Levels of Granularity in Cognitive Modeling:

Towards a “grand unified theory” of cognition

Eric Nichols (epnichol@indiana.edu)
Center for Research on Concepts and Cognition
Bloomington, IN 47408 USA
December 11, 2006

Introduction

Cognitive scientists have historically used a diverse group of approaches to modeling of cognitive processes, including physical symbol systems, behaviorism, connectionism, embodiment, dynamic systems, etc. Comparing such different models seems difficult, in part, due to the different levels of description used. For instance, models describing generalized behaviors without describing internal states of an organism work at a very high level but may fail to make predictions that would require more detail. Conversely, a model based on simulating many details of neural anatomy may fail to make useful predictions of overall behavior due to the engineering challenges of implementing such a detailed model on a large scale. Many cognitive models come bundled with implicit assumptions about the appropriate types of processes to model.

Within each general modeling paradigm, such as connectionism or dynamical systems, any particular model also makes a choice as to the desired level of granularity. A connectionist model might attempt to capture details of individual neurons, as in a retinal model, or else it might allow the model’s nodes to represent larger abstract groupings of functionally related neurons. This paper surveys several broad classes of cognitive models with respect to their stance on model granularity.

In addition to examining traditional cognitive models, this paper also discusses how we might attempt to overcome the granularity divide amongst these models. In particular, we point towards Holland’s work on emergent behavior in complex adaptive systems as paving the way for understanding cognition more deeply at multiple levels of granularity. We begin by motivating the discussion by examining the goals of cognitive modeling.

Purpose of Modeling

We consider cognitive science to be the study of thinking – how it is that brains can process information, manipulate ideas, and effect changes in the body attached to the brain to interact with the environment. Here we consider standard approaches to cognitive science in terms of how they model thinking. The very idea of modeling in science, however, deserves some clarification before we proceed with modeling in cognitive science.

What is the purpose of modeling in general?

The primary reason to build a model of a physical system is arguably to make prediction possible. While the best possible model of any system would be a facsimile of the system itself, such a “model” provides little predictive power. We might build a scale replica of our solar system, but Kepler’s laws of planetary motion allow us to immediately make predictions far into the future that would be impossible with the replica (Holland 1998). Of course, predictive power due to simplification of reality comes at the price of loss of predictive accuracy. Indeed, Holland claims that “the single most important factor in making an accurate prediction is the level of detail” (Holland 2002). In Holland’s context, a detailed model of a chess game would make a prediction about all the successive moves that would occur starting at a given position, while a less detailed (but still useful) model would predict the overall outcome of the game from that position, without specifying a particular series of moves. Note that while "detail" here refers to the level of specificity in prediction, we also consider the number of factors in the model and the level of granularity to be key components of the "level of detail".

A key task facing a modeler is to choose the level of detail that best balances amount of detail and predictive power. Note that, in general, including more details does not necessarily lead to more accuracy, due to the complexities inherent in simulating complex systems. Typically, given a particular modeling paradigm, the error in the model will decrease as the amount of detail is increased. Eventually, however, the amount of extra computation time required to add detail will make computation intractable. Worse, it is often the case that after reaching a minimum the amount of error will actually begin to increase due to effects of numerical instability or chaotic dynamics.

Due to complexity, prediction with a high degree of accuracy may not be possible in a model uninformed by a requisite amount of *understanding*. The modeler needs to select not only the amount of detail to include in a model, but also the general types of structures that make up the model. A model of the weather, for instance, might explicitly include such high-level structures as storm fronts or the jet stream. A forecaster might either hope for these features to emerge from a lower-level model or else build these into the model up-front. The choice has important ramifications: a model based on the higher-level granularity of storm fronts might be more reliable in some situations, while a lower-level model based on features like air pressure and moisture content might perform better in other cases. Deep understanding of the system under consideration is necessary to posit higher-level features and to describe their dynamics. A secondary motivation for modeling is to gain the understanding necessary to develop richer models to aid the primary goal of prediction. Similarly, increased understanding of a system may be a necessary prerequisite to the development of more advanced theories.

What is the goal of cognitive modeling?

As a cognitive scientist it seems reasonable that the goal of cognitive modeling is to better understand the internal mechanisms of thought, to see how thinking "really" works. The study of cognition has particular significance as it is closely related to deep philosophical questions about the nature of mind and consciousness. But we can productively sidestep the issue of philosophy of mind by recalling the secondary motivation of modelers above, and focus on understanding thought mechanisms in service of building better-informed predictive models of cognitive processes.

Marr proposes three different levels of explanation for use in understanding cognitive systems: 1) computational theory (input-output analysis), 2) representation and algorithm, and 3) implementation (Marr 1982). Clark argues that although this division into three distinct levels may be "too neat", it points out how understanding a single level such as the neural substrate of the brain is not sufficient to understand how computation is organized at larger levels of granularity.

The word "level" is used here in different contexts, so it is important to keep the uses distinct. Marr's three levels refer to the three distinct types of analysis listed above. The level of granularity in a model typically refers to the relative size of the units of representation; thus this sort of granularity differentiates between properties within Marr's level two. Holland's discussion of "level of detail" (as in predicting chess moves vs. prediction the outcome of an entire game) invokes the idea of granularity of the model output (as opposed to the input or internal granularity of a model). This idea of level thus involves the granularity within Marr's level one.

Keeping in mind our discussion of modeling as a predictive tool as well as Marr's levels of explanation, let us review several important approaches to cognitive science.

Survey of Cognitive Models

Behaviorism

The behaviorist school is more of a psychological approach than a cognitive model, but we can briefly consider it in the terms described above. Behaviorism stipulates that an organism's behavior can be studied without considering its internal (cognitive) states. Of primary importance is the effect of the environment on organism. Skinner acknowledges that internal states of an organism exist, writes that "the organism is not empty, and it cannot adequately be treated simply as a black box" "Something is done today which affects the behavior of the organism tomorrow" (Graham 2005). In Marr's terms, this is analysis on the first level alone; any internal representations and implementation are left out of the picture. Indeed, the lack of representation is "one of the defining features of traditional behaviorism" (Graham 2005). In terms of model granularity, behaviorism can be considered a model where the size of the postulated internal states has been increased to its limit: each structure in the model is simply equated to an external behavior – there is no room for any smaller internal component in the model.

To make this more clear, imagine model granularity as controlled by a knob. Turning the knob to the left, say, decreases granularity until the model contains only representations of individual cells in the body, such as neurons. Continuing to turn the knob left results in models based on individual molecules, then atoms, and finally subatomic particles. Turning the knob to the right, however, increases granularity as we see models based on larger assemblies of neurons in the brain, then entire gross anatomical segments of the brain, the entire brain itself, and finally the organism as a whole. Behaviorism lies near this knob setting, where the organism as a whole and its immediate environmental stimulus are considered. Continuing to turn he knob the right yields models based on groups of individuals, entire populations, interactions between species, etc.

Philosophically, behaviorism is quite directly aligned with the goal of model-based prediction – behaviorism studies the output of an organism based on its environmental inputs. However, the additional choice of avoiding internal states makes the behaviorist's task difficult. For example, behaviorists will model behavior that is learned over time as a response to environmental stimuli, but this learning must be situated, from a modeling perspective, in the precarious position of a variable that exists outside of the organism. Behaviors such as those related to basic survival instinct might be predicted reliably, but those that depend on an organism's idiosyncratic representation of past history would be more difficult to predict. In a given modeling situation, even if the obvious direct route to simulation and prediction would make use of internal representation, the behaviorist model would be restricted to a tricky analysis and interpretation of history and behavior with internal states only present as hidden variables lurking behind the scenes. Skinner's simultaneous acknowledgement of the existence of inner states along with the commitment to leaving them out of any model (Graham 2005) results in a Quixotic task for the behaviorists when faced with situations where behavior manifests itself spontaneously based on the processing of inner states. An extremely detailed behaviorist model accounting for an entire life history of an individual might surmount these difficulties, but this both seems terribly difficult and is not the typical behaviorist program.

Physical Symbol Systems

Turning our granularity knob back to the left from the behaviorist setting, we come across an approach that does allow internal state to factor into predictions. The Physical Symbol System Hypothesis states that "A physical symbol system has the necessary and sufficient means for general intelligent action." (Newell and Simon, 1976) That is, a sufficiently complex PSS can be "programmed" to produce intelligent behavior, and additionally anything that is intelligent must make use of a PSS. The claim of the PSS view is that thinking is computation via symbol-manipulation according to formal rules. Additionally, there is a commitment to "semantically transparent systems" in PSS (Clark 2001), where the symbols in use have a clear meaning that maps onto familiar things in the world.

Note that in addition to turning the granularity knob to allow more internal state in the model, the PSS view also emphasizes Marr's second level – representation and algorithm, while abstracting away the role of the physical implementation. Notions of Turing equivalence in computational theory are invoked to suggest that the physical details are irrelevant to the types of computation that the brain can perform. Marr's first level is also less important to this view. Indeed, problems with the nature of inputs and outputs abound when PSS comes up against Searle's Chinese Room argument.

While it seems that the PSS hypothesis might apply to any level of detail in a cognitive system, typical uses of the theory for building cognitive models act at a relatively high level where objects like restaurants, food, tables, and "going out to dinner" are the symbols under consideration (in Schank's scripts), or similarly symbols like "Eaten-by(hamburger, John)" and "Loves(John, Mary)" in a predicate calculus framework. Further, PSS research is based on the idea that "intelligence resides at, or close to, the level of deliberative thought." (Clark 2001) For example, the SOAR architecture models a "cognitive band" of thoughts flowing on a time scale of 10 ms to 10 seconds.

Predictions of PSS models have a reasonable chance to be accurate if thinking is indeed driven by events occurring at this semantically transparent level in the cognitive band. However, there is room for skepticism for several reasons. First, ignoring Marr's implementation level completely may cause the PSS modeler to miss crucial mechanisms of cognition that have been worked out by eons of evolution. For example, the parallel nature of neural computation is poorly approximated by a serial symbol crunching system. Second, the focus on high-level, folk-psychological symbols may miss important details that occur on more rapid time scales. Finally, this commitment to symbols, especially in the predicate calculus framework, generally fixes the representation of each symbol to a fixed meaning that does not respect the dynamical properties of cognition as it unfolds in time.

Connectionism

In an attempt to solve the problems with PSS models, we can continue to turn the granularity knob down, until we reach connectionist models that operate at the neuronal level. Along the way we pass through the space of high-granularity connectionist models where each node in the model corresponds to clusters of many neurons. However, models of connectionist learning are motivated by physical processes at the neural level so it is natural to initial consider connectionism as a rather fine-grained paradigm. Additionally, as we will see below, for connectionism to be considered mid-level with respect to granularity, we would need to see justification for the simplifying scale-up of connectionist learning processes.

Fine-grained connectionism stands in an opposite extreme to the behaviorist perspective, and arguably exemplifies the lowest level of granularity we should look for in a cognitive model, unless we expect that even lower-level things like individual atoms or quantum effects (such as those supported by Roger Penrose) have important bearing on cognition. Connectionism is also an opposite of behaviorism in terms of Marr's levels, as it is most directly interested in the third level, implementation. Connectionist models are directly inspired by physical features of the brain. Even though any connectionist model makes certain simplifications (to avoid modeling-via-facsimile), arguments can be made about the physical justification for various simplification choices.

The obvious problem with connectionism in regards to predictive accuracy is the sheer complexity of any neural model of a brain, with its trillions of components and massive parallelism. Thus, in addition to simplifying the model of each neuron, the entire system of neurons itself must be simplified. In so doing, connectionist models make choices about the inputs and outputs to the model (Marr's first level) – to test the model, there must be input to the system and predictions made about system output. Less obviously, even

Marr's second level of algorithm and representation is involved: single neural units in a model often correspond to groups of neurons in a cognitive system. This higher-level chunking built into the model adds a particular bias to its representations and algorithms; it is not immediately obvious that the mechanisms for neuronal signaling between individual neurons are applicable to aggregates of neurons. Also, the meaning of any neuronal activation pattern is unclear – quite the opposite of the semantic transparency of PSS models.

An analogy with physics may be useful here. When astronomers model planetary motion, they typically regard huge celestial bodies such as planets and moons as point masses. This simplifies calculation of gravitational effects and allows useful predictions to be made, with a high degree of accuracy. This simplification – turning Jupiter into a point with high mass in the context of other planets and moons relatively far away – is justifiable on physical grounds related to the square-distance term in a gravitational force equation. However, in a neural system, the requisite reduction step – from a huge collection of interconnected neurons to a single aggregate neuron with links to other aggregate neurons – is less justifiable until a formal bridge can be built between these two different scales. Connectionism may indeed provide models with good predictive capability for small neural systems, but if our goal is to model large-scale cognition, its use is problematic and unlikely to capture the complexity of real cognitive systems.

Embodiment

Embodiment does not have such an obvious setting on our granularity knob, although its typical incarnation would be at a much higher-level than connectionism. Embodiment concerns large-scale interaction of an organism in its environment. It is less of a modeling strategy as a philosophy of how to carve up the world into pieces for modeling. In particular, it argues against a complete distinction between organism and environment, and calls instead for a more unified view of both. Here Marr's first level, input & output, is being taken as critically important. Including more than just the organism seems like a realistic approach to modeling because more information about the world is accounted for, but it comes at the price of requiring tremendous complexity. Such a model, like a large-scale connectionist model, again approaches the facsimile problem of becoming a copy of the world itself. This increase in complexity reduces the potential for making usable simulations and predictions.

Still, embodiment as a philosophy is compatible with the other strategies discussed here. From our modeling and prediction perspective, we can simply take it as a reminder to seriously account for the environmental factors that are buried in the "inputs" of Marr's first level and to consider the artificiality of the input-output distinction of the model, while making careful decisions about simplifications made to both the cognitive system being modeled and the external world in which it is embodied.

Dynamical Systems

Dynamical systems theory also does not fit as comfortably on our granularity knob, and instead offers a particular mathematical framework for modeling a cognitive system in terms of a state space and transitions between points in state space. Such transitions are also typically couched in terms of differential or difference equations with respect to time. The abstract quality of the framework means that it also has less particular emphasis on any of Marr's three levels – much is left up to the modeler. However, a typical dynamical model will be less interested in the representation level (Beer 2000), and much of the algorithmic nature is dependent into the use of differential or difference equations. The role of time is critically important to dynamical systems equations, and thus is more directly modeled than in most other approaches.

From a predictive point of view, dynamical system models have the advantage of powerful mathematical tools for analysis. The design of a model is, however, difficult due to the same sorts of questions that trouble connectionist models. Dynamical systems are better understood for small numbers of variables and equations, and grow in complexity like connectionist systems do when a large system is modeled.

Summary

Each cognitive modeling approach examined above has its particular weaknesses and strengths. Additionally, most of these have a typical associated level of granularity at which the model operates. We see the reconciliation of these different levels to be of fundamental importance to future cognitive science modeling efforts. In this regard, the final section suggests possible first steps towards a "grand unified theory" of cognition.

Reconciliation of levels

Skinner's radical behaviorism was partly motivated by the desire to avoid circularity of explanation – internal processing, he argued, was a form of behavior, and thus could not itself be used to explain organism behavior (Graham 2005). While behaviorism is unpopular today, the worry about circularity requires a reasoned response. One way to attempt to come to terms with the problem is again to invoke the notion of levels. If we consider internal states to belong exclusively to the realm of Marr's levels two and three – representation, algorithm, and implementation, we can try to imagine that levels two and three involve processing that eventually leads to level-one output (behavior). Looking through the lens of Marr's levels, this seems to solve the problem, except for the "input" part of level one. Even as the organism produces output, new input is constantly fed into the system via sensory input from the environment. Thus we haven't quite escaped some of the circular character of the model that worried Skinner. Behavior indeed is involved in a feedback loop where the environment, external organism behavior, and internal behavior (mental processing) are constantly influencing each other. Skinner's objection to the modeling of inner states at least points out the complexities involved.

The dynamical systems viewpoint may be one way to answer behaviorism, by simply acknowledging this complexity and rising to the modeling challenge even when accounting for inner states. The interplay between organism, environment, and inner states, that Skinner deemed circular might be described in terms of *coupling* of variables in differential equations – the standard dynamical method for dealing with "complex casual relationships" (Clark 2001). However, it seems that today's dynamicists are not particularly concerned with modeling internal *mental* states, but rather with modeling systems of a relatively small number of variables in a state space more closely tied to external behavior than internal processing. Even though Clark (2001) points out that dynamicism represents a radical paradigm shift (exemplified by the *Radical Embodied Cognition Thesis* that "structured, symbolic, representational, and computational views of cognition are mistaken"), he characterizes extant dynamical research as "some quite low-level sensorimotor engagements with the world (finger wiggling, infant walking, Watt governing, etc.)" The philosophy and language of dynamicism suggest that it strives to deal directly with the complexities of interaction between both internal and external "behavior", but in practice research is aimed at a different level that posits a small collection of abstract state variables coupled with external behavior variables. To date, dynamics research cannot adequately answer the challenges posed by the sheer complexity of embodied cognitive systems.

In case there is any doubt about the difficulties raised by complexity in a dynamical model (or indeed in most other cognitive models), consider the problem of categorization. One might imagine a complicated model where the perception of various input features such as the shape, color, and smell of an object figure in to part of a modeled cognitive state. The dynamic perception of new object features over time would move the system state around in a high-dimensional state space, corresponding to the changing perceived categorization of the input object. Such a model already seems quite complicated. Next consider the effect of perceiving something quite unexpected. This happened to me recently while I was waiting to get off an airplane and the pilot informed us that there would be a delay due to the jetway getting a flat tire. It is unlikely given the current state of the art in dynamical system research that there would be a general system capable of modeling the cognitive surprise when this statement led me to reinterpret the “hallway/tunnel” connection between the plane and terminal as some sort of wild automobile with a hallway on top and a flat tire underneath. A dozen or so orthogonal state space variables are not up to the task. How are we to find an appropriate level of analysis capable of respecting such complexity of cognition?

The need for bridge-building

Connectionist models suffer due to complexity in much the same way as dynamical models. In the case of recurrent connectionist models in particular, something like the circular feedback loops that led Skinner to avoid explicit modeling of internal state are present and make theoretical analysis quite difficult. In addition to this complexity based on network topology, the sheer number of neurons involved in biological systems adds another type of complexity. As mentioned above, we don’t have a principled way to simplify from vast numbers of neurons down to a manageable set for use in a connectionist simulation. Such a strategy for simplification with respect to the granularity of the neural model is desirable is it would make tractable models of otherwise too-great complexity. That is, the model would predict essentially the same behavior with a greatly-reduced number of components. This type of simplification could also be useful in the domain of other cognitive approaches such as dynamical systems and physical symbol systems. In these cases the desired simplification would give a mapping from a more complex model to one with many fewer states (state space variables in the case of dynamical systems, or symbols and rules in the case of PSS.)

Even more interesting is the possibility for building connections between different modeling strategies. In addition to just simplifying a connectionist model by reducing the number of nodes, we would like to see a theory of how a connectionist or dynamical model could produce something like a PSS at a higher level. Hofstadter (1986) also seeks such a bridge between levels:

> The best traditional AI (and cognitive psychology) is something like classical physics: true on all scales, but in some sense “irrelevant” on a large scale. The problem is therefore to link these two vastly different levels. In physics, the “correspondence principle” says that the equations of quantum mechanics must turn into their classical counterparts in the limit of large quantum numbers. In that sense, a beautiful bridge is rigorously established between the unfamiliar micro-world and the familiar macro-world. I would like see such a bridge established between connectionism and the study of cognition itself, which includes traditional AI, cognitive psychology, linguistics, and the philosophy of mind.

In astronomy, the bridge-building problem was as simple as calculating a center of mass and discarding negligible gravitational effects. How can we find analogous cognitive centers of mass?

Emergence

We have encountered several difficulties in imagining how to deal with complexity. In connectionist networks, we would like to build bridges between low-level neural models and higher-level models of clusters of neurons. Similarly, we seek a connection between lower level connectionist models and higher-level models such as PSS. Models of symbol systems would benefit from a principled way to pick out their constituent symbols and rules in such a way that dynamic behavior is more easily modeled. Finally, we see dynamical models as confounded by the complexity of high dimensional systems, and suspect that it is difficult to find the appropriate level at which to build a dynamical model while restricting the number of variables to a manageable set. Despite their apparent dissimilarity, we suspect that theories of *emergence* can shed light on all these modeling issues.

Chalmers (2006) clarifies two different notions of the term emergence, which he terms “strong emergence” and “weak emergence”. Weak emergence is the notion that higher-level phenomena may arise in an unexpected way from rules of a lower-level system; much work or simulation may be required to show how the higher-level effect occurs as a result of the lower-level rules, but the connection exists. This is the notion we are concerned with here. Chalmers indicates that weak emergence is the notion that is typically invoked by proponents of emergence in complex systems theory. Strong emergence is a more radical position that some phenomena can not be explained in terms of constituent lower-level systems, and thus *require* understanding in terms of completely new rules that work at the higher level. Chalmers argues that the main example, and indeed perhaps the *only* example, is that of consciousness as a strongly emergent property.

In the philosophy of cognitive science, the term emergence is often used in reference to problems of consciousness (O'Connor 1994). This sort of emergence plays into discussions of free will and (strong) downward causation (Chalmers 2006, O'Connor 1994). In Chalmers' terms, the emergence discussed by philosophers such as Samuel Alexander seems most closely affiliated with strong emergence, while others such as Mill, Broad, and Clark (O'Connor 2006) are interested in something more akin to weak emergence, as we are here in the context of modeling cognitive systems. Chalmers describes how weak emergence is useful in terms of explanatory power:

> Of course, weak emergence may still have important consequences for our understanding of nature. Even if weakly emergent phenomena do not require the introduction of new fundamental laws, they may still require in many cases the introduction of further levels of explanation above the physical level in order to make these phenomena maximally comprehensible to us. Further, by showing how a simple starting point can have unexpected consequences, the existence of weakly emergent phenomena can… accommodate all sorts of unexpected richness at higher levels, as long as explanations are given at the appropriate level. (Chalmers 2006)

We see the potential for applying weak emergence to cognitive modeling because it promises to give insight into multiple layers of explanation and select the appropriate structures to include in models. Additionally, emergence naturally seems to account for some of the complexity typically modeled in a dynamic systems paradigm: the higher-level structures key to an emergent theory may emerge from dynamic-type interactions at lower levels. A main problem with emergence at present, however, is the lack of a mature theory of the subject. Strong and weak emergence, as philosophical concepts, are more well-understood than is how to apply emergence to a modeling task. For some hints as how such a theory could proceed to develop, we turn to Holland and Clark.

In his study of complex adaptive systems, Holland (1988) provides many ideas about how to study weakly emergent complex behavior. The link between lower and higher cognitive levels, such as a link between connectionism and PSS, can be viewed as understanding how symbols can emerge from the interaction of vast numbers of neurons. Both Holland and Clark (2001) emphasize the role of *persistent patterns* in the study of emergence. Holland writes that "Only persistent patters will have a directly traceable influence on future configurations in generated systems… the persistent patterns are the only ones that lend themselves to a consistent observable ontogeny."

Persistent patterns provide a first step in our search for an application-oriented theory of emergence. In a complex system, if patterns emerge, they can be used as the basis for prediction. A model that accounts for higher-level patterns can thus escape the facsimile trap and generate predictions without simply simulating every small detail of the original system. These emergent patterns may pave the way towards bridge-building – if connectionism and a higher-level theory like PSS are to be consistent, the emergent patterns of a connectionist network will likely match up with the higher-level symbols.

One worry about prediction in a complex system is the worry that emergent phenomena are "uncompressible" in terms of predictable patterns, that is, they are "those phenomena for which prediction requires simulation." (Clark 2001) Clark counters this pessimistic attitude, however, writing that

> My intuition, by contrast is that emergent phenomena are often *precisely* those phenomena in which complex interactions yield robust, salient patterns capable of supporting prediction and explanation, i.e., that lend themselves to various forms of low-dimensional projection.

There is, then, room for optimism. However, it is worth noting that a viable, predictive theory of emergent behavior in complex systems is still a distant goal. Holland describes a general interdisciplinary roadmap for studying emergence, in which he first advocates simply collecting and studying examples of the phenomenon in order to look for underlying principles at work in different contexts. He writes that "the difficulty, it seems to me, stems more from the daunting diversity of emergent phenomena. Like consciousness, life, or energy, emergence is ever present, but protean in form." (Holland 1998) Given a large collection of examples of emergence, he advocates a special form of reduction where complex systems are explained in terms of *interactions* between simpler components. The specific goal is to use the collected examples of such interactions to discover and control "the essential conditions for emergence." (Holland 1998) He goes on to speculate about future theorems that a theory of emergence might look for:

> To build a competent theory one needs deep insights, not only to select a productive, rigorous framework (a set of mechanisms and the constraints on their interactions), but also to conjecture about theorems that might be true (conjectures, say, about *lever points* that allow large, controlled changes in aggregate behavior through limited local action.) (Holland 1998)

Although the very existence of such a theory is speculative, something like a theory of emergence would likely help select the appropriate objects to embed as primitives in any given modeling context (persistent patterns seem likely candidates). It would also help build bridges between granularity levels (the idea of an emergent behavior is indeed precisely this sort of level connection). Similarly, emergent patterns could provide key links between different modeling paradigms.

Given the current disparate approaches to modeling in cognitive science, it can be difficult to reconcile the differences as we search for a more fundamental unifying cognitive theory. The geologist Thomas Chamberlin proposed the *method of multiple working hypotheses* (Chamberlin 1965) to provide guidance in this sort of the search for clarity. To gain scientific understanding of a phenomenon in the face of multiple explanations, this method advocates the common-sense notion that instead of viewing the world through a single "ruling theory" or single hypothesis, one should instead keep a set of possible hypotheses open and under active consideration. Emergence may provide a useful supplement to Chamberlin's method as a scientific tool for investigating relationships between compatible multiple working hypotheses and different modeling paradigms. For instance, demonstrated emergence of high-level symbols from a connectionist substrate might help unify the multiple working PSS and connectionist paradigms. Ideally, such unification would result in the refinement of both theories in order to make them compatible enough for application of the tools of emergence theory. Even though applied emergence is still in its infancy, it is imperative that cognitive science embrace and support the development of such a pioneering theory, as the lack of bridges between levels and modeling paradigms will remain a serious obstacle to cognitive science for the foreseeable future.

References

Beer, R. “Dynamical approaches to cognitive science.” *Trends in Cognitive Sciences*, **4**(*3*), March 2000, pp 91-98.

Chalmers, D. (2006). “Strong and Weak Emergence,” in Clayton and Davies (2006).

Chamberlin, T. (1965). “The method of multiple working hypotheses.” *Science*, **148**, May, pp. 754-759.

Clark, A. (2001). *Mindware*. New York: Oxford University Press.

Clayton, Philip, and Davies, Paul, ed., (2006). *The Re-Emergence of Emergence*, Oxford: Oxford University Press.

Graham, G. "Behaviorism", *The Stanford Encyclopedia of Philosophy (Fall 2005 Edition)*, Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/fall2005/entries/behaviorism/>.

Hofstadter, D. R. (1986). “Impressions of AI”. Internal memo, University of Michigan.

Holland, J. (1998). *Emergence*. New York: Basic Books.

Holland, J. (2002). What Is to Come and How to Predict It. In Brockman, J. (ed.), *The Next Fifty Years*. New York: Vintage Books.

Marr, D. (1982). *Vision*. San Francisco: W.H. Freeman.

Newell, A. and Simon, H. (1976). In Haugeland, J. (ed.), *Mind Design II*, Cambridge, MA: MIT Press, 1997, pp. 81-110.

O'Connor, T. (1994). “Emergent Properties”. *American Philosophical Quarterly*, **31**, pp. 91-104.

O'Connor, T., Wong, H. Y. "Emergent Properties", *The Stanford Encyclopedia of Philosophy* (Winter 2006 Edition), Edward N. Zalta (ed.), forthcoming URL = <http://plato.stanford.edu/archives/win2006/entries/properties-emergent/>.