UC Merced
Proceedings of the Annual Meeting of the Cognitive Science Society

Title
Models of Cognition: Neurological Possiblity Does Not Indicate Neurological Plausibility

Permalink
https://escholarship.org/uc/item/26t9426g

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 27(27)

ISSN
1069-7977

Authors
Kriete, Trenton E.
Noelle, David C.

Publication Date
2005

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

Models of Cognition: Neurological possibility does not indicate neurological plausibility.

Peter R. Krebs (peterk@cse.unsw.edu.au)
Cognitive Science Program
School of History & Philosophy of Science
The University of New South Wales
Sydney, NSW 2052, Australia

Abstract

Many activities in Cognitive Science involve complex computer models and simulations of both theoretical and real entities. Artificial Intelligence and the study of artificial neural nets in particular, are seen as major contributors in the quest for understanding the human mind. Computational models serve as objects of experimentation, and results from these virtual experiments are tacitly included in the framework of empirical science. Cognitive functions, like learning to speak, or discovering syntactical structures in language, have been modeled and these models are the basis for many claims about human cognitive capacities. Artificial neural nets (ANNs) have had some successes in the field of Artificial Intelligence, but the results from experiments with simple ANNs may have little value in explaining cognitive functions. The problem seems to be in relating cognitive concepts that belong in the 'top-down' approach to models grounded in the 'bottom-up' connectionist methodology. Merging the two fundamentally different paradigms within a single model can obfuscate what is really modeled. When the tools (simple artificial neural networks) to solve the problems (explaining aspects of higher cognitive functions) are mismatched, models with little value in terms of explaining functions of the human mind are produced. The ability to learn functions from data-points makes ANNs very attractive analytical tools. These tools can be developed into valuable models, if the data is adequate and a meaningful interpretation of the data is possible. The problem is, that with appropriate data and labels that fit the desired level of description, almost *any* function can be modeled. It is my argument that small networks offer a *universal* framework for modeling any conceivable cognitive theory, so that neurological *possibility* can be demonstrated easily with relatively simple models. However, a model demonstrating the possibility of implementation of a cognitive function using a distributed methodology, does not necessarily add support to any claims or assumptions that the cognitive function in question, is neurologically *plausible*.

Introduction

Several classes of computational model and simulation (CMS) used in Cognitive Science share common approaches and methods. One of these classes involves artificial neural nets (ANNs) with small numbers of nodes, particularly feed forward networks (Fig. 1) and simple recurrent networks (SRNs)[1] (Fig. 2). Both of these architectures have been employed to model high level cognitive functions like the detection of syntactic and semantic features for words (Elman, 1990, 1993), learning the past tense of English verbs (Rumelhart and McClelland, 1996), or cognitive development (McLeod et al., 1998; Schultz, 2003). SRNs have even been suggested as a suitable platform "toward a cognitive neurobiology of the moral virtues" (Churchland, 1998). While some of the models go back a decade or more, there is still great interest in some of these 'classics', and similar models are still being developed, e.g. Rogers and McClelland (2004). I argue that many models in this class explain little at the neurological level about the theories they are designed to support, however I do not intend to offer a critique of connectionism following Fodor and Pylyshyn (1988).

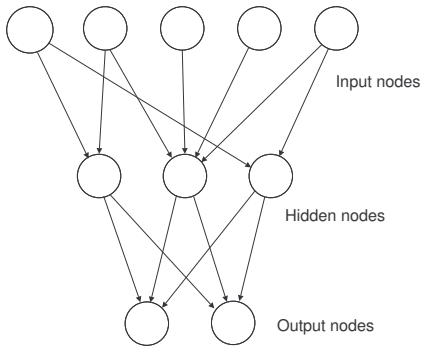

Figure 1. Feed forward network architecture

Instead, this paper concerns models where ANNs act merely as mathematical, or analytical, tools. The fact that mathematical functions can be extracted from a given set of data, and that these functions can be successfully approximated by an ANN (neurological possibility), does not provide any evidence that these functions are capable of being realized in similar fashion inside human brains (neurological plausibility).

Bridging the Paradigms

Theories in Cognitive Science fall generally into two distinct categories. Some theories are offered as explanations of aspects of human cognition in terms of *what*

[1]SRNs have a set of nodes that feed some or all of the previous states of the hidden nodes back. The nodes are often described as *context* nodes. They provide a kind of 'short term memory' that becomes part of the input in the next step of the simulation.

brains do, and the implementation at neural level is usually of little concern. Arguably, these theories are *all* about psychological phenomena and the physical brain should not even be considered in the context of this approach. Bennett and Hacker (2003), for example, believe that

> [...] it makes no sense to ascribe such psychological functions [i.e. perceiving and thinking] to anything less than the animal as a whole. It is the animal that perceives, not parts of its brain, and it is human beings who think and reason, not their brains. The brain and its activities *make it possible* for *us* - not for *it* - to perceive and think, to feel emotions, and to form and pursue projects (Bennett and Hacker, 2003, 3).

Essentially, the *top-down*[2] approach deals with high level cognitive functions, and the brain, or entire being, is viewed as a single black box, or a collection of black boxes with certain *functional* properties.

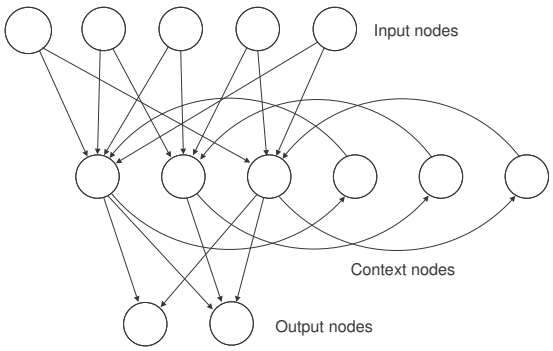

Figure 2. SRN architecture

The *bottom-up* approach, in contrast, deals with the base elements, namely neurons, and their physiological and functional properties and processes. Functional aspects of brains, or parts of brains, are investigated by looking at individual neurons and structures of groups and networks of neurons. Cognitive Neuroscience and some work in Artificial Intelligence is concerned with *how* cognitive functions might be implemented in brains.
Currently, methods are explored to connect the *top-down* and the *bottom-up* approaches in attempts to ground high-level psychological phenomena in neuro-physiology. One such research program concerns the mapping or localizing of cognitive functions in the brain. Modern technologies such as PET and fMRI[3] are commonly used for that purpose although there are many technical and conceptual issues unresolved[4].
Other attempts to bridge the divide between the two paradigms involve computational models which aim to explain how higher cognitive functions could possibly be supported by a distributed architecture. In order to achieve this, descriptive elements from different levels are brought together in an attempt to present unified and coherent CMSs of cognitive processes. In the case of models that are based on simple feed forward ANNs and SRNs, theoretical and conceptual elements are subjected to a set of neurologically inspired *mathematical* tools. An important contribution to the apparent success of these models is that the analysis and interpretation of experimental results can be framed in the language of the theoretical and conceptual entities concerning the cognitive function. Building models using ANNs is not a difficult task, particularly if the ANN is small, because many of the technical and methodological details need not to be dealt with[5].

A Universal Framework

Artificial neural networks are trained using algorithms that adjust the weights between units, i.e. model neurons, so that the error between the ANN's computed output and the expected output is minimized for the given input. This process is repeated for all possible input-output pairs many times over. For example, to implement the *XOR*-function

$$O_n = (I_1 \wedge \neg I_2) \vee (\neg I_1 \wedge I_2)$$

the network will be presented with values for I_1 and I_2, i.e. '0, 0', '0, 1', '1, 0' and '1, 1'. The weights are adjusted using an appropriate algorithm to minimize the error between the network's output and the output of the training set, i.e. '0', '1', '1', and '0' respectively. Once the network is trained, it will compute the output O_n from the inputs I_1 and I_2 according to the *XOR*-function. In many discussions about ANNs in the context of cognitive modeling, the inputs are labeled with terms other than '0's or '1's. Because we can use these labels freely, there is always the danger of introducing 'wishful' terminology not only for labels, but also for methodological terms. But, as Fodor and Pylyshyn (1988) have pointed out,

> [...] the labels play *no role at all* in determining the operation of a Connectionist machine; in particular, the operation of the machine is unaffected by the syntactic and semantic relations that hold among the expressions that are used as labels. To

[2] I am using the terms *top-down* and *bottom-up* in favor of *high-level* and *low-level* to suggest that these are not static research programs, but that they are dynamic endeavors aiming to close the divide between them.

[3] Positron Emission Tomography (PET) and functional Magnetic Resonance Imaging (fMRI) are based on the assumption that mental activity causes an increase in the metabolic rate of neurons, and therefore an increase in the flow of blood. PET detects the locations where positrons are emitted from decaying atoms of a radioactive tracer (typically H_2O^{15}), while fMRI detects different levels of oxygenated and deoxygenated hemoglobin.

[4] See for example Uttal (2001) for a discussion of issues surrounding current methods in mapping cognitive functions onto the brain.

[5] Many computer programs that implement various ANNs are freely available, and little technical expertise is required to use them.

> put this another way, the node labels in a Connectionist machine are not part of the causal structure of the machine (Fodor and Pylyshyn, 1988, 13).

Only the activation levels of the input nodes and the connection strengths in the network matter for an ANN to produce the appropriate output for the function it is trained to approximate. Nevertheless, labels for the nodes and terminology for other parts of the networks are introduced whenever models are constructed. Schultz (2003), for example, maps terms from neural nets onto terms from developmental psychology (here, Piagetian theory).

> *Accommodation*, in turn, can be mapped to connection-weight adjustment, as it occurs, in the output phase of cascade-correlation learning. [...] More substantial qualitative changes, corresponding to *reflective abstraction*, occur as new hidden units are recruited into the network. [...] Then the network reverts back to an output phase in which it tries to incorporate the newly achieved representations of the recruited unit into a better overall solution. This, of course, could correspond to Piaget's notion of *reflection* (Schultz, 2003, 128, original italics).

The terminology from Piagetian theory clearly belongs to a higher level of description than the true descriptions of the network's structure and dynamics.

Churchland (1998) suggests that a recurrent network could model more challenging cognitive functions. He considers that a recurrent network may have appropriate architecture for simulating the acquisition of moral virtues in humans. He argues that a network would be able to map concepts like *cheating*, *tormenting*, *lying*, or *self-sacrifice* within a n-space of classes containing dimensions of *morally significant*, *morally bad*, or *morally praiseworthy* actions, by learning through "repeated exposure to, or practice of, various *examples* of perceptual or motor categories at issue" (Churchland, 1998, 83). Churchland says that

> [t]his high-dimensional similarity space [...] displays a structured family of categorical "hot spots" or "prototype position", to which actual sensory inputs are assimilated with varying degree of closeness (Churchland, 1998, 83).

It is beyond the scope of this paper to discuss whether a model of such a calculus of moral virtues is appropriate, but Churchland certainly demonstrates that it *can* at least in principle be modeled with an ANN. However, feed forward networks and SRNs with suitable numbers of inputs and outputs and a reasonable number of hidden units[6] can be trained to implement almost *any* function. The point here is that a network can implement almost anything we want to model, as long as we have the appropriate training set for the particular selection of labels for the inputs and outputs of the ANN. As a result we will have to accept that even simple ANNs provide a *universal*, but uninformative, framework for cognitive models.

There is a further methodological issue to consider. It may be surprising to learn that neural nets in some models are not necessarily composed of neurons. Elman et al. (1998) offer as a "note of caution" that

> ... [m]ost modelers who study higher-level cognitive processes tend to view the nodes in their models as equivalent not to single neurons but to larger populations of cells. The nodes in these models are *functional* units rather than *anatomical* units (Elman et al., 1998, 91).

Can models still be considered as 'bottom-up' neural nets, if they are composed of functional units? I suggest that such models do not belong in the realm of connectionism, because the replacement of model neurons with "functional units" re-introduces exactly those black boxes that we trying to eliminate in the 'bottom-up' approach.

The kinds of models that I am describing here, i.e. simple feed forward ANNs and small SRNs, do not rely on special neural 'circuitry', unlike *structured models* in which the models' architectures reflect a particular part of brain physiology. The architectures are universal in the sense that only the number of neurons and connections vary from model to model. The diversity of models that have been described in the literature is the product of applying ANNs as analytical tools to a diverse set of problems where suitable data sets for training of the networks are available. The universal architecture and the freedom to choose labels and terminology fitting the particular model explains the proliferation of ANN inspired models. Traditional mathematical (symbol based) models may be more constrained as far as the selection of representations is concerned[7]. How then are explanatory links maintained between representations in distributed models and real world phenomena?

Symbols and Representations

Classic CMS are representational systems using symbols, which carry arbitrarily assigned semantic content. Haugeland (1985, 1981) and others have argued that these semantics remain *meaningful* during processing, as long as the syntactical structure is appropriate and suitably maintained. Haugeland notes that in an interpreted formal system with true axioms and truth-preserving rules, the semantics will take care of itself, if you take care of the syntax (Haugeland, 1981). The symbol '5', for example, carries different semantics in a positional number system. Whether '5' means '500' in '1526', or '50' in '1257' is a function that is governed by the syntactic and semantic rules of the number system. Tying

[6] Note that the number of hidden nodes and the number of connections within the network are largely determined by experience and experiment. There are no definite methods or algorithms for this.

[7] Dretske (1981, 1988), among others, has dealt with questions of representations and their semantics in representational systems.

semantics to symbols is much more problematic in connectionist models, because it is part of the connectionist doctrine that representations (symbols) are distributed in the structures of neural nets. It is therefore more difficult to produce a trace of what happens semantically in an ANN, because no syntactical structure exists.

ANNs are usually described as having distinct and discrete inputs and outputs[8], each labeled as having a distinct and discrete meaning. Such labels may be words, like *boy*, *girl*, *read*, *book*, or, the labels may be concepts such as *phonemes*, or *visual inputs*. Such labels have their own set of problems associated with them. Attaching the value '*grandmother*' to one of the input nodes illustrates my concern. While nearly everyone rejects the existence of a *grandmother-neuron* in the brain as a rather naïve concept, *boy-*, *girl-*, or *book-* neurons are willingly accepted in models.

Localized representations are no longer available once the focus shifts on to hidden nodes within the network, and the 'representations' are now described in terms of weights, or synaptic strengths, between individual units. However, for a meaningful interpretation of the network and its dynamics, it is necessary to convey content and meaning in terms of non-distributed (localized) symbols, because is not sufficient for a discussion of what goes on in ANNs to assign semantic content merely to inputs and outputs. In order to track the flow of information through the networks, some *descriptions* are needed, because explaining the processes in the ANN in terms of connection-weights between neurons is *tedious* and *unsuitable* for the kinds of models in question. Discussing representations in terms of connection weights is tedious, because the number of connections can be considerable, even in small networks[9]. A distributed representation R, i.e. the activation pattern for a particular input $I_{1\ldots k}$, could be specified in the form of a matrix, or as a vector, with as many elements as there are connections in the network.

$$R(I_{1\ldots k}) = (.8234, .9872, .1290, \ldots, .0012).$$

In any case, it it necessary to specify all of the numeric values to capture every single activation pattern. Representations and descriptions in this form are unsuitable, because they reveal little in terms of the cognitive function that is modeled. Where do new and helpful descriptions come from?

Interpreting models

The representations for words, concepts, phonemes, visual inputs, and so on, are usually coded in binary, or as real values, in paired input and output vectors in the training set for the ANN. During the training the *relationships* between the input and output vectors are *encoded* in the hidden layers of the ANN, or as Fodor and Pylyshyn (1988) put it, "the weights among connections are adjusted until the system's behavior comes to model the *statistical properties of its inputs*" (my italics).

Elman (1990), for example, presented 29 words in the human language one at a time to a simple recurrent network in the form of binary vectors $I_1 \ldots I_n$, such that a single bit represented a particular word. The words themselves were presented in sequences forming two and three word sentences that had been generated according to a set of 15 fixed templates. A cluster analysis of the hidden nodes revealed that the trained network exhibits similar activation patterns for inputs (words) according to their relative position in the sequence (sentence) and their probability of occurring in relation to other words. The analysis of these activation patterns allowed for the classification of inputs into categories like *nouns* or *verbs*. Moreover, the categories of internal representations could be broken down into smaller groups like *human*, *non-human*, *large animals*, or *edibles*, and so on.

Cluster analysis is used as a method to gain insights into the internal representations of ANNs, but is not without some conceptual problems. Clark (2001) argues that cluster analysis is an analytic technique to provide answers to the crucial question of what kinds of representations the network has acquired. However, cluster analysis does not reveal anything that is not already contained in the raw data of the model. The relationships and patterns in the input datasets and training datasets become embedded in the structure of the network during training[10]. What counts are the mathematical and statistical relations that are contained in the *training* datasets. In many cases the relations may just be tacitly accepted. In other models these relations are purposefully introduced from the outset. Under these conditions, the relations are part of the model's *design*. Elman (1990), for example, states that "13 classes of nouns and verbs were chosen" for generating the datasets. Whether the relations in the data are introduced by design, or whether the experimenter is unaware of these statistical artifacts, there should be no surprise that the analysis will reveal these relations later during the experiment. The implementation of a model as an ANN and the subsequent extraction of results that are already in the data may have little value in terms of obtaining *empirical* evidence. The training set of pairs of input and output vectors already contains all there is to the model, and the ANN does not add anything that could not be extracted from the training sets through *other* mathematical or computational methods.

Green (2001) argues that

> these results are just as analytic as are the results of a mathematical derivation; indeed they *are* just mathematical derivation. It is *logically* not possible that [the results] could have turned out other than they did (Green, 2001, 109).

[8] Inputs and outputs of ANNs can also have continuous values. The kinds of models I am discussing here have typically discrete values.

[9] A fully connected feed forward network with 20 input nodes, 10 hidden nodes, and 5 output nodes has 250 connections.

[10] The patterns and relationships in these datasets can either be carefully designed or might be an unwanted by-product.

A trained ANN implements a mapping from the input nodes (I_{1-n}) to the output nodes (O_{1-i}). The power of the ANN is in its ability to implement some function

$$O_{1-i} = f(I_{1-n})$$

from the training data set. Hoffmann (1998) emphasizes this point and says that

> [t]he greatest interest in neural nets, from a practical point of view, can be found in engineering, where high-dimensional continuous functions need to be computed and approximated on the basis of a number of data points (Hoffmann, 1998, 157).

The modeler does not need to specify the function f, in fact, the modeler does not even need to know anything about f. Knowledge extraction (KE) from ANNs is concerned with providing a description of the function f that is approximated by the trained ANN. The extraction of the function "lies in the desire to have explanatory capabilities besides the pure performance" (Hoffmann, 1998, 155). The ability to determine f may or may not add to the explanatory value of the model.
For moderately sized networks and relatively simple functions it is quite feasible to describe the model in a series of simple logic statements or with some high level programming language. In this step by step description of the network in terms of its input-output relations, knowledge of the function that will be ultimately implemented is not necessary. A particular relationship could be expressed, albeit awkwardly, in the form

```
if I1 = 0 and I2 = 1 then On = 1
else if I1 = 1 and I2 = 0 then On = 1
else if I1 = 1 and I2 = 1 then On = 0
else if I1 = 0 and I2 = 0 then On = 0
```

The simple *XOR*-function is easily recognized in this example. This approach may not be practical for more complicated functions, however it would be possible in principle. ANNs can even offer a convenient way of implementing complicated functions approximately, if some data points of these functions are known. The number of data points that are available for the training of the network determine how close the approximation of the functions can be, unless the function is known to be linear. The ability to process even relatively large data sets make ANNs valuable analytical tools to reveal something about the data. Even employing KE methods that may help to determine the function f does not overcome the limitation that the ANN cannot deliver anything new for the cognitive model. Regression analysis (curve fitting) performed on the training dataset will provide a more exact description of f than to teach an ANN and to perform KE subsequently.
There is a further complication as the data revealed in the cluster analysis is not accessible within the model. In other words, the results of the analysis are not furnished by the ANN. Rather, they are *interpretations* of the internal structures at a different level of description. The *actual* role of the network is that of a predictor, where the trained network attempts to guess the next output following the current input[11]. The analysis of the experiment is framed in the language of the higher cognitive function that is the subject of the model. For the interpretation and the analysis of the results, the output nodes are neglected and new 'output' for the model is generated by methods that belong to a higher level of description than the ANN. New 'insights' are synthesized from distributed representations by means and methods external to the ANN. Figure 3 illustrates the disconnectedness of the ANN and the newly gained 'insights' emerging from the network's internal representations. The experimenter performs the task of extracting information about the activation pattern using a new tool, cluster analysis for example, however the network has no part in this - ANNs do not perform cluster analysis. The work is clearly performed by the *modeler's* neurons with the aid of a statistical procedure and not by the *model's* neural structure.

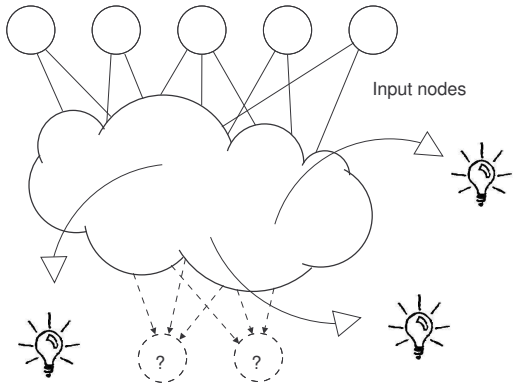

Figure 3. Actual model architecture

If the ANN is meant to be a model of what might happen at the neural level, then the question arises, what mechanism could be responsible for the equivalent (cluster) analysis of activation patterns in the brain? In order to make this information accessible to the rest of the brain, we will have to introduce some *other* neural circuit to do such an analysis of the hidden nodes. Such a new addition to the network could possibly categorize words into *verbs* and *nouns*, but then we need another circuit to categorize words into *humans*, *non-humans*, *inanimates*, or *edibles*, and another to categorize words into *monosyllabic* and *multi-syllabic*. In fact, we will need an very large number of neural circuits just for the analysis of word categories, provided the training dataset contains the appropriate relations to allow for such categorizations.
The class of simple ANNs that I have discussed here cannot provide any new 'insights' in any meaningful symbolic[12] or coded form on some output nodes. This, however, would have to be a crucial function of the model

[11] The desired output, which follows the input in the training set, is used as the target to determine the error for back propagation during the training phase.

[12] I do not think that 'distributed' or 'sub-symbolic' representations are helpful here. Moreover, this alternative ap-

to be considered neurologically *plausible*. For a model to be neurologically plausible, it would need to deduce new information about itself. More importantly, it would be necessary to signal the newly obtained knowledge to other neurons by changing the state of some nodes. Both cluster analysis and current methods of KE clearly fail to do this, although more recent developments in KE can deliver much more accurate description of f. However the renewed and possibly accurate synthesis of relations that were present in a training dataset does not warrant claims that the ANN 'discovered', 'learned', or 'recognized' something or other, even if these relations were not evident to the experimenter before. The ability to determine a function f that is contained in some dataset illustrates the power of ANNs as analytical tools. However, it should be clear that a different analytical tool could also have been used to detect the function f. We must conclude then that the model has failed to explain any processes at the neural level. Instead, the network model has only succeeded in offering an alternative method to encode the data, and the cluster analysis provides an alternative method to analyze the data.

Conclusions

Computational models and simulations, and models using ANNs in particular, are commonly used in support of theories about aspects of human cognition. Some models deal with high level psychological functions where the operations at the neural level are of little interest, and some models are concerned with the implementation of cognitive functions at neural level. I have argued that neurological *possibility* can be demonstrated for nearly any conceivable psychological theory due to the universality of simple ANNs. However using the language and symbolism of neural nets does not support any claims for neurological plausibility. The mistake, I believe, is to bring the top-down psychological model and the bottom-up neural environment together and to treat the result as a coherent and meaningful demonstration. ANNs can be used successfully as models, provided a clear description of the aims, assumptions and claims are presented. However, when simple ANNs with small numbers of nodes are employed to model complex high level cognitive functions, the experimenter should evaluate whether the simplicity of the network can provide a *plausible* implementation, because it is all too easy to provide a neurologically *possible* model.

Acknowledgments

I would like to thank Anthony Corones for comments and valued suggestions on earlier drafts of this paper.

References

Bennett, M. R. and Hacker, P. M. S. (2003). *Philosophical Foundations of Neuroscience*. Malden, Mass.: Blackwell.

Branquinho, J. (2001). *The foundations of cognitive science*. New York: Oxford UP.

Churchland, P. M. (1998). *Toward a Cognitive Neurobiology of the Moral Virtues*. In Branquinho (2001).

Clark, A. (2001). *Mindware: An Introduction to the Philosophy of Cognitive Science*. New York: Oxford UP.

Cummins, R. and Delarosa Cummins, D. (2000). *Minds, Brains, and Computers: The Foundations of Cognitive Science*. Malden, Mass.: Blackwell.

Dretske, F. (1981). *Knowlegde & the Flow of Information*. Cambridge, Mass.: MIT Press.

Dretske, F. (1988). *Representational Systems*. In O'Connor and Robb (2003).

Elman, J. (1993). Learning and development in neural networks: The importance of starting small. *Cognition*, 48:71–99.

Elman, J. L. (1990). Finding structure in time. *Cognitive Science*, 14:179–211.

Elman, J. L., Bates, E. A., Karmiloff-Smith, A., Parisi, D., and Plunkett, K. (1998). *Rethinking Innateness: A Connectionist Perspective on Development*. Cambridge, Mass.: MIT Press.

Fodor, J. and Pylyshyn, Z. (1988). Connectionism and cogintive architecture: A critical analysis. *Cognition*, 28:3–71.

Green, C. D. (2001). Scientific models, connectionist networks, and cognitive science. *Theory & Psychology*, 11(1):97–117.

Haugeland, J. (1981). *Semantic Engines: An Introduction to Mind Design*. In Cummins and Delarosa Cummins (2000).

Haugeland, J. (1985). *Artificial Intelligence: the Very Idea*. Cambridge, Mass.: MIT Press.

Hoffmann, A. (1998). *Paradigms of Artificial Intelligence*. Singapore: Springer.

McLeod, P., Plunkett, K., and Rolls, E. T. (1998). *Introduction to Connectionist Modelling of Cognitive Processes*. Oxford: Oxford UP.

O'Connor, T. and Robb, D. (2003). *Philosophy of Mind: Contemporary Readings*. London: Routledge.

Rogers, T. T. and McClelland, J. L. (2004). *Semantic Cognition: A Parallel Distributed Processing Approach*. Cambridge, Mass.: MIT Press.

Rumelhart, D. E. and McClelland, J. L. (1996). *On Learning the Past Tense of English Verbs*. In Cummins and Delarosa Cummins (2000).

Schultz, T. R. (2003). *Computational Developmental Psychology*. Cambridge, Mass.: MIT Press.

Uttal, W. R. (2001). *The New Phrenology: The Limits of Localizing Processes in the Brain*. Cambridge, Mass.: MIT Press.

proach of dealing with network structures is usually not employed by modelers either.