented. Such an external observer could, at least in principle, assess and compare occurrent representations with dependencies and histories. But no system can be in a position of being such an external observer on itself. Interactive representation does not require such external perspectives; interactive representation does permit system detectable representational error.

Interactive System Organization

An interactive system must in general have some way of indicating the *possibilities* of various interactions that is distinct from initiating *engagement* in those interactions. This is so because there will in general be more than one possible interaction available in a particular environment, and that availability must be indicated so that a selection can be made of which interaction to initiate. A frog that sees a nearby fly, for example, has available the possibility of flicking its tongue and eating, but if it also sees the shadow of a hawk, it is likely to avail itself of the possibility of jumping into the water instead. Human beings have available at any moment myriads of possible interactions.

The selection of which interaction to engage in should, in general, occur in terms of anticipated outcomes of the indicated interactions. So the system must, in addition to indications of interaction *potentialities*, have indications of anticipated or anticipatable interaction *outcomes*.

I will be modeling representation in terms of these two forms of indication — indications of interactive potentialities and indications of consequent outcomes of those interactions. If either one of these forms of indication must be realized in a manner that requires representation, then the approach being explored is doomed to circularity — to representation being modeled in terms of representation. The proximate task, then, is to model how these two forms of indication could occur in a strictly functional manner, without presupposing the representationality that is being modeled.

Indications of Interactive Potentiality

What must be indicated are the potentialities of interaction types, not the details of interaction tokens. The types can be specified as tightly as needed by the system, but the details may be both irrelevant and dependent on as yet undetected characteristics of the environment. Interaction types are easily specified by the functional or control structure organizations that would engage in those interactions, should the system select them. Interaction types, then, can be indicated by indicating subsystem organizations, like subroutines or servomechanisms.

How, then, can system components be indicated? The simple answer is with a pointer. A collection of pointers in a privileged location that point to subsystems will suffice to indicate the interactions that would be engaged in by those subsystems as currently available. This is not the only way to model this function of indication, but it suffices for current purposes: all I need is some way to model such indication that does not commit a representational circularity.

With regard to the function of selecting an interaction (type) to engage in, we can simply presuppose that the selection always takes place within the set of possibilities being pointed to. This is a simple functional restriction, and poses no circularity problems. There remains the problem of how interaction outcomes can be indicated without representational circularity, which I will address immediately below, and the interesting question of how the pointers to possibilities — the indications of interactive potentialities — get set up and updated over time in the first place. That question I will defer till later in the paper.

Indications of Interaction Outcomes

If the interaction outcomes to be indicated are external outcomes, then they must be represented, and the fatal circularity is upon us. If those outcomes are strictly internal, however, this circularity does not necessarily occur. In particular, if the indicated outcomes are themselves possible internal functional states of the interactive system, then:

1) those states can be pointed to as possibilities without being represented,

2) those states can be associated with the interaction types that might yield them, again via pointers and without representation,

3) interaction types can be functionally selected on the basis of the indicated internal outcomes, without representation being required,

4) the error or lack of error of such an outcome indication is constituted by the system either entering that internal functional state or not entering it, and

5) in either case, that is a functional fact in the system, available to influence further processing in the system, without anything being represented.

There is no representational circularity in any of these functional relationships.

Presupposition and Implicitness

An indication that some interaction type is currently possible and that it will yield one of some indicated set of possible internal outcomes is an indication that may hold or may fail. In order for it to hold, the environment must possess some set of properties of response to the interaction subsystem that will support the course of the overall interaction achieving the indicated internal outcome.

There is nothing in the indication per se that explicitly represents what those environmental interactive properties are. They are the truth conditions of the indication, but they are not explicitly represented truth conditions. Instead, they are functional presuppositions of the interaction-outcome indications. In that sense, they are *implicitly* represented, rather than explicitly.

This presuppositional implicitness is the basic form of interactive representation. It constitutes a kind of skill intentionality or praxeological intentionality. It is quite different from the standard explicit representations of symbols in a symbol manipulation system.

The Adequacy of Interactive Representation

Interactive implicit representation is different enough from standard conceptions of representations that it poses a host of questions about the adequacy of interactive representation. Even if the basic interactive model is accepted as an explication of some form of representation, how could it possibly handle ... and there follows a rather long list: Physical objects? Abstractions, such as electron or number? Perception? Imagery? Memory for events? Language? And so on.

Objects. I cannot tackle that programmatic set of questions here (Bickhard, 1980, 1987, 1991, 1992a, 1992c, 1993; Bickhard & Campbell, 1992; Bickhard & Richie, 1983; Campbell & Bickhard, 1986, 1992), but I will adumbrate approaches to two of the questions in order to show that there are such approaches. Objects first. If some internal outcome to an interaction were to be obtained, that functional state may, in turn, indicate the possibility of still further interactions, with their own potential outcomes. Indications of interactions and outcomes, in other words, can branch and iterate. In branching and iterating, there emerges the possibility of nets of conditional indications, and, in particular, subnets that close on themselves. This would be the case if some class of potential interactions and outcomes all indicate the collective possibility of the entire class — the indications are closed within such a class.

Furthermore, such a closed organization of indications may itself remain invariant as a structure of interaction possibilities under various interactions. For example, a toy block offers many interaction possibilities to a child, ranging from visual scans, to manipulations, to throwings, to droppings, to chewings, and so on. Further, any one of these indicates the availability of all the rest — it is closed under the indicated interactions. And still further, the entire structure of possibilities remains invariant under many kinds of interactions, such as manipulations, locomotions of the child, storing in the toy box, and so on. It does not remain invariant, however, under burning, crushing, and so on.

The basic proposal is that object representation in its most primitive form is just such invariances of closed subnets of interactive indications. This is, clearly, a basically Piagetian notion of object representation (Piaget, 1954). Piaget, in fact, participates in the general pragmatic approach (Piaget, 1971, 1977) — the descent is roughly Peirce to James to Baldwin to Piaget. I will not elaborate on this proposal nor defend it further; my basic point is made, that object representation does not present an aporia to a pragmatic approach to representation.

Numbers. What about abstractions, such as number? This is the second prima facie problem that I will address. The core insight here is to note that the properties of the interactive systems themselves are more abstract than the environments which those systems interact with. One property of an interaction type, for example, may be to iterate some subsystem, perhaps till some internal criterion is met. If that iteration occurs, say, three times, then the ordinal three is a property of the interaction that is not necessarily a property of whatever is being interacted with.

If there is a second level interactive system that can interact with the first level system organization, then that second level system could represent such properties as the three-ness of some iteration organization in the first level architecture. Any such second level system, in turn, will have properties that might be represented from a third level system, and so on. Here we have resources for abstractions that are unbounded — an unbounded hierarchy of levels of potential interactive system.

Again, there are many secondary questions that immediately arise: How many such levels might we find in human beings? How could an organism ascend such levels? And so on. Again, I will not pursue them here. I will note, however, that this approach to abstraction is not ad hoc: it converges with a model of developmental stages with its own empirical and logical support (Bickhard, 1992b, 1992d; Campbell & Bickhard, 1986). This too is generally Piagetian, though with stronger differences from Piaget than in the case of the model of object representations (Bickhard, 1988; Bickhard & Campbell, 1989).

These models of object and abstraction representations require much development. Here, however, I only want to block the superficial appearance of immediate aporia. The claim that interactive representation might be adequate to all forms of representation is at least still viable. For more detailed developments of the model, see, for example, regarding perception, Bickhard (1992a, 1992c; Bickhard & Richie, 1983), rationality, Bickhard (1991a, 1992a, in preparation-b; Hooker, 1995), language, Bickhard (1980, 1987, 1992a, 1992c, 1995, in preparation; Bickhard & Campbell, 1992; Bickhard & Terveen, 1995; Campbell & Bickhard, 1992), contemporary artificial intelligence and cognitive science, Bickhard & Terveen (1995), development, Bickhard (1980b, 1988, 1991b, 1992a, 1992b, 1992d; Bickhard & Campbell, 1989, in preparation; Bickhard & Christopher, 1994; Campbell & Bickhard, 1986), personality, Bickhard (1989; Bickhard & Christopher, 1994), and the nature of persons more broadly, Bickhard (1980b, 1992a, 1992b, in preparation).

Functions

I have modeled representation in functional terms. Representational error emerges as a particular kind of functional error, and, therefore, representation emerges as a particular kind of function — the function of indicating possibilities of further interactive process. I claim that this model of representation has many virtues, among which is the possibility of system detectable error, and, therefore, of system error guided processes such as goal-directedness and learning. The notion of function, however, poses its own problems.

Function, and, therefore, the distinction between function and dysfunction, is commonly modeled in terms of various learning and evolutionary histories (Dretske, 1988; Millikan, 1984, 1993). If function were dependent on such histories, then the interactive model of representation would require assessment of such histories, and comparisons with those histories, in order to determine function and dysfunction, and, therefore, to determine representational error and lack of error — as do Dretske's and Millikan's models of representation, and for similar reasons. This would violate the criterion of system detectable error.

In order to satisfy the criterion of system detectable *representational* error, then, the interactive model of representation requires a model of function that satisfies the criterion of system detectable *functional* error. Intrinsic dependence of function on history violates this criterion.

The issues regarding function are complex, and the literature on function is extensive. I will not attempt here an exhaustive treatment of function and its embedding in the current literature. Instead, I will outline a framework for the modeling of function that is plausible, and that clearly satisfies the criterion of system detectability of error. In particular, I will outline an approach to function that is not historical, though it does have important and strong connections to issues of history (Christensen, 1995, forthcoming; Bickhard, 1993; Bickhard & Terveen, 1995; Hooker & Christensen, in preparation).

The central criterion for a model of function is to distinguish between function and dysfunction in a way that is strictly natural, not dependent on any external ascriptions. Functions are consequences (Wimsatt, 1972), so some basis must be modeled for distinguishing consequences that are functional from those that are not, and for asymmetrically distinguishing between the successful yielding of those consequences and the failure to yield those consequences.

The paradigm domain for intuitions about functions is the biological domain. We want to be able to explain the sense in which hearts and kidneys and so on typically serve functions. The intuition, of course, is that they contribute to the continuation of the life of the organism, or of the species. It has proven remarkably difficult, however, to explicate these paradigm cases (Bechtel, 1986; Bigelow & Pargetter, 1987; Block, 1980, 1980b; Boorse, 1976; Cummins, 1975; Neander, 1991; Wimsatt, 1972, 1976; Wright, 1973). Current approaches present models of function that are dependent on the evolutionary and learning histories of the systems involved, but this would make the differentiation between serving a function and failing to serve that function a matter of comparison between current process and past history — a comparison that systems, in general, are not in a position to make. Such models of function, therefore, at best explicate the ascription of function by an external observer of the organism and its species. This does not suffice as a naturalistic model of function. Nevertheless, the intuition of functions being consequences that contribute to the survival of the system can be maintained without adverting to history.

Consider a process that is far from thermodynamic equilibrium, such as a flame. The continued existence of such a process is dependent on continuous interchange with the environment. The flame is intrinsically an open system: if it is cut off from the environment, the interchange, and, therefore, the process, ceases. This constitutes a fundamental difference with some other processes — the dance of nucleons and electrons that constitutes a water molecule, for example — which continue in existence even if isolated from their environments.

Some open systems — for example, a chemical bath with continuous flows of active agents into the bath — are completely dependent on the external persistence of necessary environmental conditions and processes for the continued persistence of the open system itself. In the example of the chemical bath, something must maintain the inflows of chemical agents. Other open systems, such as a flame, generate properties and processes inherent to the open system process that contribute to the maintenance of the open system process itself. In the case of the flame, the combustion process generates heat, which contributes to the maintenance of the necessary condition of above combustion threshold temperature for the flame to continue. In a normal atmosphere and gravitational field, this also contributes to convection, and thereby to the maintenance of the presence of oxygen — another condition necessary for the maintenance of the flame.

Whether or not the flame process continues is a natural phenomenon with natural consequences. Neither the continuation of the flame, or lack thereof, nor any of its consequences are dependent on any observer assessment or ascription. Similarly, the contributions that the flame makes to its own continued existence — or the failure to make those contributions — are natural phenomena with natural consequences.

The intuition of "function" that I propose is that system processes, or subprocesses, serve a function relative to an open system insofar as they contribute to the continued existence of that open system. This is an occurrent notion of function that is not dependent on any history. A lion that just popped into existence, then, would on this view have a heart that serves a function (cf. Millikan, 1984).

This intuition provides only a minimum outline of an approach to modeling function: a distinction between consequences that are functional and those that are not, and an asymmetric distinction between serving a function and failing to do so. I am leaving many central issues unaddressed here.

Perhaps most important among them is the distinction between "serving a function" and "having a function". A particular kidney may have a function even if it does not and never did serve that function. The property of having a function is one that a particular subsystem bears in virtue of its membership of some class of similar subsystems which, as a class, inherit that property. The move from "serving a function" to "having a function" requires moving from the singular case to the typical case, with many issues along the way. These are explored a little further elsewhere (Bickhard, 1993, in preparation).

In current approaches, "having a function" is taken as the primary notion to model, with "serving a function" or failing to do so on the part of a singular subsystem being derivative. The property of having a function, in turn, is modeled in terms of the learning and evolutionary history of the type of element or subsystem involved. It is this focus on the history of types as primary that renders these models non-natural — history cannot have naturalistic current consequences except in terms of current processes. By reversing the order of relationship between "serving a function" and "having a function" — by making the occurrent notion of "serving a function" primary — I intend to rescue the notion of function from being merely ascriptive on the part of some observer who can assess such histories.

More broadly, I assume that functions *can* be naturalized, whether or not in the manner in which I have outlined. The model of interactive representation, then, is a natural model insofar as the functional notions of pointers and so on — and their many architecturally different functional alternatives (Bickhard & Terveen, 1995) — are natural. That cannot be, however, by virtue of their histories, on pain of non-naturalism.

The dependence of open systems on the persistence of far-from-equilibrium conditions necessary for the continued existence of open systems provides a framework for making the necessary distinctions.

Avoiding Other Aporias

The Disjunction Problem. I have organized this presentation of the interactive model of representation around the criterion of system detectable error, and argued that alternative modeling approaches do not even address this criterion. Insofar as interactive representation does make sense of system detectable error, then it certainly accounts for representational error per se. In particular, the disjunction problem does not arise for interactive representation. The disjunction problem is a red herring produced solely by attempting to account for representation as correspondence. Similarly, other problematics to be found regarding standard approaches to understanding representation also do not arise for the interactive model.

Too Many Correspondences. Another serious problem for correspondence approaches to representation is that there are far too many correspondences in the universe — every instance of every lawful relation, for example, is also an instance of a correspondence — and most of them, at least, are not representational at all. Furthermore, even if attention is restricted to correspondences with mental states, and even to causally induced such correspondences, we still have that any mental element in correspondence with a table in front of an organism's eyes will also be in correspondence with retinal chemical activities, light patterns, interactions between light and electrons in the surface of the table, the presence of the table at that position yesterday, the movement of the table to that position whenever that occurred, the manufacture of the table, the production of the raw materials for the table, the creation of the raw materials for the table in an ancient supernova, and so on all the way back to the Big Bang. There are still too many correspondences (Coffa, 1991). How is the organism, or machine, supposed to determine which of these correspondences is supposed to be the representational one? A common answer is in terms of some sort of further functioning of the system that is "appropriate" to one privileged such correspondence (Bogdan, 1988a, 1988b, 1989; B. C. Smith, 1985, 1987). But even this move could at best pick out one among a host of alternative representational contents, one for each other end of the myriad of correspodences, and so it presupposes the existence of prior representations for the other end of the correspondences. This is circular.

Again, the problem of too many correspondences is a product solely of attempting to model representation in terms of correspondences. The problem simply does not arise for interactive representation. The implicit definition of an interactive property specifies that property. Elements or events in correspondence do not specify what they are in correspondence with, but, again, such representational specification is precisely the problem of representation to be addressed.

Wide and Narrow Contents. If the content of a (correspondence) representation is whatever is on the other end of the correspondence, then Twin Earth arguments show that conditions internal to an epistemic system cannot uniquely specify that external, or wide, content. It is always possible that all internal conditions would be the same, but the environment different such that the correspondences were with something different (Fodor, 1987, 1990a; Loewer & Rey, 1991). This produces problems for notions that representational content ought to be functionally or causally efficacious in the mental processes of the organisms and systems involved. Such efficaciousness seems blocked if the content is external, not internal, and can never be uniquely specified internally.

Something internal is necessary to play a role in at least partially specifying the external or wide content — commonly dubbed narrow content — and it is therefore available for influencing internal processes of the system. But narrow content cannot do what is needed because it does not specify uniquely what it might be in correspondence with. There is, in fact, no model available for how narrow content could specify anything at all about what its wide content might be in a way that could be functionally efficacious.

For a third time, this problem never arises for interactive representation. The content of an interative representation is implicit in the organization of the interactive system. It is internal, and is thus available for influencing internal processing. There is no mystery about the functional efficacy of representational content in this model.

An interactive representation does not specify uniquely what it is a representation of, but, for interactive representation, this is part of its strength, not a weakness of the model. Interactive representation is emergent in implicit definitions and differentiations, and on indicated relationships between them. Interactive representation is not built up out of particulars such as sense data, but is a process of differentiating the environment in ways that are relevant to further interactive possibilities for the system. Differentiations are intrinsically open and underspecify what they differentiate.

Emergence. One of the many scandals of encoding notions of representation is that there is no account of original representational content. Encodings can be defined in terms of other representations, including other encodings, but foundational encodings cannot be defined within the constraints of a strict encoding modeling approach. There is no way to model the nature or origin of the representational contents that would make foundational encodings into encodings at all.

A partial recognition of this is to posit that all grounding encodings must be innate, since we cannot account for their learning or development (Bickhard, 1991b, 1993; Fodor, 1975, 1981). But, if the problem is that encodingism cannot account for the origin of grounding encoding representational contents at all, as a logical matter, then evolution cannot generate representational content either. In fact, unless this logical aporia of the impossibility of emergence of representational content is somehow dissolved, it is impossible for representation to have emerged at any time and in any way since the Big Bang. Since it is fairly clear that no representations existed at the moment of the Big Bang, it follows that representation is not possible at all. If something cannot come into existence, then it cannot exist. This is a clear reductio — something has to be wrong.

I claim that what is wrong is the assumption that representation is a species of correspondence. Correspondence does not announce what it is in correspondence with, and no restrictions to sub-classes of correspondences (e.g., correspondences that are causally induced, that are followed by particular functional processes, and so on) can solve that problem. Correspondence models make the emergence of representation impossible. Interactive representation, on the other hand, emerges with complete naturalism out of certain sorts of functional organizations. There is no mystery of representational origin in this model.

Furthermore, for any system, biological or otherwise, of sufficient complexity that selections of further interactions cannot be simply triggered by current inputs, but must be made on the basis of anticipations of the further consequences of those interactions, interactive representation serves a clear function. So, not only is the *possibility* of the emergence of interactive representation clear, so also is the explanation for the *actual* emergence of interactive representation in living beings, and in machines (Bickhard & Terveen, 1995).

What about Input Encodings?

It would seem that the pervasive conception of representation as encodings of inputs, such as sensory encodings (Carlson, 1986), could not be completely wrong: how could it succeed empirically as well as it does? In other words, how does the interactive approach save the empirical results that are normally taken as supportive of the standard conceptions of representations?

The answer to this question brings us back to the earlier question of how indications of interactive potentialities are set up. Sensory "encodings" are correspondences that occur between internal states, such as firing rates on particular axons, and external conditions, such as visual or auditory inputs. Classes of such input patterns may be differentiated by neural processing of various sorts, such as might differentiate conditions in which there is a fly present from conditions in which there is not a fly present. The results of such processing, in turn, create correspondences with flies.

So far, so good. But at this point, the standard interpretation is that those internal conditions somehow *represent* that fly to the frog, and we enter into the myriad aporetic red herrings. The factual correspondences are interpreted as representational, encoding, correspondences (Coffa, 1991; Fodor & Pylyshyn, 1981; Hanson, 1990; Pylyshyn, 1984). In contrast, all that the interactive model needs is that those conditions that are in fact in correspondence with fly conditions set up pointers indicating the potentiality of tongue-flicking-and-eating. Indications of interactive potentialities need to be sensitive to environmental conditions, but that sensitivity need only be a factual sensitivity, such as informational or functional sensitivity — precisely what we actually find — not a representational sensitivity. What are normally taken as *being* representations are in the interactive model taken as constituting the functional conditions under which representational indications are set up. And, correspondingly, what is represented is not the other ends of the correspondences, not the detections or differentiations from processing inputs, but the interactive potentialities that might follow from those functional detections. The frog represents potentialities for tongue-flicking-and-eating, not flies.

Connectionism

To this point, I have not mentioned connectionism. How does it fare with respect to these issues? A net is trained to detect instances of classes of input patterns (Bickhard & Terveen, 1995; Churchland, 1989; Clark, 1989, 1993; Horgan & Tienson, 1988; McClelland & Rumelhart, 1986; Rumelhart, 1989; Rumelhart & McClelland, 1986; Smolensky, 1986, 1988; Waltz & Feldman, 1988). A net performs the sorts of functions of detection or differentiation based on inputs that we find in the sensory systems (the full sensory case, however, can be more complicated than a typical net: Bickhard & Richie, 1983). In classical approaches, such detection is presumed to occur via transducers: transducers encode input categories, and make those encodings available for further processing (Fodor & Pylyshyn, 1981; see Bickhard & Richie, 1983, for a critique). In both the neural net and the classical case, however, what actually occurs is a differentiation of instances of classes of input patterns, and, in both cases, such differentiations are construed as being representations. Transducers are evolved or engineered, while connectionist nets can be trained, but what they end up doing is the same sort of task, and is subject to the same sort of misinterpretation as constituting representation. Neither transducers nor nets, however, are capable of system detectable error concerning what they take to be on the other end of their input correspondences — neither one takes *anything* to be on the other end of its inputs. Connectionism, in other words, does not address this basic problematic of representation *for the system* either.

The interactive model, in contrast, needs exactly such differentiators — whether transducers or nets or hybrids or whatever — in order to set up its interactive indications successfully. Transducers and nets are functional for the setting up of representations, but they are not representations per se — not for the system itself.

Conclusion

System detectable error is a meta-epistemological criterion that current informational approaches to representation fail. Interactive representation naturally satisfies this criterion, as well as that of error per se, of the possibility of emergence, and other such criteria (Bickhard, 1993; Bickhard & Terveen, 1995). Further, interactive representation manifests the possibility of being able to account for other prima facie problematic forms of representation, such as of objects and numbers, and, therefore, shows a programmatic possibility of being the fundamental form of all representation.

References

Bechtel, W. (1986). Teleological functional analyses and the hierarchical organization of nature. In N. Rescher (Ed.) *Current Issues in Teleology*. Landham, MD: University Press of America, 26-48.

Bickhard, M. H. (1980). *Cognition, Convention, and Communication.* New York: Praeger.

Bickhard, M. H. (1980b). A Model of Developmental and Psychological Processes. *Genetic Psychology Monographs, 102*, 61-116.

Bickhard, M. H. (1987). The Social Nature of the Functional Nature of Language. In M. Hickmann (Ed.) *Social and Functional Approaches to Language and Thought* (pp. 39-65). New York: Academic.

Bickhard, M. H. (1988). Piaget on Variation and Selection Models: Structuralism, Logical Necessity, and Interactivism *Human Development, 31*, 274-312.

Bickhard, M. H. (1989). The Nature of Psychopathology. In L. Simek-Downing (Ed.) *International Psychotherapy: Theories, Research, and Cross-Cultural Implications*. (115-140). New York: Praeger.

Bickhard, M. H. (1991a). A Pre-Logical Model of Rationality. In L. Steffe (Ed.) *Epistemological Foundations of Mathematical Experience* New York: Springer-Verlag, 68-77.

Bickhard, M. H. (1991b). The Import of Fodor's Anticonstructivist Arguments. In L. Steffe (Ed.) *Epistemological Foundations of Mathematical Experience*. New York: Springer-Verlag, 14-25.

Bickhard, M. H. (1992a). How Does the Environment Affect the Person? In L. T. Winegar, J. Valsiner (Eds.) *Children's Development within Social Contexts: Metatheory and Theory.* Erlbaum, 63-92.

Bickhard, M. H. (1992b). Scaffolding and Self Scaffolding: Central Aspects of Development. In L. T. Winegar, J. Valsiner (Eds.) *Children's Development within Social Contexts: Research and Methodology.* Erlbaum, 33-52.

Bickhard, M. H. (1992c). Levels of Representationality. *Conference on The Science of Cognition*. Santa Fe, New Mexico, June 15-18.

Bickhard, M. H. (1992d). Commentary on the Age 4 Transition. *Human Development*, 182-192.

Bickhard, M. H. (1993). Representational Content in Humans and Machines. *Journal of Experimental and Theoretical Artificial Intelligence*, *5*, 285-333.

Bickhard, M. H. (1995). Intrinsic Constraints on Language: Grammar and Hermeneutics. *Journal of Pragmatics, 23*, 541-554.

Bickhard, M. H. (in preparation). *The Whole Person: Toward a Naturalism of Persons.* Harvard.

Bickhard, M. H. (in preparation-b). Critical Principles: On the Negative Side of Rationality. In Herfel, W., Hooker, C. A. (Eds.) *Beyond Ruling Reason: Non-formal Approaches to Rationality.*

Bickhard, M. H., Campbell, R. L. (1989). Interactivism and Genetic Epistemology. *Archives de Psychologie, 57*(221), 99-121.

Bickhard, M. H., Campbell, R. L. (1992). Some Foundational Questions Concerning Language Studies: With a Focus on Categorial Grammars and Model Theoretic Possible Worlds Semantics. *Journal of Pragmatics, 17(5/6),* 401-433.

Bickhard, M. H., Campbell, R. L. (in preparation). Topologies of Learning and Development.

Bickhard, M. H., Christopher, J. C. (1994). The Influence of Early Experience on Personality Development. *New Ideas in Psychology*, *12*(3), 229-252.

Bickhard, M. H., Richie, D. M. (1983). *On the Nature of Representation: A Case Study of James J. Gibson's Theory of Perception*. New York: Praeger.

Bickhard, M. H., Terveen, L. (1995). *Foundational Issues in Artificial Intelligence and Cognitive Science — Impasse and Solution*. Amsterdam: Elsevier Scientific.

Bigelow, J., Pargetter, R. (1987). Functions. *Journal of Philosophy, 84*, 181-196.

Block, N. (1980). Introduction: What is functionalism? In N. Block (Ed.), *Readings in philosophy and psychology* (Vol. I). (171-184). Cambridge: Harvard.

Block, N. (1980b). Troubles with functionalism. In N. Block (Ed.) *Readings in philosophy and psychology* (Vol. I). (285-305). Cambridge: Harvard.

Bogdan, R. (1988a). Information and Semantic Cognition: An Ontological Account. *Mind and Language, 3*(2), 81-122.

Bogdan, R. (1988b). Mental Attitudes and Common Sense Psychology. *Nous, 22*(3), 369-398.

Bogdan, R. (1989). What do we need concepts for? *Mind and Language, 4*(1,2), 17-23.

Boorse, C. (1976). Wright on Functions. *Philosophical Review, 85*, 70-86.

Burnyeat, M. (1983). *The Skeptical Tradition*. Berkeley: University of California Press.

Campbell, R. L., Bickhard, M. H. (1986). *Knowing Levels and Developmental Stages.* Basel: Karger.

Campbell, R. L., Bickhard, M. H. (1992). Clearing the Ground: Foundational Questions Once Again. *Journal of Pragmatics*, *17(5/6),* 557-602.

Carlson, N. R. (1986). *Physiology of Behavior*. Boston: Allyn and Bacon.

Christensen, W. (1995). From Descriptive to Normative Functionality. Australasian Association of Philosophy Conference, **

Christensen, W. (forthcoming). A Complex Systems Theory of Teleology. *Biology and Philosophy*.

Churchland, P. M. (1989). *A Neurocomputational Perspective*. MIT.

Clark, A. (1989). *Microcognition*. MIT.

Clark, A. (1993). *Associative Engines*. MIT.

Clay, M., Lehrer, K. (1989). *Knowledge and Skepticism*. Westview.

Coffa, J. A. (1991). *The Semantic Tradition from Kant to Carnap*. Cambridge.

Cummins, R. (1975). Functional Analysis. *Journal of Philosophy, 72*, 741-764.

Dretske, F. I. (1981). *Knowledge and the Flow of Information*. Cambridge, MA: MIT.

Dretske, F. I. (1988). *Explaining Behavior*. MIT.

Dreyfus, H. L. (1967). Why Computers Must Have Bodies in order to be Intelligent. *Review of Metaphysics, 21*, 13-32.

Dreyfus, H. L. (1982). Introduction. In H. L. Dreyfus (Ed.) *Husserl: Intentionality & Cognitive Science*. (1-27). MIT.

Dreyfus, H. L. (1991). *Being-in-the-World*. MIT.

Dreyfus, H. L., Haugeland, J. (1978). Husserl and Heidegger: Philosophy's Last Stand. In M. Murray (Ed.) *Heidegger & Modern Philosophy* (222-238). Yale University Press.

Fodor, J. A. (1975). *The Language of Thought*. New York: Crowell.

Fodor, J. A. (1981). The present status of the innateness controversy. In J. Fodor, *RePresentations* . Cambridge: MIT Press (pp. 257-316).

Fodor, J. A. (1987). *Psychosemantics*. Cambridge, MA: MIT Press.

Fodor, J. A. (1990a). *A Theory of Content and Other Essays*. MIT.

Fodor, J. A. (1990b). Information and Representation. In P. P. Hanson (Ed.) *Information, Language, and Cognition*. (175-190). University of British Columbia Press.

Fodor, J. A., Pylyshyn, Z. (1981). How direct is visual perception?: Some reflections on Gibson's ecological approach. *Cognition*, *9*, 139-196.

Guignon, C. B. (1983). *Heidegger and the Problem of Knowledge*. Indianapolis: Hackett.

Hanson, P. P. (1990). *Information, Language, and Cognition*. University of British Columbia Press.

Heidegger, M. (1962). *Being and Time*. New York: Harper & Row.

Hooker, C. A. (1995). *Reason, Regulation, and Realism: Towards a Regulatory Systems Theory of Reason and Evolutionary Epistemology*. SUNY.

Hooker, C. A., Christensen, W. (in preparation). Very Simple Minds.

Hoopes, J. (1991). *Peirce on Signs*. Chapel Hill.

Horgan, T., Tienson, J. (1988). Settling into a New Paradigm. The Spindel Conference 1987: Connectionism and the Philosophy of Mind. *The Southern Journal of Philosophy, XXVI*(supplement), 97-114.

Loewer, B., Rey, G. (1991). *Meaning in Mind: Fodor and his critics*. Blackwell.

McClelland, J. L., Rumelhart, D. E. (1986). *Parallel Distributed Processing: Explorations in the Microstructure of Cognition: Vol. 2 Psychological and Biological Models*. Cambridge: MIT.

Millikan, R. G. (1984). *Language, Thought, and Other Biological Categories*. MIT.

Millikan, R. G. (1993). *White Queen Psychology and Other Essays for Alice*. MIT.

Neander, K. (1991). Functions as Selected Effects: The Conceptual Analyst's Defense. *Philosophy of Science, 58*(2), 168-184.

Okrent, M. (1988). *Heidegger's Pragmatism*. Cornell.

Piaget, J. (1954). *The Construction of Reality in the Child*. New York: Basic.

Piaget, J. (1971). *Biology and Knowledge*. Chicago: University of Chicago Press.

Piaget, J. (1977). The Role of Action in the Development of Thinking. In W. F. Overton, J. M. Gallagher (Eds.) *Knowledge and Development: Vol. 1*. New York: Plenum.

Pylyshyn, Z. (1984). *Computation and Cognition*. MIT.

Rescher, N. (1980). *Scepticism*. Totowa, NJ: Rowman and Littlefield.

Rosenthal, S. B. (1983). Meaning as Habit: Some Systematic Implications of Peirce's Pragmatism. In E. Freeman (Ed.) *The Relevance of Charles Peirce*. Monist, 312-327.

Rumelhart, D. E. (1989). The Architecture of Mind: A Connectionist Approach. In M. I. Posner (Ed.) *Foundations of Cognitive Science*. MIT, 133-160.

Rumelhart, D. E., McClelland, J. L. (1986). *Parallel Distributed Processing. Vol. 1: Foundations*. MIT.

Smith, B. C. (1985). Prologue to "Reflections and Semantics in a Procedural Language" In R. J. Brachman, H. J. Levesque (Eds.) *Readings in Knowledge Representation*. (31-40). Los Altos, CA: Morgan Kaufmann.

Smith, B. C. (1987). *The Correspondence Continuum*. Stanford, CA: Center for the Study of Language and Information, CSLI-87-71.

Smith, J. E. (1987). The Reconception of Experience in Peirce, James, and Dewey. In R. S. Corrington, C. Hausman, T. M. Seebohm (Eds.) *Pragmatism Considers Phenomenology*. (73-91). Washington, D.C.: University Press.

Smolensky, P. (1986). Information Processing in Dynamical Systems: Foundations of Harmony Theory. In Rumelhart, D. E., McClelland, J. L. (Eds.) *Parallel Distributed Processing: Explorations in the Microstructure of Cognition. Vol. 1: Foundations*. Cambridge, MA: MIT, 194-281.

Smolensky, P. (1988). On the Proper Treatment of Connectionism. *Behavioral and Brain Sciences, 11*, 1-74.

Waltz, D., Feldman, J. A. (1988). Connectionist Models and Their Implications. In D. Waltz, J. A. Feldman (Eds.) *Connectionist Models and Their Implications*. Norwood, NJ: Ablex, 1-12.

Wimsatt, W. C. (1972). Teleology and the Logical Structure of Function Statements. *Studies in the History and Philosophy of Science, 3*, 1-80.

Wimsatt, W. C. (1976). Reductive Explanation: A functional account. In R. S. Cohen, C. A. Hooker, A. C. Michalos, J. Van Evra (Eds.) *PSA-1974. Boston Studies in the Philosophy of Science*. (Vol. 32, pp. 671-710). Dordrecht: Reidel.

Wright, L. (1973). Functions. *Philosophical Review, 82*, 139-168.