

AD-A218 941

ISSUES IN SYMBOL PROCESSING

Technical Report - 19

David S. Touretzky

Computer Science Department
Carnegie Mellon University
Pittsburgh, PA 15213

29 September 1987

The Artificial Intelligence and Psychology Project

Departments of
Computer Science and Psychology
Carnegie Mellon University

Learning Research and Development Center
University of Pittsburgh

DTIC
ELECTE
MAR 14 1990
S B D

Approved for public release; distribution unlimited.

90 03 12 047

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION Unclassified		1b. RESTRICTIVE MARKINGS	
2a. SECURITY CLASSIFICATION AUTHORITY		3. DISTRIBUTION / AVAILABILITY OF REPORT Approved for public release; Distribution unlimited	
2b. DECLASSIFICATION / DOWNGRADING SCHEDULE			
4. PERFORMING ORGANIZATION REPORT NUMBER(S) AIP - 19		5. MONITORING ORGANIZATION REPORT NUMBER(S)	
6a. NAME OF PERFORMING ORGANIZATION Carnegie-Mellon University	6b. OFFICE SYMBOL (if applicable)	7a. NAME OF MONITORING ORGANIZATION Computer Sciences Division Office of Naval Research (Code 1113)	
6c. ADDRESS (City, State, and ZIP Code) Department of Psychology Pittsburgh, Pennsylvania 15213		7b. ADDRESS (City, State, and ZIP Code) 800 N. Quincy Street Arlington, Virginia 22217-5000	
8a. NAME OF FUNDING / SPONSORING ORGANIZATION Same as Monitoring Organization	8b. OFFICE SYMBOL (if applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER N00014-86-K-0678	
8c. ADDRESS (City, State, and ZIP Code)		10. SOURCE OF FUNDING NUMBERS p400005up101; 7-4-86	
		PROGRAM ELEMENT NO N/A	PROJECT NO N/A
		TASK NO. N/A	WORK UNIT ACCESSION NO N/A
11. TITLE (Include Security Classification) ISSUES IN SYMBOL PROCESSING			
12. PERSONAL AUTHOR(S) David S. Touretzky			
13a. TYPE OF REPORT Technical	13b. TIME COVERED FROM 86Sept15 to 91Sept15	14. DATE OF REPORT (Year, Month, Day) 87 September 29	15. PAGE COUNT 5
16. SUPPLEMENTARY NOTATION			
17. COSATI CODES		18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)	
FIELD	GROUP	SUB-GROUP	
		Artificial Intelligence, symbols, connectionism	
19. ABSTRACT (Continue on reverse if necessary and identify by block number) SEE REVERSE SIDE			
20. DISTRIBUTION / AVAILABILITY OF ABSTRACT ☐ UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT ☐ DTIC USERS		21. ABSTRACT SECURITY CLASSIFICATION	
22a. NAME OF RESPONSIBLE INDIVIDUAL Dr. Alan L. Meyrowitz		22b. TELEPHONE (Include Area Code) (202) 696-4303	22c. OFFICE SYMBOL N00014

ABSTRACT

The notion of mind as symbol processor is fundamental to AI and cognitive science, but some connectionists are now arguing against it. Eliminative connectionism challenges the validity of formal symbol manipulation as a level of mental description. We review some of the claims that have been made, and argue that most connectionist models, especially those constructed by learning algorithms, are operating at the level of pattern classifiers. Their (rather limited) success using non-symbolic representations demonstrates that they have not yet even approached the tasks which symbol processing models attempt to solve. Continued progress in connectionist research may require reimplementaion rather than rejection of the symbolic level.

Accession For	
NTIS GRA&I	☒
DTIC TAB	☐
Unannounced	☐
Justification	
By	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A-1	

Issues in Symbol Processing

David S. Touretzky
Computer Science Department
Carnegie Mellon University
Pittsburgh, PA 15213

Abstract

The notion of mind as symbol processor is fundamental to AI and cognitive science, but some connectionists are now arguing against it. Eliminative connectionism challenges the validity of formal symbol manipulation as a level of mental description. We review some of the claims that have been made, and argue that most connectionist models, especially those constructed by learning algorithms, are operating at the level of pattern classifiers. Their (rather limited) success using non-symbolic representations demonstrates that they have not yet even approached the tasks which symbol processing models attempt to solve. Continued progress in connectionist research may require reimplementation rather than rejection of the symbolic level.

1. Introduction

The first thirty years of AI research proceeded on the entirely plausible assumption that the mind was a symbol processor. This idea has recently been challenged by a group of cognitive scientists known as connectionists, who construct neural network models which they claim have no equivalent description at the formal symbolic level. Rumelhart and McClelland (1986a) and McClelland and Rumelhart (1986) provide a good introduction to the connectionist approach, also known as "parallel distributed processing."

The controversy surrounding connectionism has been heating up recently in response to three important papers. Smolensky (in press) provides the definitive statement of the hypotheses underlying connectionism. Pinker and Prince (1987) and Fodor and Pylyshyn (1987) challenge the connectionist view, both by attacking general claims that others have made about these networks and by criticizing particular models that have appeared in the literature. A review of the current status of connectionist symbol processing will allow the interested reader to follow the debate as it unfolds.

As in Pinker and Prince (1987, pp. 4-7), we divide connectionism into three schools. "Implementational connectionism" is concerned with how massively parallel archi-

tectures might implement the classical notion of symbol processing. Examples include Touretzky and Hinton's neural network production system interpreter (Touretzky and Hinton, 1985) and Ballard's connectionist implementation of resolution (Ballard, 1986). "Eliminative connectionism," on the other hand, denies the validity of symbolic-level descriptions. (Smolensky's paper serves as a sort of "Eliminativist Manifesto.") Eliminativists believe that to accurately explain what goes on in the mind one must shift to a hypothetical sub-symbolic level, more abstract than the neural level, but also fundamentally different from (and not merely an implementation of) symbolic-level computation. Finally, "revisionist-symbol-processing connectionism" is suggested by Pinker and Prince as a middle ground where discoveries might lead to fundamental changes in our understanding of symbol processing without forcing us to abandon the classical commitment to symbolic-level descriptions.

This paper is primarily concerned with the eliminative position, not because it is preferred, but because as the most radical version of connectionism it is most at odds with the classical account of intelligence.

2. The Symbolic-Level Paradigm

The symbolic-level paradigm underlies most research in AI and cognitive science. Language, commonsense reasoning, and conscious problem solving can all be described at this level in terms of structures, composed of symbols, that are manipulated by formal rules. Parse trees, semantic nets, frames, scripts, and axiom sets are examples of composite symbol structures. Their manipulation can be specified in various ways, some of which are computational, such as production rules, theorem provers, and Lisp functions, and others which are descriptive but not necessarily computational, such as the rules linguists write to describe syntactic or phonological processes.

The claim made by the symbolic-level paradigm is that intelligent behavior can be adequately explained purely in terms of formal operations on symbol structures. Newell (1980a) calls this the Physical Symbol System Hypothesis.



In other words, the mind contains symbol structures for concepts, goals, intentions, memories, and so forth, and intelligence derives from the effective manipulation of these structures. Eliminative connectionism denies this. Before getting into what the connectionists would have in place of symbols and rules, I should emphasize a key point in the definition of the symbolic-level paradigm. It isn't just a claim that the mind works by manipulating symbols; it is a claim that the structures the mind manipulates can be directly identified with the elements of mental life: they are words, thoughts, percepts; not arbitrary, meaningless atoms.

Consider a thermostat with setpoint T_0 whose behavior is governed by the following rule:

```
IF  $T < T_0$ 
  THEN turn-on(furnace)
  ELSE turn-off(furnace)
```

This rule constitutes a symbolic-level theory of thermostats. It is expressed in terms of ambient temperature, setpoint, and furnace activity: the language of the thermostatic domain. It does not refer to the individual atoms that make up the thermostat, or to the motions of particles in the atmosphere of the room. It is a formal theory because it can be implemented, and it accurately predicts the thermostat's behavior. According to the symbolic-level paradigm, mental processes can also be explained by formal theories, without reference to phenomena such as neuron firings that exist only at a lower level of description.

3. Are There Rules?

If there are explicit representations of rules in the head, there must be an interpreter to execute the rules as thinking proceeds. Conscious problem solving behavior does appear to be rule-based. For example, John Anderson's ACT* model, which learns new production rules as it gains experience at tasks such as proving geometry theorems, offers a good account of how humans behave when performing the same tasks (Anderson, 1983; Anderson, in press). But it is important to distinguish between conscious, deliberate behavior and intuitive behavior. The latter is not explainable by introspection, nor is it decomposable into consciously-accessible steps such as occur in problem solving.

Intuitive-level phenomena certainly include such things as vision and motor control, which operate almost entirely below the level of conscious thought. Language and common sense reasoning also proceed largely at the subconscious level, and appear to be intuitive. Smolensky suggests that our linguistic facility actually serves as the rule interpreter for conscious problem solving, and that what we perceive as consciousness is a series of snapshots of the state of an intuitive processor that is not itself rule-based. Note, however, that rule-based behavior isn't necessarily conscious. Newell (1980b) used the production rule formalism to speculate about mental implementations of the Harpy speech understanding system.

In linguistics, the goal has been to explain phenomena

by deriving the most economical set of rules that account for the data. Linguists shy away from claiming that these formal rules, with their associated interpreter, are what is actually in the head (Stabler 1983; Thompson, 1983). However, models of linguistic development (as opposed to competence) are often phrased in terms of rule acquisition and revision, which may require an explicit representation for rules.

The symbolic-level paradigm is about the *description* of behavior in terms of symbols and rules; it says nothing about the explicit representation of rules. In the case of the thermostat, which clearly has no rule interpreter inside it, the physical structure of the device induces certain causal relationships between the ambient temperature, setpoint, and furnace activity which are accurately summarized by the rule we gave previously. That is all the symbolic-level paradigm requires.

4. Symbolic-Level Connectionism

Some connectionist models identify symbols with particular units. These are known as *localist* models, to distinguish them from the *distributed* models that are the focus of this paper. Cottrell (1985) and Waltz and Pollack (1985) use a localist representation in which individual units stand for words or word senses; Shastri (1985) uses units to denote classes and properties in an inheritance hierarchy; Selman (1985) and Fianty (1986) use units to denote input atoms and grammatical tokens in networks that parse context-free languages. Localist models may have interesting dynamical properties, e.g., when units denoting competing word senses inhibit each other, they are in effect "fighting" to settle on the most plausible interpretation of the input. In Pollack's model of the garden path sentence "The astronomer married the star," the ASTRONOMER unit causes the HEAVENLY-BODY unit to have a higher initial activation than MOVIE-STAR, but HEAVENLY-BODY eventually loses out due to constraints imposed by MARRIED, much as humans revise their initial interpretation the first time they hear the sentence.

Fodor and Pylyshyn's criticism of the localist approach focuses on the inability to compose symbols when they are tied directly to processing units. They point out that although one can designate individual units to stand for P, Q, and P&Q, the fact that P&Q *references* P cannot be expressed. An excitatory connection from P&Q to P would allow the network to infer P whenever P&Q is asserted to be true. But the network cannot decompose P&Q to get P, nor can it compose new structures such as P&Q&R from already existing ones. But Fodor and Pylyshyn go too far when they claim distributed models suffer the same difficulty; Touretzky (1986) and Touretzky and Geva (1987) provide counterexamples.

Most connectionists view the localist approach as a temporary compromise that allows them to conveniently explore certain dynamic constraint satisfaction phenomena. When the full power of the classical symbol processing model has been implemented in a distributed connection-

ist architecture, the localist approach may no longer be attractive.

5. The Sub-symbolic Paradigm

In distributed connectionist models, symbols and symbol structures are represented by patterns of activity over a collection of units, rather than by individual units. Symbols are then points in a high-dimensional metric space, with a natural similarity measure being the dot product. Although, as Fodor and Pylyshyn note (p. 58), one may impose arbitrary similarity measures on conventional symbol systems, in the connectionist case the similarity effects are rooted in the causal structure of the model.

A large class of distributed connectionist models are concerned with pattern classification or pattern transformation. For example, Sejnowski and Rosenberg's celebrated NETtalk model maps input patterns that represent a seven letter window of text to output patterns that represent a phoneme (Sejnowski and Rosenberg, 1987). Rumelhart and McClelland's verb learning model maps phonemic representations of present tense verbs to phonemic representations of the corresponding past tense, e.g., "hug" to "hugged" and "go" to "went" (Rumelhart and McClelland, 1986b). The weights in both of these models are derived by connectionist learning procedures from repetitive exposure to example inputs. Rumelhart and McClelland used a version of the perceptron learning algorithm, while Sejnowski and Rosenberg used the more recent back propagation algorithm of Rumelhart, Hinton, and Williams (1986).

Another large class of models perform constraint satisfaction by relaxation, such as Hopfield nets (Hopfield, 1982), the Boltzmann machine (Hinton and Sejnowski, 1986), and harmony theory (Smolensky, 1986). Boltzmann machines and harmony theory are stochastic models that relax by simulated annealing, in analogy with statistical mechanics. These networks also have learning algorithms.

An observation connectionists are fond of making is that there are no explicit rules in distributed models: all the knowledge is in the connection strengths. Since individual units are not meaningful as symbols (only activity patterns taken as a whole are meaningful), the connections between units cannot be regarded as symbolic-level rules, and the connectionist model's behavior is not rule-based (Smolensky, *in press*; Derthick and Plaut, 1986).

The natural (but incorrect) counter to this argument is that it could be made about any symbol manipulation system if we choose too low a level of description, e.g., describing a digital computer by the behaviors of individual transistors. The flaw in this reasoning is that the computer's circuitry is constrained a priori to implement a logically-designed instruction set. Therefore one can abstract away from the transistor level to an instruction-level of description without loss of information about the computational behavior of the machine. In contrast, distributed connectionist models are not constrained to implement machines with symbolic-level descriptions. They are typically constructed by learning procedures whose only goal is to min-

imize an error measure by modifying connection strengths. Connectionists therefore claim that since the learning procedure is not required (or even trying) to implement a machine that possesses a symbolic-level description, it is unlikely that the networks they construct will have such descriptions. To the extent that these networks exhibit intelligent behavior, their intelligence is at the subsymbolic level, not at the level of formal operations on symbol structures; the latter is at best an approximate description of the computation taking place.

In the remainder of this paper I will argue against the connectionist position, beginning with a reexamination of the symbolic-level theory of thermostats.

6. Beyond Pattern Transformation

Consider a graph of ambient temperature vs. setpoint. We can draw a line with slope 1 that divides the graph into two regions, one labeled "furnace on," the other "furnace off." Given any point specified by T and T_0 , we can predict from the region the point falls in whether the thermostat will turn the furnace on or off. Furthermore, using back propagation we can train a one-unit connectionist network to do this, and it will automatically "generalize" to points not in the training set.

This example demonstrates that the thermostat can be faithfully modeled as a one-neuron linear discriminator rather than as a symbol processing device. More demanding discriminations, involving multiple regions with complex shapes, would require more units and several processing layers, but they are not fundamentally different from this example. Connectionist learning schemes are apparently quite good at learning to make pattern discriminations; they are often better than previously-known pattern recognition techniques. Furthermore, connectionist models can learn to transform patterns rather than merely classify them; under certain conditions a properly-trained network can transform a novel pattern into another novel pattern, once it's learned the "rule" for doing so. But this has little to do with symbol processing.

If we examine the connectionist models that have been held up as evidence against the symbolic paradigm, we see that rather than attacking the symbol manipulation problem head on to demonstrate the illusory nature of symbol processing, they have instead been trivializing complex behaviors to get simple tasks that can be solved by pattern transformation or just pattern recognition. The (rather limited) success of these models merely proves that symbol processing isn't required for such tasks, just as it isn't required to implement a thermostat.

There are several reasons why intelligent behavior should not be dismissed as simply a pattern transformation problem. First, as Fodor and Pylyshyn point out, language and thought have a highly combinatorial, compositional structure. Whether or not such structure is reflected at some hypothetical sub-symbolic level, connectionist models must at least *behave* as if they had such structures inside them. Pattern transformation systems do not meet this criterion

unless they are trained on practically every structure they will ever encounter, as in (Allen, 1987).

Second, the notion of "variables" is essential to intelligent behavior, as it permits the manipulation of structures that were not specified in advance. One example is filling in the participants in a script. If we see John go into a restaurant where Mary works, we know that it is Mary who will bring the menu and John who will read it. This is not conscious problem solving, it's the sort of common sense reasoning that takes place at the intuitive level. The restaurant script includes the variables Customer and Waitress, and our intuitive comprehension of the situation leads us to conclude that John and Mary play those respective roles; this in turn allows John and Mary to be instantiated elsewhere in the script to predict things like who reads the menu. Another example, due to Pinker and Prince, is morphological reduplication, which copies an entire stem as a unit, yielding forms such as "dum-dum" or "boom-boom". The variable, in this case, references the stem to be copied.

Third, as Drew McDermott notes in (McClelland *et al.*, 1986), the ability to contrast one thing with another (as in "She is more outgoing with her friends than with me, her advisor") is an important part of reasoning. McDermott calls this property "thirdness". How can a connectionist model formulate such propositions without doing symbol manipulation?

7. Conclusion

The nature of thought and language would appear sufficient to impose a symbol-processing level on connectionist networks, *provided* we choose a task rich enough to require this level rather than one so simple it can be solved by pattern transformation alone.

On the other hand, if the eliminative hypothesis is correct, connectionists should be able to point to mental phenomena that cannot be explained by symbol processing models, but are explained by connectionist ones. They are nowhere near the point where they can do this convincingly. Connectionist models are too primitive to reproduce even basic symbol processing behavior, and the space of symbolic-level models is huge, making it difficult to prove that no such model could ever account for a particular behavior.

Even if the eliminative hypothesis is disproved, leaving connectionism an implementation technique instead of an alternative paradigm for cognition, the distributed connectionist approach promises to be a source of many valuable insights. For example, Derthick's Micro-KLONE, a connectionist version of KL-ONE (Brachman and Schmolze, 1985), shows how counterfactual reasoning by constructing plausible models can be formulated as a massively parallel constraint satisfaction problem involving thousands of sub-symbolic micro-inferences (Derthick, 1987). This approach to reasoning is a distinct departure from traditional AI methods. The parallel application of knowledge is a serious problem that AI has largely failed to address.

In conclusion, connectionist models are too new for us to determine the validity of the eliminative hypothesis. But if this revolutionary idea is to have any chance of success, connectionists must first construct more complex models that go beyond simple pattern transformation and relaxation. They must at least approximate the powerful linguistic and inferential abilities human beings are known to possess.

Acknowledgements

This work was sponsored by the Office of Naval Research under contract number N00014-86-K-0678. I thank Paul Smolensky, James McClelland, David Rumelhart, and George Lakoff for useful discussions.

- [1] Allen, R. (1987) Natural language and back-propagation: demonstratives, analogies, pronoun reference, and translation. *Proceedings of the IEEE First Annual International Conference on Neural Networks*, San Diego, CA.
- [2] Anderson, J. R. (1983) *The Architecture of Cognition*. Cambridge, MA: Harvard University Press.
- [3] Anderson, J. R. (in press) Methodologies for studying human knowledge. *Behavioral and Brain Sciences*, to appear.
- [4] Ballard, D. H. (1986) Parallel logical inference and energy minimization. Technical report TR-142, Department of Computer Science, University of Rochester, Rochester, NY.
- [5] Brachman, R. J., and Schmolze, J. G. (1985) An overview of the KL-ONE knowledge representation system. *Cognitive Science* 9(1):171-216.
- [6] Cottrell, G. W. (1985) A connectionist approach to word sense disambiguation. Doctoral dissertation available as technical report TR-154, Department of Computer Science, University of Rochester, Rochester, NY.
- [7] Derthick, M. A. (1987) Counterfactual reasoning with direct models. *Proceedings of AAAI-87*, Seattle, WA, 346-351.
- [8] Derthick, M. A., and Plaut, D. C. (1986) Is distributed connectionism compatible with the physical symbol system hypothesis? *Proceedings of the Eighth Annual Conference of the Cognitive Science Society*, Amherst, MA, 639-644.
- [9] Fianty, M. A. (1986) Context-free parsing with connectionist networks. In J. S. Denker (ed.), *Neural Networks for Computing*. AIP Conference Proceedings number 151. New York: American Institute of Physics.
- [10] Fodor, J. A. and Pylyshyn, Z. W. (1987) Connectionism and cognitive architecture: a critical analysis. Manuscript draft, May, 1987.
- [11] Hinton, G. E., and Sejnowski, T. J. (1986) Learning and relearning in Boltzmann machines. In Rumel-

- hart and McClelland (eds.), *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, volume 1. Cambridge, MA: Bradford Books/MIT Press.
- [12] Hopfield, J. J. (1982) Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences USA* 79:2554-2558.
- [13] McClelland, J. L. and Rumelhart, D. E. (1986) *Parallel Distributed Processing: Explorations in the Microstructure of Cognition. Volume 2: Psychological and Biological Models*. Cambridge, MA: Bradford Books/MIT Press.
- [14] McClelland, J. L., Feldman, J., Bower, G., and McDermott, D. (1986) Connectionist models and cognitive science: goals, directions and implications. Report of a workshop on connectionism convened by NSF. Available as a technical report from the Carnegie Mellon University Department of Psychology.
- [15] Newell, A. (1980a) Physical symbol systems. *Cognitive Science* 4:135-183.
- [16] Newell, A. (1980b) Harpy, production systems and human cognition. In R. Cole (ed.), *Perception and Production of Fluent Speech*. Hillsdale NJ: Erlbaum.
- [17] Pinker, S. and Prince, A. (1987) On language and connectionism: analysis of a parallel distributed processing model of language acquisition. Occasional paper #33, MIT Center for Cognitive Science, Cambridge, MA. Revised version to appear in *Cognition*.
- [18] Rumelhart, D. E. and McClelland, J. L. (1986a) *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, volume 1. Cambridge, MA: Bradford Books/MIT Press.
- [19] Rumelhart, D. E., and McClelland, J. L. (1986b) On learning the past tenses of English verbs. In J. L. McClelland and D. E. Rumelhart (eds.) *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, volume 2. Cambridge, MA: Bradford Books/MIT Press.
- [20] Rumelhart, D. E., Hinton, G. E., and Williams, R. J. (1986) Learning internal representations by error propagation. In Rumelhart and McClelland (eds.), *Parallel Distributed Processing: Explorations in the Microstructure of Cognition. Volume 1: Foundations*. Cambridge, MA: Bradford Books/MIT Press.
- [21] Sejnowski, T. J., and Rosenberg, C. R. (1987) Parallel networks that learn to pronounce English text. *Complex Systems* 1(1):145-168.
- [22] Selman, B. (1985) Rule-based processing in a connectionist system for natural language understanding. Technical report CSRI-168, University of Toronto Computer Systems Research Institute. Toronto, Canada.
- [23] Shastri, L. (1987) *Semantic Networks: an Evidential Formalization and Its Connectionist Realization*. Los Altos, CA: Morgan Kaufman.
- [24] Smolensky, P. (1986) Information processing in dynamical systems: foundations of harmony theory. In Rumelhart and McClelland (eds.), *Parallel Distributed Processing: Explorations in the Microstructure of Cognition. Volume 1: Foundations*. Cambridge, MA: Bradford Books/MIT Press.
- [25] Smolensky, P. (in press) On the hypotheses underlying connectionism. *Behavioral and Brain Sciences*, to appear.
- [26] Stabler, E. P., Jr. (1983) How are grammars represented? *Behavioral and Brain Sciences* 6(3):391-420.
- [27] Thompson, H. (1983) Computation misrepresented: the procedural/declarative controversy resumed. *Behavioral and Brain Sciences* 6(3):415.
- [28] Touretzky, D. S., and Hinton, G. E. (1985) Symbols among the neurons: details of a connectionist inference architecture. *Proceedings of IJCAI-85*, Los Angeles, CA, 238-243.
- [29] Touretzky, D. S. (1986) BoltzCONS: reconciling connectionism with the recursive nature of stacks and trees. *Proceedings of the Eighth Annual Conference of the Cognitive Science Society*, Amherst, MA, 522-530.
- [30] Touretzky, D. S., and Geva, S. (1987) A distributed connectionist representation for concept structures. *Proceedings of the Ninth Annual Conference of the Cognitive Science Society*, Seattle, WA, 155-164.
- [31] Waltz, D., and Pollack, J. (1985) Massively parallel parsing: a strongly interactive model of natural language interpretation. *Cognitive Science* 9(1):51-74.