Title

The Proper Treatment of Cognition

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 14(0)

Author

van Gelder, Timothy

Publication Date

1992

"The Proper Treatment of Cognition"

Tim van Gelder
Department of Philosophy, Indiana University
Bloomington IN 47405
(812) 855 7088 tgelder@ucs.indiana.edu

Abstract

The apparent contradiction between Smolensky's claim that connectionism is presenting a dynamical conception of the nature of cognition as an alternative to the traditional symbolic conception, and Giunti's recent elaboration of computational systems are special cases of dynamical systems can be resolved by adopting a framework in which (a) cognitive systems are dynamical systems, (b) cognition is state-space evolution in dynamical systems, and (c) differences between major research paradigms in cognitive science are differences in the kind of dynamical systems thought most appropriate for modeling some aspect of cognition, and in the kinds of concepts, tools and techniques used to understand systems of that kind.

If cognition consists of those internal, knowledge-based processes which underlie sophisticated human or animal behavior, then the primary question that cognitive scientists address is: what kind of processes are these?[1] A wide range of answers have been proposed, varying with the particular cognitive domain (vision, language processing, etc) under consideration and the level of abstraction at which the answer is framed. It is now becoming increasingly apparent, however, that most if not all such answers can be subsumed under one very general empirical hypothesis: cognition is state-space evolution in dynamical systems. This hypothesis follows naturally from two key insights, discussed below. The first is Smolensky's realization that a dynamics-based conception of cognition provides a deep alternative to traditional computational approaches. The second, paradoxically, is Marco Giunti's demonstration that traditional computational systems are special cases of dynamical systems. The apparent contradiction is resolved by seeing cognitive systems and models as drawn from a wide range of possible kinds of dynamical systems. The deep contrast is not between computational systems on one hand and dynamical systems on the other; it is between *kinds* of dynamical system, and corresponding kinds of concepts, tools and techniques for analyzing them.

1. The Proper Treatment of Connectionism

In his widely-read and influential article The Proper Treatment of Connectionism (PTC) (1988), Smolensky's aim was to articulate the connectionist approach to cognitive science, and to contrast it with the traditional "symbolic" approach. Since the latter approach has been described in detail in many places (e.g., Pylyshyn 1984), I will not elaborate on it here; suffice to say that, for current purposes, it can be summarized as the view that cognition is essentially computation: (something like) the rule-governed manipulation of symbolic representations with "conceptual" level semantics. Smolensky discussed many points of contrast between the symbolic and the connectionist approaches, but of particular concern here is the general account of the nature of cognition itself that he claimed to find embodied in connectionist

---

[1] This characterization of cognitive science is not intended to exclude the detailed study of actual human or animal performance. As Chomsky for one pointed out, often the most appropriate first stage in the study of cognition is to gain an adequate description of the performance itself. This characterization is also not intended to beg any questions about the extent to which those processes underlying sophisticated performance need to be knowledge-based.

work. A synthesized version of that account is summarized in the following claims:

PTC Dynamical Cognition Hypothesis

(1) Connectionist networks are high-dimensional, continuous and non-linear dynamical systems consisting of networks of interconnected units.
(2) Cognitive systems are connectionist networks.
(3) Cognition is state-space evolution within connectionist networks.
(4) The most appropriate tools for the study of cognition are dynamical modelling and dynamical systems theory.

It turns out that this cluster of claims has been largely ignored in subsequent discussion; for example, almost no mention of these themes is made in the interdisciplinary peer commentary that accompanied PTC in *Behavioral and Brain Sciences*. This is somewhat surprising, since Smolensky is here articulating, apparently for the first time, a deep and exciting new description of the nature of cognition, one very different from the dominant symbolic conception.

The PTC approach can be assessed from at least two directions: as a description of connectionism and its conceptual innovations, on one hand, and as an hypothesis concerning the nature of cognition on the other. As an account of connectionism it is in some ways misleading. It is probably true nowadays that most connectionist networks are high-dimensional, continuous and non-linear, but of course there have been and still are strains of connectionist work that reject properties such as continuity, or non-linearity. More importantly, only a relatively small portion of connectionist researchers bring genuinely dynamical methods to bear in their descriptions of network functioning or cognitive processes, at least in any extensive or systematic way. Indeed, most connectionist researchers seem to shy away from dynamical methods even though the networks they set up are in fact dynamical systems defined by differential or difference equations. There are at least two fairly standard strategies for doing this. One is to observe the behavior of the system only over very few time steps - often, as few as one or two, in standard feed-forward backpropagation networks. The other strategy is to focus attention at any given time only on restricted portions of the state-space - e.g., on the possible activity patterns over the hidden units, or those over the output units. Indeed, it is standard practice to *shift* attention from one portion of the state-space to another at each time step, as when one observes how an input pattern is transformed into a pattern over the first hidden layer, and so forth. Most connectionists critically depend on manoeuvres such as these, but they both represent ways of avoiding thinking of cognitive processing as general state-space evolution, and consequently enable and encourage connectionists to use analytical techniques quite different from those standard in dynamical modeling and dynamical systems theory. In short, the PTC perspective may be true to most connectionist work in some respects, but in others describes only a small portion of it. It is probably best regarded as at least partly normative: as describing, in other words, what might be thought of as the "most interesting" connectionist approaches.[2]

As an account of a novel conception of the nature of cognition the PTC perspective is also somewhat misleading. A fundamental component of PTC is a shift to a dynamics-based conception of cognition. Smolensky implies that this is a distinctively connectionist contribution. However, there are increasing numbers of researchers adopting dynamical approaches to the study of various aspects of cognition without being connectionists (see, e.g., van Geert 1991, Townsend 1989, Skarda & Freeman 1987). These researchers deploy dynamical systems theory in constructing models of cognitive processes, even though those models do not come in the form of networks of interconnected processing units. They believe that cognitive systems are dynamical systems, and that cognition is state-space evolution; they see the importance of properties such as continuity and non-linearity as much as any connectionist. From their point of view, connectionist networks are *just one way* to implement genuinely dynamical approaches.

The upshot of these points is that the dynamical conception of cognition Smolensky articulates in PTC (a) accurately characterizes only part of connectionist work, but (b) is held in common with various other non-connectionist strands of research. Together, these suggest that it would be wrong to tie the exciting idea that cognition is a dynamical phenomenon too closely to connectionism in particular. There is a

---

[2] For a brief discussion of the use of dynamical explanatory methods in connectionist psychological modelling, see van Gelder 1991.

different and more natural conceptual boundary to be drawn. It does not identify the dynamical conception of cognition with connectionism, but rather uses the dynamical conception as the central commitment tying together a diverse group of researchers which *includes* some connectionists. We can thus think of "dynamicists" as those researchers committed to something like the following very general claims:

General Dynamical Cognition Hypothesis:

(a) Cognitive systems are [non-computational; see below] dynamical systems.
(b) Cognition is state-space evolution within such dynamical systems.
(c) The most appropriate tools for the study of cognition are dynamical modelling and dynamical systems theory.

It is important to see that this hypothesis has two sides. One specifies (in very abstract terms) the nature of cognitive processes. The other is methodological: it recommends certain kinds of tools as most appropriate for the detailed investigation of cognition. As will become more clear below, these two sides are complementary.

2. Computational Systems as Dynamical Systems

In a recent PhD dissertation Marco Giunti has exhaustively elaborated the thesis that computational systems, including those deployed in the mainstream symbolic approach to the study of cognition, are special cases of dynamical systems (see Giunti 1991). Here I will only illustrate his position with a coarse description of a paradigm example of computational systems, the Turing Machine, as a dynamical system. The overall state S of a Turing Machine at time t is fully specified when we know the contents of every cell on the tape, the current head state, and the current location of the head. Since the tape is unbounded in both directions, a useful way to represent this overall state is S(t) = ...aaaaqaaaa.... where each "a" designates the contents of a cell of the tape, q is the head state, and the head is positioned over the cell immediately to the right of q. The evolution equation F for the Turing Machine is a general specification of behavior of the machine, i.e., a specification of what state the machine will go to at time t+1 depending on the state it is in at time t, as depicted in the following schema:

$$\begin{array}{ccc} t & F & t+1 \\ \text{....aaaaqaaaa....} & \Longrightarrow & \text{....aaaaqaaaa....} \end{array}$$

Each state transition in a Turing Machine involves three elementary changes: writing in the current cell, changing head state, and moving the head either right or left. The exact nature of the state transition depends on the contents of the current cell and the current head state, in a way that is specified in the machine table. Thus the machine table really is the evolution equation, though encoded in a somewhat unusual form. The general form of an evolution equation for a discrete system is S(t+1) = F(S(t)). In this case the equation is a tedious conditional easily reconstructible from the machine table. The table below, for example, gives the evolution equation for Minsky's seven state, four symbol universal Turing Machine (Minsky 1967).

Of course, there is nothing distinctive about Turing

S(t+1) = F(S(t)) =	....aaa1a_aaa..... if S(t) = ...aaaa1Yaaa.....	....aaa7a1aaa.... if S(t) = ...aaaa41aaa....
	....aaa1a_aaa..... if S(t) = ...aaaa1_aaa.....	....aaa4a1aaa.... if S(t) = ...aaaa4Aaaa....
	....aaa2a1aaa..... if S(t) = ...aaaa11aaa.....	....aaaaY5aaa.... if S(t) = ...aaaa5Yaaa....
	....aaa1a1aaa.... if S(t) = ...aaaa1Aaaa....	....aaa3aYaaa.... if S(t) = ...aaaa5_aaa....
	....aaa1a_aaa.... if S(t) = ...aaaa2Yaaa....	....aaaaA5aaa.... if S(t) = ...aaaa51aaa....
	....aaaaY2aaa.... if S(t) = ...aaaa2_aaa....	....aaaa15aaa.... if S(t) = ...aaaa5Aaaa....
	....aaaaA2aaa.... if S(t) = ...aaaa21aaa....	....aaaaY6aaa.... if S(t) = ...aaaa6Yaaa....
	....aaaaY6aaa.... if S(t) = ...aaaa2Aaaa....	....aaa3aAaaa.... if S(t) = ...aaaa6_aaa....
	....aaa3aYaaa.... if S(t) = ...aaaa3Yaaa....	....aaaaA6aaa.... if S(t) = ...aaaa61aaa....
(halt)	....aaaa3_aaa.... if S(t) = ...aaaa3_aaa....	....aaaa16aaa.... if S(t) = ...aaaa6Aaaa....
	....aaa3aAaaa.... if S(t) = ...aaaa31aaa....	....aaaa_7aaa.... if S(t) = ...aaaa7Yaaa....
	....aaa4a1aaa.... if S(t) = ...aaaa3Aaaa....	....aaaaY6aaa.... if S(t) = ...aaaa7_aaa....
	....aaa4aYaaa.... if S(t) = ...aaaa4Yaaa....	....aaaa17aaa.... if S(t) = ...aaaa71aaa....
	....aaaaY5aaa.... if S(t) = ...aaaa4_aaa....	....aaaa_2aaa.... if S(t) = ...aaaa7Aaaa....

Machines in this regard, although their relative familiarity makes it particularly easy to illustrate the point. From this perspective, computation - a particular sequence of symbol manipulations within a computational system - turns out to be a matter of state-space evolution within the particular kind of discrete state space offered by a digital computer. (Indeed, we might say that computation is a matter of *touring* the state-space.) Consequently, when the symbolic approach to cognition construes cognitive processes as computational processes, it also is construing them as state-space evolution within (computational) dynamical systems.

3. The Space of Cognitive Systems

The fact that computational systems can be described as dynamical systems has important implications for the discussion in the first section. The PTC dynamical cognition hypothesis and its more general counterpart were both intended as presenting *alternatives* to the symbolic conception of cognition. The deep difference between the symbolic approach and the dynamical alternatives cannot, however, be a contrast between symbol manipulation on one hand and state-space evolution in dynamical systems on the other, for the former is a special case of the latter. Rather, *the significant differences must lie in the kind of dynamical system employed, and the kinds of concepts and tools one uses in describing these systems.*

Occasionally, snippets of the official rhetoric of the computational approach to cognition has been dynamical in flavor; recall, for example, Newell & Simon's definition of a *physical symbol system* as "a machine that produces through time an evolving collection of symbol structures"[3]. Typically, however, the tools, techniques and concepts of dynamical modelling and dynamical systems theory are completely absent from standard discussions of computational systems. Why is this, if computational systems are special cases of dynamical systems? The answer is that certain ways of thinking about the behavior of systems lend themselves most naturally to certain kinds of systems. Various deeply different ways of understanding behavior can be applied to dynamical systems, but are most effective when applied to systems of particular kinds. Further, it is in the nature of standard computational systems, as dynamical systems, to encourage algorithmic rather than dynamical ways of thinking.

Consider the Turing machine again. This kind of computational device originated as Turing's own formalization of the process of elementary arithmetical calculations using pencil and paper. It is from this humble origin that the extremely simple nature of basic Turing machine operations derive. Consequently, Turing Machines as dynamical systems are fundamentally:

(1) Discrete. Each transition change takes place at a distinct point in "time".

(2) Digital. State transitions involve a jump from one unambiguously identifiable state to another. The symbols which can appear in the cells, the head states, and the head positions, are all digital in character.

(3) Deterministic. From each state there is only one next state to which the machine can change.

(4) Low interdependency of state variables. The Turing Machine system contains an unbounded number of state variables, but, in general, the change in any given state variable depends on only a very small number of these. For example, each cell on the tape corresponds to a distinct state variable. Will the value of that variable change in a given state variable? That depends on the values of only two other variables - i.e., on head position, and head state.

(5) Local. Each state transition involves changes in only three of the unbounded variables. The outcome of each transition is another point very "close" in state space.

It is, of course, no accident that the Turing Machine exhibits this particular combination of features. Basically, they make it possible to think of the behavior of the machine as the following of an *algorithm*. The low interdependency of state variables and local nature of state transitions enable one to *ignore* most of the state of the machine at any given time; after all, any change depends on only two variables and affects at most three. These features, in other words, encourage one to think of processing steps not as transitions from one total state of the system to another, but rather as localized alterations in particular

---

[3] Newell & Simon (1981) p.40. In Human Problem Solving (1972; pp.11-12) they maintain that "the explanations of cognitive science are not in principle different from the explanations of any other science which is concerned with the dynamical behavior of some system.".

variables. They encourage thinking of the behavior of the system not in the geometrical sense of how the system is moving through its state space, but in the mechanical or syntactic sense of how particular constituents are being manipulated. Consider then the effect of adding the other three major features - discreteness in time, and digital and deterministic state transitions. Combined, these have the effect that each of these highly local alterations can be specified by its own simple rule. The total behavior of the machine is then a sequence of elementary rule-governed steps. By careful ordering of these steps, the desired overall effect is achieved - i.e., the behavior of the machine is specifiable by an algorithm. That is, it is in the very nature of a Turing machine, as a dynamical system, that all its basic state transitions can be *micro-managed* by the designer of the machine. The whole idea is to enable the designer to achieve controlled complexity in the overall behavior of the machine by orderly sequencing of carefully defined elementary operations. Complexity of global behavior is supposed to flow from simplicity and order at the base.

My claim, then, is that fundamental features of the Turing Machine as a dynamical system directly facilitate thinking about its behavior in algorithmic terms - and algorithmic modes of analysis are deeply different from dynamical ones. Generalizing, I ambitiously assert that what is true here of Turing Machines holds true of computational systems more generally. Computational systems (von Neumann machines, LISP machines, production systems, etc) form a natural class by virtue of sharing certain fundamental characteristics that enable us to most effectively describe their behavior in basically algorithmic terms. Conversely, the kind of connectionist systems that Smolensky had in mind when he formulated the PTC conception of cognition share certain other fundamental characteristics which render dynamical techniques fundamentally appropriate in their analysis.

Waxing metaphorically, we can think of particular dynamical systems as falling into a vast space of possible kinds of dynamical systems. The axes of this space are the fundamental properties that such systems can have - properties such as continuity vs discreteness, degree of interdependence of state variables, linearity vs non-linearity, and so on. Typical computational systems possess a certain characteristic set of properties and so "cluster" in one region of the space of possible systems. The symbolic approach to cognition can then be seen as the empirical hypothesis that real cognitive systems are dynamical systems which also fall into that particular region of the space - somewhere relatively "close" to Turing Machines. PTC, by contrast, focuses on systems that are "connectionist" - typically high-dimensional, continuous and non-linear - and is committed to the empirical hypothesis that real cognitive systems belong in this corner. The more general dynamical conception proposed at the end of Section 1 can then be seen as the suggestion that the PTC perspective circumscribes the corner into which cognitive systems fall a little too narrowly, though even the wider area embraces only systems which demand dynamical methods in their analysis. Most generally of all, each of these perspectives shares the fundamental assumption that real cognitive systems are located *somewhere* in this space of possibilities. This is equivalent to the the broad empirical hypothesis that cognition is state-space evolution in dynamical systems.

If this is right, it poses at least three major questions for further research. First, what really are the key dimensions of the space of possible dynamical systems? What are the deep properties which make for fundamental differences among kinds of dynamical systems? To stretch the spatial metaphor to its limits, what *principal components* can we abstract from the kinds of dynamical systems we already know about? A variety of important issues have already figured in the discussion so far (i.e., continuity, non-linearity, degree of interdependence of state variables, number of state variables, digital, deterministic), but there are also many other relatively obvious candidates (e.g., systems might have numerical vs arbitrary symbolic state variables, or be time-invariant, homogeneous, reversible, or chaotic), and no doubt a variety of not-so-obvious ones as well. From the perspective being advanced here, properly understanding the possible forms that cognitive processes might take, and the relationships between different research programs in cognitive science, presupposes clearly understanding the most basic kinds of properties that dynamical systems can have.

Second, what are the natural clusters within this space of possibilities? This is really the question: what are the natural kinds of dynamical systems (if any), based on their deep properties? It seems plausible that, for example, classical computational systems and perhaps connectionist systems (or various sub-categories of them) cohere into identifiable types, on the basis of which more specific hypotheses concerning the nature of cognitive processes can be framed. But are there other candidates as well? Perhaps certain types of analog computers, or the kinds of systems deployed by the non-connectionist dynamicists mentioned above, are equally candidates.

Third, what arguments can be formulated for supposing that real cognitive systems belong to a given kind? Mainstream computational cognitive science can be understood as making a bet - underwritten by some respectable arguments - that real cognitive systems belong in their corner of the space of possible dynamical systems, and hence that computational systems will provide the best models, and computational methods will provide the best analyses. From the point of view of others, such as dynamicists, the computational corner looks more like a ghetto, a particularly narrow and confining nook which people stay in not out of choice but because of unfortunate historical contingencies. Their bet is that real cognitive systems are to be found in relatively remote regions inhabited by systems which demand genuinely dynamical techniques if they are to be properly understood. The general arguments in favor of this position are yet to be worked out in detail, but at least one intuition is worth mentioning at this stage: if computational systems are attractive cognitive models for agents conceived as *abstract reasoners*, certain kinds of non-computational systems appear more deeply suited for models of agents conceived as *situated actors*.

References

Giunti, M. 1991. *Computers, Dynamical Systems, Phenomena, and the Mind.* PhD. Dissertation, Department of History and Philosophy of Science, Indiana University.

Minsky, M. 1967. *Computation: Finite and Infinite Machines.* Englewood Cliffs, NJ: Prentice-Hall.

Newell, A., & Simon, H. 1981. Computer science as an empirical enquiry. in Haugeland, J. ed. *Mind Design.* Montgomery VT: Bradford Books.

Newell, A. & Simon, H. 1972. *Human Problem Solving.* Englewood Cliffs, NJ: Prentice-Hall.

Pylyshyn, Z. 1984. *Computation and Cognition.* Cambridge MA: MIT Press.

Skarda, C.A. & Freeman, W.J. 1987. Brain makes chaos to make sense of the world. *Behavioral and Brain Sciences* 10:161-195.

Smolensky, P. 1988. On the proper treatment of connectionism. *Behavioral and Brain Sciences* 11:1-74.

Townsend, J.T. & Busemeyer, J.R. 1989. Approach-Avoidance: Return to Dynamic Decision Behavior. in Hizuko Izawa ed. *Current Issues in Cognitive Processes.* Hillsdale NJ: L.Erlbaum Associates.

van Geert, P. 1991. A dynamic systems model of cognitive and language growth. *Psychological Review* 98:3-53.

van Gelder, T.J. 1991. Connectionism and Dynamical Explanation. *Proceedings of the 13th Annual Conference of the Cognitive Science Society.* Hillsdale N.J.: L.Erlbaum Associates; 499-503.