Is Cognition an Autonomous Subsystem?

Mark H. Bickhard

In standard approaches, cognition is assumed to be a distinct functional module from action and interaction, on the one hand, and motivation, on the other hand. I argue that the apparent possibility of such modularization is itself a consequence of a false underlying presupposition concerning the nature of representation. An alternative model of representation is outlined, and it is shown that action and motivation emerge naturally and necessarily as aspects of one single underlying ontology of the interactive agent.

In standard views, representation consists of elements in correspondence with what they represent. Only certain kinds of correspondences will do — not all correspondences are representational — but correspondence is the basic category within which representations are differentiated (B. C. Smith, 1987). The representational elements are taken to encode that which they are in correspondence with. This assumption regarding the nature of representation is held in common among those who consider such encodings to be transduced (Fodor & Pylyshyn, 1981) or innate (Fodor, 1981) or designed (Newell, 1980) or trained into a net (Churchland, 1989) — these positions differ regarding the presumed *origin* of encoding correspondences, not regarding their basic representational nature.

In such a view of representation, cognition is taken to consist of various stages of the input and processing, and sometimes the output, of such encodings. The fundamental backbone of cognition is assumed to be the sequence from perception to cognitive processing to the re-encoding into utterances in some language.

This view lends itself to, if not forces, a strong modularization of models of the mind. This modularization is *in addition* to the potential modularizations within cognition itself (Fodor, 1983). Cognition per se is seen as being organized around one or more storage banks for encodings. Various processes enter data into such banks, process the contents of such banks, and re-encode selected contents into utterances.

For any model of a real agent, with a real mind, additional modules are required. Some subsystem is required to control action. Another subsystem interprets the cognitive contents for their relevance to action. Another engages in the motivational selections of actions to be performed, again informed by the cognitive contents. Still further fragmentations of mentality are common, but I will focus in this paper on these three: representation, action, and motivation.

In particular, I will argue that the standard view of representation as some kind of correspondence, as an encoding, is wrong. I outline an alternative model of representation that emerges naturally in agents, biological or designed, that actually engage the world (Beer, 1990, 1995, in press; Beer, Chiel, Stirling, 1990; Bickhard, 1980, 1993; Bickhard & Terveen, 1995; Brooks, 1991a, 1991b, 1991c; Cherian & Troxell, in press; Malcolm, Smithers, Hallam, 1989; Smithers, 1994). One primary consequence of this alternative model of representation — called interactivism — is that functions that are standardly taken to reside in separate modules, such as representation, action, and motivation, are inherently integrated as separate functional *aspects* of one single underlying ontology. They are not inherently distinct modules. If standard models that permit such modularization are in error, then so are such modularizations per se.

Encoding Models of Representation.

Encoding models of the nature of representation are wrong. The deepest errors, however, are not perspicuous, and attempting to locate them can lead to explorations of endless mazes of blind alleys and fruitless pursuit of red herrings. The consequences are fatal to any aspirations of understanding mental processes, and can be devastating even to some strictly practical design goals in artificial intelligence (Bickhard & Terveen, 1995).

I will not be arguing that encodings do not exist. They clearly do: Morse code is a paradigm example. What I do argue, however, is that encodings cannot be the basic nature of representation. Genuine encodings must be derivative from some other, more fundamental, form of representation. If the assumption is made that the fundamental nature of representation is encodings (Palmer, 1978; cf. Bickhard & Terveen, 1995) — whether as an explicit assumption or a deeply buried implicit assumption — an incoherence results.

There is a large family of corollary arguments showing the incoherence of encodingism — of the assumption or presupposition that representation is encoding. I will outline only a few.

External forms of representation, such as Morse code or a statue or a blueprint, require interpretation and understanding. The dots and dashes of Morse code must be understood and interpreted into characters; the statue must be understood as a statue, and relevant resemblances noted and interpreted. This is unexceptional so long as the requirement for such an interpreter is unproblematic.

When attempting to account for the inner representations of such an interpreter, however — the inner mental representations of any real mind — presupposition of still another interpreter is lethal. The assumption that internal representations are of the same nature as external representations yields infamous regresses of interpretations requiring interpretations which require further interpretations, and so on, with the corresponding regress of interpretive homunculi to perform these feats. Whatever internal representation is, it cannot be the same kind of thing as external representation (Bickhard & Terveen, 1995; Clancey, 1991).

Considering Morse code in a different respect, we note that the dot and dash patterns of the code are stand-ins for the characters that they encode. They stand-in for those characters in the sense that we define them and use them and interpret them that way. It is useful to do so because, for example, dots and dashes can be sent over telegraph wires while characters cannot.

Such stand-in relationships are what constitute the Morse encodings as representations at all. The dot and dash patterns borrow their representational content — the specification of what they represent — from what they stand-in for. "Dot dot dot" represents the same thing as does "s". Again, this is not problematic so long as there is something to borrow representational content from.

But if we assume that all representations are encodings, are stand-ins, then we encounter an incoherence. The stand-in relationship can be iterated. **X** can be defined in terms of **Y,** which can be defined in terms of **Z**, but this iteration must be finite, so long as we are considering finite systems. Therefore, there must be a bottom level — a level of representations that do not stand-in for some other level.

If we assume that this level is constituted as encodings, however, we encounter the incoherence. Consider some presumed element, say "**X**", of this presumed grounding level of encodings. It cannot be defined in terms of any other representation by assumption. The only alternative is to define it in terms of itself, but that yields something like " "**X**" represents **X**" or " "**X**" stands-in for "**X**" ". This does not succeed in providing "**X**" with any representational content at all, and, therefore, does not succeed in constituting "**X**" as a representation at all. But there are, by the encodingism assumption and the assumption that "**X**" is an element of the grounding level of encodings, no further resources available for making "**X**" an encoding. Encodingism requires such a grounding level of encodings, so the encodingism assumption per se has yielded an incoherence: it assumes a grounding level of encoding elements, which cannot exist.

There are many other corollary arguments against encodingism (Bickhard, 1993; Bickhard & Terveen, 1995), but I will not pursue them here. The incoherence argument per se logically suffices to refute encodingism, but all I require for current purposes is a prima facie case that encodingism has serious problems — all I require is a motivation for considering an alternative model of representation.

Representation as Isomorphism. There are not only many corollary arguments against encodingism, which I will not pursue here, there are also many apparent rejoinders to the general critique of encodingism. For similar reasons of space, I cannot pursue most of those either (see Bickhard, 1993; Bickhard & Terveen, 1995), but there is one rejoinder that is inherent in one of the foundational frameworks for Artificial Intelligence — specifically, the Physical Symbol System Hypothesis — that I will address briefly.

Within the Physical Symbol System Hypothesis, symbols are defined in terms of pointers (*access*) to some other entity in the system (Newell, 1980; Bickhard & Terveen, 1995). This a strictly functional notion, strictly internal to a machine, and, as such, is unobjectionable. Problems emerge, however, when the attempt is made to extend the model to representation — in particular, to representation of external entities. The relationship between a symbol and what it is supposed to designate is fundamentally different in this external case — to construe that relationship in terms of pointer access is to presuppose what is to be explained. In particular, access is a primitive function built into a machine, but it is not primitive in the relationships between machine and the world. Representation *is*, in some sense, epistemic access, so access cannot be simply presupposed in a model of representation.

The alternative that is proposed is a relationship of the isomorphism of patterns: patterns in the world are what is designated, and they are designated by patterns in the system that are isomorphic to the designated external patterns (Newell, 1980; Vera & Simon, 1993). This is a richer concept than simple correspondence, but it, nevertheless, is still a version of encodingism, and is still logically unworkable.

Pattern isomorphy is still a version of correspondence — isomorphic, or structural, correspondence as well as single point to point correspondence. As such, it is subject to all of the problematics of encodingism:

- There is nothing about the internal pattern that carries knowledge of the fact of any such isomorphy, nor about what any such isomorphy might be with. There is no representational content specifying what the isomorphy is with — about the other end, the presumed represented end, of the isomorphic relationship.
- Isomorphy is multifarious: isomorphic correspondences can be defined ubiquitously, with almost anything. Which of these is the "representational" isomorphy?
- Isomorphy is transitively unbounded: any isomorphy with one pattern will also constitute an isomorphy with patterns causally prior to that one (as well as noncausal accidental isomorphisms all over the universe), and prior again, and so on. Which of *these* is the "representational" isomorphy?

- Isomorphy is transitive, but representation is *not* transitive: merely knowing the label of a map will not permit you to travel in the mapped territory — the label of the map does not represent the territory.
- Isomorphy is symmetric, but representation is not symmetric: the table does not represent your mental representation of the table.
- If representation is constituted as isomorphy, then, if the isomorphy exists, the representation exists, and it is correct. But if the isomorphy does not exist, then the representation does not exist, *and it cannot be incorrect*! Correspondence models, including isomorphy models, cannot account for representational error (Dretske, 1988; Fodor, 1987, 1990; Loewer & Rey, 1991; Millikan, 1984; B. C. Smith, 1987).
- Still more deeply, correspondence models, including isomorphy models, cannot account for the possibility of representational error (error of such correspondence) that is *detectable by the system itself*. But, without system detectable representational error, representational learning, among other error guided processes, is not possible (Bickhard & Terveen, 1995).

Isomorphy models of representation are versions of informational approaches to representation, of the assumption that the representational relationship is a version of the informational relationship (Fodor, 1987, 1990; Loewer & Rey, 1991; Hanson, 1990; B. C. Smith, 1987). An informational relationship, in turn, is just one version of a correspondence relationship. Informational approaches to representation are certainly the dominant approaches today, but, if the arguments against encodingism are correct, they are ultimately unworkable. It is clear that no one knows today how to make such an approach work: “we haven’t got a ghost of a Naturalistic theory about [encoding]” (Fodor, 1987b, p. 81) “The project of constructing a representational theory of the mind is among the most interesting that empirical science has ever proposed. But I’m afraid we’ve gone about it all wrong.” (Fodor, 1994, p. 113). The Physical Symbol System Hypothesis, then, does not provide a solution to the problems of encodingism. More generally, isomorphic correspondences are no improvement over correspondences per se in attempting to model representation.[1]

Interactivism.

For an agent interacting with its world, the ability to anticipate what interactions might be possible next would be a useful resource (Bickhard, 1993). Anticipations would permit the agent to select among those possibilities those that are most suited to its current internal conditions, or to select those that are most to be avoided. Such possibilities for further interaction will vary as the situation of the agent varies, so some process for constructing and updating the anticipations would be required.

The critical property of such anticipations for my current purposes is that they might be wrong, and might be discoverable to be wrong by the system itself: if the system engages in an indicated possible interaction, and the interaction fails — fails to proceed as indicated — then the anticipation was in error. Anticipations — indications of possible interactions — can be false, and can be discovered to be false by the system. Anticipations have truth values, for the system. Possessing truth values is at least one fundamental property of representations.

---

[1] I should point out that, although there are many important and useful properties of connectionist nets, the trained correspondences that are supposed to constitute representations in standard connectionism are no improvement over the designed or isomorphic correspondences that are supposed to constitute representations in GOFAI (Bickhard & Terveen, 1995).

I argue that these primitive truth values are in fact foundational to all representation. Indications of potential system interactions are the most primitive, the foundational, form of representation, out of which all other representation is constructed.

There are many aspects and promissory notes involved in this claim. I will address only two here:

1. Can this notion of anticipation be explicated in purely functional terms?
2. How could interactive anticipations account for such representations as those of objects?

Examples of other issues that I will *not* address here include: How could this notion of representation account for abstract representations, such as number? How could such a model of representation account for perception, for language? How could it be consistent with rational thought? And so on. These are all addressed elsewhere (Bickhard, 1980, 1993; Bickhard & Richie, 1983; Bickhard & Terveen, 1995: Hooker, 1995). For my current purposes, I need only a prima facie plausibility of interactive representation, not a demonstrated adequacy in all senses, because I am primarily aiming at the implications of such an interactive model of representation for issues of modularity. In particular, I will argue that, in the interactive model, issues of action and of motivation as action selection are most fundamentally intrinsic aspects of anticipatory interactive systems, not separate modules.

Functional Anticipations.

Rendering the necessary notion of anticipation in functional terms is not difficult: pointers and subroutines suffice. In particular, a pointer to a subroutine can indicate the potentiality of the interactions that would be engaged in by that subroutine, while further pointers to internal outcomes should that subroutine be in fact executed constitute the anticipations that are detectable by the system. If the system does engage that subroutine and the internal outcome of the interaction is not one of those indicated, then the indications have been falsified — and falsified for the system itself. There are other architectural frameworks in which the requisite anticipations can be modeled (Bickhard & Terveen, 1995), but demonstrating the adequacy of pointers and subroutines suffices to demonstrate that no non-functional notions are necessary.

Interactive anticipation yields the possibility of system detectable error. System detectable error, in turn, is necessary for error guided activities, such as goal directedness or learning. System detectable error, then, is a necessity for any but the most simple and primitive forms of life or artificial agents.

Representing Objects.

Primitive interactive anticipations have truth values, and, thus, constitute primitive forms of representation. They implicitly predicate to the environment whatever interactive properties would support the indications of internal outcomes (Bickhard, 1993). Such primitive representation, however, is appropriate primarily to simple organisms and simple artificial agents. More complex agents involve more complex representations, and that complexity must be accounted for.

There are two primary resources in the interactive model of representation that permit it to model complex representations, such as objects. These resources are conditional indications, and iterated indications.

All indications are conditional at least in the sense that they are evoked in certain internal system conditions, and not in others. That is, they are conditional on particular internal states of the system. In turn, interactions actually engaged in yield subsequent internal conditions as the internal outcomes of those interactions. This yields the possibility of iterated indications: interaction I_7 is possible given current system states and will yield outcomes O_1, O_2, or O_3, while, if O_1 is reached, then interactions I_{10}, I_{23}, and

I_{34} will be possible, which would yield outcomes O_{88}, ... and so on. Indications can branch and iterate, forming potentially complex webs of indicated interactive potentiality.

One possibility for such webs is that they might close into loops and other reciprocal indication relationships. A subweb might even exhibit the property of closure: all states in the web are reachable from all other states via some class of interactions that relate those states. With one additional property, I claim that we now have the necessary resources for modeling simple object representation. That additional property is invariance.

In particular, a typical object, say a child's toy block, offers many possible interactions — visual scans from various perspectives, multiple manipulations, dropping, chewing, and so on. Furthermore, every one of these possibilities indicates all the others, perhaps with necessary intervening interactions (for example, turning the object around to create the possibility of the visual scan from that angle). That is, the organization of the possibilities is closed. Still further, that overall pattern of possibilities, together with its closure, is invariant with respect to a large class of interactions. Clearly it is invariant with respect to each interaction in the web itself, but the interactive pattern of the block, for example, is also invariant under various throwings, locomotions of the infant, storing in the toy chest, and so on. It is *not* invariant, however, under burning, crushing, dissolving in acid, and so on.

Epistemically, objects just *are* closed invariant patterns of physical interaction. That's all they can be to infants and monkeys. Accounting for objects in terms of theories about objects, in terms of earth, air, fire, and water, for example, or in terms of atoms and molecules, is a much higher order accomplishment. Accounting for those higher order possibilities requires, among other things, addressing the representation of abstractions — something I will not undertake here.

The claim is that interactive representation is capable of accounting for representations of physical objects, in the generally Piagetian manner just outlined (Piaget, 1954). The concluding claims of this section, then, are that interactive representation is renderable in strictly functional terms, and that it has at least a prima facie initial plausibility of serving as an adequate approach to all representation.

Pragmatics and Representation.

Interactivism is a version of the general pragmatic approach to representation (Hookway, 1985; Houser & Kloesel, 1992; Rosenthal, 1983, 1990; Thayer, 1973). Such approaches share the assumption that representation is an emergent of action, and that classical correspondence, or "spectator", models of representation are inadequate (J. E. Smith, 1987). Although pragmatism has to date had relatively little influence on studies of cognition and artificial intelligence, there are exceptions. Most of these derive their pragmatist influences from Jean Piaget (influenced by James Baldwin, who, in turn, was influenced by Peirce and James). Interactivism shares a general pragmatism with such approaches, and shares a more specific influence from Piaget. The differences, consequently, are more subtle than they are with encoding models — all genuine pragmatist approaches *share* a rejection of correspondence models of representation (or at least an attempted rejection. In a number of cases, I claim that the attempt to avoid encodingist presuppositions has in fact not been fully successful). Within Artificial Intelligence, perhaps the best known of pragmatist approaches is the work of Drescher (1986, 1991), so I turn now to a brief comparison between interactivism and Drescher's model.

Drescher. Drescher's model of representation is essentially that of Piaget. There are strong commonalities between the interactive model and Piaget's model (Bickhard & Campbell, 1989), and, therefore, there are strong commonalities between the interactive model (Bickhard, 1980, 1993; Bickhard & Richie, 1983; Bickhard & Terveen, 1995) and

Drescher's (1986, 1991). Most of these commonalities follow from the basic framework of modeling representation as an emergent of action systems.

For example, if representation is construed as correspondence, there is a temptation to think that the world could impress itself into a passive but receptive mind, leaving behind representational correspondences. This is essentially Aristotle's model of perception, and is still with us in the technologically more sophisticated, but logically no better, notions of passive transduction and induction (Fodor & Pylyshyn, 1981; Bickhard & Richie, 1983). On the other hand, if representation is an emergent of action systems, there is no such temptation to think that action systems — interactively competent control structures — could possibly be impressed by the environment into a passive mind. If representation is emergent out of action, then perception and learning and development must all be active constructive processes (Bickhard, 1992). Furthermore, since such knowledge constructions cannot be assured of being correct — if they could, then the knowledge would already be present — they must be subject to being tried out and eliminated if they fail. That is, an action framework for understanding representation forces a variation and selection constructivism, an evolutionary epistemology (D. T. Campbell, 1974).

For another example, consider again just what it is that is most fundamentally being represented in an action based model. Representation is most fundamentally of future potentialities for further action and interaction. Pragmatic representation is future looking instead being backward looking down the stream of inputs coming into the organism — pragmatic models are not models of a spectator looking into the past of that input stream. In particular, pragmatic representation is intrinsically representation of *possibilities*. Pragmatic representation is intrinsically *modal*. This is one of many fundamental differences between pragmatic models and correspondence models (correspondence models, in fact, have in principle difficulties handling modal representation). It is worth noting that representation in children does not begin with strictly "actual" representation and then later add a layer of modality, as standard logic and encoding frameworks would suggest. Instead, children begin with representation that is intrinsically undifferentiated between various aspects of modality, and actuality, possibility, and necessity must be progressively differentiated in development over the course of some years (Bickhard, 1988; Piaget, 1986, 1987).

There are many more commonalities between interactivism and Piaget's genetic epistemology, and, therefore, with Drescher's model — striking commonalities, especially in the context of the contrary but dominant encoding orientations. There are also, however, important differences (Bickhard, 1992, 1988b; Bickhard & Campbell, 1989; Campbell & Bickhard, 1986). Piaget's notion of representation is action based and modal, but it is still subtly a correspondence model. For Piaget, concepts are structures of potential actions that are isomorphic with structures of potentialities in the environment — a kind of modal isomorphism of patterns (Bickhard, 1988b; Bickhard & Campbell, 1989; Campbell & Bickhard, 1986; Chapman, 1988) — and Piaget's model of perception is straightforwardly an encoding model (Piaget, 1969). This is action based, constructive, and modal — all different from standard approaches — but it is still a correspondence notion. There is still no way for those mental structures to pick out or to specify what they are supposed to represent — to provide or constitute knowledge that they are in isomorphism or what they are in isomorphism with. There is still no way for these mental structures to avoid the problems of encodingism.

These problematics carry over into Drescher's model. His is similarly action based, constructive, and modal. He recognizes the necessity of action feedback for the construction of representation, for learning — in particular, pragmatic error feedback of when an action does not work as anticipated. But he still construes representation itself — that which is learned — in terms of correspondences between "items" in the system and conditions in the world (cf. Dretske, 1988). “Drescher has recognized the importance of pragmatic error for learning, but has not recognized the emergence of representational error, thus representational content, out of pragmatic error. In the interactive model, in contrast, representational error is *constituted* as a special kind of pragmatic error.” (Bickhard & Terveen, 1995, p. 281). Drescher’s model is a momentous advance over standard approaches in the literature of artificial intelligence and cognitive science. I suggest, however, that there remain residual problems — problems that are avoided in the interactive model.

In presenting and discussing the interactive model, I have presented brief but focused contrastive discussions with the Physical Symbol System Hypothesis and with Drescher’s model. Clearly there are innumerable additional models and approaches that could be examined, and deserve to be examined, such as Cyc, SOAR, PDP, machine learning, autonomous agents, dynamic systems approaches, and so on, as well as further cognitive issues that deserve attention, such perception, learning, language, rationality, instantiations in the central nervous system, and so on. I cannot address these here, but would suggest alternative sources to the interested reader (Bickhard, 1991, 1995, in preparation; Bickhard & Terveen, 1995; Hooker, 1995). I will turn now from the interactive model of representation per se to one of its interesting consequences: an inherent integration of issues of representation, action, and motivation.

Representation, Action, and Motivation.

Encoding representations represent in virtue of some correspondence between them and that which they represent. Typically, those correspondences are assumed to be constructed or invoked via some sort of processing of inputs (Fodor, 1990; Newell, 1980), but even that is not logically necessary: Representational correspondences are intrinsically atemporal. In particular, encodings do not require any agent in order to exist; they are not dependent on action — however much it may be that action is taken to be dependent on (interpreting) them.

Interactive representation, in contrast, is an emergent of certain forms of organization of an interactive agent. Interactive representation cannot exist in a passive system — a system with no outputs. Interactive representation is the anticipatory, the implicit predicational, aspect of interactive systems. Representation and interaction are differing functional aspects of one underlying system organization similarly to the sense in which a circle and a rectangle are differing visual aspects of one underlying cylinder. Action and representation are not, and cannot be, distinct modules.

A similar point holds for motivation too, but to see this requires a brief digression on motivation per se. Encoding representations are consistent with models of completely passive systems; correspondingly, the typical assumption is that the default condition of the system is inactivity. In such a view, the basic question of motivation is “What makes the system do something rather than nothing?” Motivation is a matter of pushing or pulling the system out of its default inactivity.

Living systems, however, are not passive. They are constantly in activity of some sort: to cease activity is to become dead. For living systems, then, the question of motivation is mis-stated: instead of “What makes the system do something rather than nothing?” the proper question of motivation is “What makes the system do this rather than that?” That is, the question of motivation is a question of action and interaction *selection*, not of action and interaction activation or stimulation. The system is always doing something, the question is what determines what it does (Bickhard & Terveen, 1995).

In this form, however, motivation becomes a functional matter — the function of interaction selection. That function, in turn, is precisely what interaction indications are useful for. Indicated interactions with their indicated potential internal outcomes form a primary resource for the system to select what interactions to engage in next. That is, interactive indications and their associated internal outcomes not only implicitly predicate interactive properties of the environment, they also serve the motivational function of interaction selection. Motivation and representation are both aspects, along with interactive competency per se, of one underlying ontology of interactive system organization (Bickhard & Terveen, 1995; Cherian & Troxell, 1995).

In complex organisms, and other complex systems, it is possible for relatively specialized and dedicated subsystems to develop that subserve complex functions of representation or of motivation. But, if the interactive model is correct, such specializations must arise out of, and on the foundation of, the basic interactive competence, interactive representation, and interactive motivational selection aspects of underlying interactive system organization. Certainly we do not find specialized such subsystems in simple organisms, only in complex organisms.

Conclusions.

The three phenomena of action and interaction, representation, and motivation, then, do not form separate functional modules that can simply be pasted together in a more complicated system if a more complex design is desired, or if modeling those additional complexities in a natural agent is desired. To the contrary, neither interaction nor representation nor motivation can be correctly modeled without, at least implicitly, modeling all three.

Conversely, if such modularization *is* possible within some modeling approach, then that approach is almost certainly assuming or presupposing an encodingism toward representation. That is, if such modularization is possible in a modeling approach, then that approach is almost certainly founded on a logical incoherence.

Cognition, then, is *not* an autonomous subsystem — and any approach or programme that permits cognition to seem autonomous is foundationally flawed. Such a foundational incoherence, in turn, can have myriad and ramified pernicious consequences throughout the programme or programmes involved — and programmatic errors can be extremely difficult to diagnose and to avoid (Bickhard & Terveen, 1995). Nevertheless, encodingism and its associated modularizations dominate contemporary artificial intelligence and cognitive science. It will not be possible to understand or to design beings with minds within such an approach — artificial intelligence and cognitive science are dominated by programmatic assumptions that make their own highest level programmatic aspirations impossible (Bickhard & Terveen, 1995).

Mark H. Bickhard
Department of Psychology
17 Memorial Drive East
Lehigh University
Bethlehem, PA 18015
mhb0@lehigh.edu

References

Beer, R. D. (1990). *Intelligence as Adaptive Behavior.* Academic.

Beer, R. D. (1995). Computational and Dynamical Languages for Autonomous Agents. In R. Port, T. J. van Gelder (Eds.) *Mind as Motion: Dynamics, Behavior, and Cognition*. MIT.

Beer, R. D. (in press). A Dynamical Systems Perspective on Agent-Environment Interaction. *Artificial Intelligence*.

Beer, R. D., Chiel, H. J., Sterling, L. S. (1990). A Biological Perspective on Autonomous Agent Design. In P. Maes (Ed.) *Designing Autonomous Agents.* (169-186). MIT.

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

Bickhard, M. H. (1982). Automata Theory, Artificial Intelligence, and Genetic Epistemology. *Revue Internationale de Philosophie, 36*(142-143)*,* 549-566.

Bickhard, M. H. (1988). The Necessity of Possibility and Necessity: Review of Piaget's Possibility and Necessity. *Harvard Educational Review, 58*, No. 4, 502-507.

Bickhard, M. H. (1988b). Piaget on Variation and Selection Models: Structuralism, Logical Necessity, and Interactivism *Human Development, 31*, 274-312. Also in L. Smith (Ed.) (1992). *Jean Piaget: Critical Assessments*. Routledge, chapter 83, 388-434.

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

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

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). Critical Principles: On the Negative Side of Rationality. In Hooker, C. A., Brown, H. I. (Eds.) *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., 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*. Elsevier Scientific.

Brooks, R. A. (1991a). New Approaches to Robotics. *Science, 253*(5025), 1227-1232.

Brooks, R. A. (1991b). How to Build Complete Creatures Rather than Isolated Cognitive Simulators. In K. VanLehn (Ed.) *Architectures for Intelligence*. (225-239). Erlbaum.

Brooks, R. A. (1991c). Challenges for Complete Creature Architectures. In J.-A. Meyer, S. W. Wilson (Eds.) *From Animals to Animats*. (434-443). MIT.

Campbell, D. T. (1974). Evolutionary Epistemology. In P. A. Schilpp (Ed.) *The Philosophy of Karl Popper*. (413-463). LaSalle, IL: Open Court.

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

Chapman, M. (1988). *Constructive Evolution: Origins and Development of Piaget's Thought*. Cambridge: Cambridge University Press.

Cherian, S., Troxell, W. O. (1995). Interactivism: A Functional Model of Representation for Behavior-Based Systems. Submitted to ECAL95.

Cherian, S., Troxell, W. O. (in press). Intelligent behavior in machines emerging from a collection of interactive control structures. *Computational Intelligence*.

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

Clancey, W. J. (1991). The Frame of Reference Problem in the Design of Intelligent Machines. In K. VanLehn (Ed.) *Architectures for Intelligence: The Twenty-Second Carnegie Symposium on Cognition*. (357-423). Hillsdale, NJ: Lawrence Erlbaum Associates.

Drescher, G. L. (1986). Genetic AI: Translating Piaget into Lisp. MIT AI Memo No. 890.

Drescher, G. L. (1991). *Made-Up Minds*. MIT.

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

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

Fodor, J. A. (1983). *The Modularity of Mind: An Essay on Faculty Psychology.* Cambridge, Mass.: MIT.

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

Fodor, J. A. (1987b). A Situated Grandmother? *Mind and Language, 2*, 64-81.

Fodor, J. A. (1990). *A Theory of Content*. Cambridge, MA: MIT Press.

Fodor, J. A. (1994). Concepts: A Potboiler. *Cognition, 50*, 95-113.

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

Hanson, P. P. (1990). *Information, Language, and Cognition*. Oxford.

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

Hookway, C. (1985). *Peirce*. Routledge.

Houser, N., Kloesel, C. (1992). *The Essential Peirce. Vol. 1*. Indiana.

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

Malcolm, C. A., Smithers, T., Hallam, J. (1989). An Emerging Paradigm in Robot Architecture. In T Kanade, F.C.A. Groen, & L.O. Hertzberger (Eds.) *Proceedings of the Second Intelligent Autonomous Systems Conference*. (pp. 284-293). Amsterdam, 11--14 December 1989. Published by Stichting International Congress of Intelligent Autonomous Systems.

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

Newell, A. (1980). Physical Symbol Systems. *Cognitive Science, 4*, 135-183.

Palmer, S. E. (1978). Fundamental aspects of cognitive representation. In E. Rosch & B. B. Lloyd (Eds.) *Cognition and categorization*. (259-303). Hillsdale, NJ: Erlbaum.

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

Piaget, J. (1969). *The Mechanisms of Perception*. New York: Basic.

Piaget, J. (1986). Essay on Necessity. *Human Development, 29*, 301-314.

Piaget, J. (1987). *Possibility and Necessity*. *Vols. 1 and 2*. Minneapolis: U. of Minnesota Press.

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

Rosenthal, S. B. (1990). *Speculative Pragmatism*. Open Court.

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.

Smithers, T. (1994). On Behaviour as Dissipative Structures in Agent-Environment System Interaction Spaces. Presented at the meeting "Prerational Intelligence: Phenomenology of Complexity in Systems of Simple Interacting Agents," November 22-26, 1993, part of the Research Group, "Prerational Intelligence," Zentrum für Interdisziplinäre Forschung (ZiF), University of Bielefeld, Germany, 1993/94.

Thayer, H. S. (1973). *Meaning and Action*. Bobbs-Merrill.

Vera, A. H., Simon, H. A. (1993). Situated action: A symbolic interpretation. *Cognitive Science, 17*(1), 7-48.