Connectionist Models of Rule-Based Reasoning

Ron Sun
Brandeis University
Computer Science Department
Waltham, MA 02254
rsun@cs.brandeis.edu

Abstract

We investigate connectionist models of rule-based reasoning, and show that while such models usually carry out reasoning in exactly the same way as symbolic systems, they have more to offer in terms of commonsense reasoning. A connectionist architecture for commonsense reasoning, CONSYDERR, is proposed to account for commonsense reasoning patterns and to remedy the brittleness problem in traditional rule-based systems. A dual representational scheme is devised, utilizing both localist and distributed representations and exploring the synergy resulting from the interaction between the two. CONSYDERR is therefore capable of accounting for many difficult patterns in commonsense reasoning. This work shows that connectionist models of reasoning are not just "implementations" of their symbolic counterparts, but better computational models of commonsense reasoning.

Introduction

Rule-based reasoning is the most prominent paradigm of symbolic AI. Whether connectionist models can be a viable alternative to symbolic AI depends, to some extent, on their ability to account for rule-following behaviors and rule-based reasoning.

To account for rules in connectionist models, various approaches have been explored, and several different systems have been implemented that can carry out rule-based reasoning rather completely. However, most of them are straight "implementations", without any fundamental difference from symbolic systems in terms of logical capabilities or reasoning capacities.

Can connectionist models do more and better in terms of accounting for robust, flexible, and multifaceted human commonsense reasoning, especially in the types of reasoning that are most often dealt with by rule-based paradigms? Can those data (such as in Collins & Michalski 1990) that are difficult to account for by symbolic rule-based systems be explained by connectionist rule-based reasoning models? We try to answer these questions by building a connectionist reasoning system to account for common patterns in commonsense reasoning. This work shows that connectionist models of reasoning are not just an "implementation" of their symbolic counterparts; rather, they are better computational models of commonsense reasoning, taking into consideration the approximate, evidential and flexible nature of rule-based reasoning, and similarity-based reasoning (or a limited form of analogy), and also accounting for the spontaneity and parallelism in reasoning processes.

Below we will first look into various existing connectionist models of rule-based reasoning. Then we will identify some problems that are difficult for them to solve. We will move on to develop a new connectionist architecture that can better deal with these problems. In this paper, we will concentrate more on motivational issues than on technical details.

Background Reviews

Let us look into some previous work in rule-based reasoning, especially connectionist ones. Rule-based systems have a long history in AI and cognitive science. The early successful work such as those described in Buchanan & Shortliffe (1984) demonstrated the promise of this overall approach, which adopts a simple representation with modular units called rules composed of a small set of conditions and conclusions. Many elaborate cognitive theories of learning, problem solving and memory, etc. were built based on this paradigm (cf. Klahr et al 1987) [1].

Because of this success, when connectionism came along, one of the main challenges for connectionism was how to implement rule-based reasoning in a network fashion. Touretzky & Hinton (1985) is the first work that tackles this problem. They basically emulate the structure of a symbolic rule-based (production) system, with separate modules for working memory, rules, and facts. An elaborate pull-out network is designed to match working memory data against rules

[1] Nevertheless the paradigm has long been plagued by the *brittleness* problem for large scale systems, as will be discussed.

and to decide which matching rule is to fire. Competition and winner-take-all are used for the matching purpose. The result is the equivalent of a simple sequential symbolic rule-based system.

Barnden (1988) represents another early attack on this problem. In his system, data reside in grid-like networks (called Configuration Matrices), coded with the help of adjacency relations and highlighting techniques. Hardwired rules are used to detect the presence of data that match particular rules, and an "Action Part" module can be used to add a new data structure representing the conclusion from the matched rule. Although there is some parallelism, it is mostly a sequential rule-based system, carrying out symbolic processing.

From these examples, it is quite clear now that connectionist models are capable of implementing rule-based reasoning in a variety of ways [2]. Now the question is: Can connectionist models account better for *commonsense* reasoning? The evidential, robust, flexible, and multifaceted nature of commonsense reasoning is evident from various studies (such as in Collins & Michalski 1990) yet they are all absent from these above models. What is really needed for a connectionist model of rule-based reasoning to be able to model commonsense reasoning adequately? We confront the above questions, by analyzing real protocol data and then actually building a new kind of connectionist models as a computational mechanism for commonsense reasoning.

Common Reasoning Patterns

Allan Collins collected a number of protocols of commonsense reasoning. He indicated the inadequacy of traditional logic in explaining those reasoning patterns, and argued for the use of different formalisms or frameworks in the study of *common reasoning patterns* found in various commonsense reasoning tasks. Collins & Michalski (1990) have done an impressive job in terms of analyzing the data and establishing a unifying framework for explaining them. What is needed is a computational mechanism from which various inference patterns contained in the data can emerge into existence. We believe that the mechanism ought to be analytically simple, structurally unified, and mechanistically sound. Connectionist models in general fit these above descriptions very well, so they might provide such a mechanism.

Let's look at some examples from Collins & Michalski (1990). One protocol is as follows:

Q: Do you think they might grow rice in Florida?

R: Yeah. I guess they could, if there were an adequate fresh water supply, certainly a nice, big, warm, flat area.

[2] Also, Dolan & Smolensky (1988), Sun (1989) (see also Sun & Waltz 1991), Ajjanagadde & Shastri (1989), and Lange & Dyer (1989) are works basically belonging to this same category.

In this example, the person answering the question deduced an uncertain conclusion based on partial information, with a piece of crucial information (fresh water) missing. This example also indicates the need for an additive procedure for accumulating evidence.

Another case is as follows:

Q: Is the Chaco the cattle country?

R: It is like western Texas, so in some sense I guess it's cattle country.

Here because there is no known knowledge (or no applicable rules), an uncertain conclusion is drawn based on similarity with known knowledge (rules).

Yet another case is

Q: Are there roses in England?

R: There are a lot of flowers in England. So I guess there are roses.

Here the deduction is based on property inheritance (flower LOCATION England; rose IS-A flower; so rose LOCATION England), and the conclusion is partially certain and can be drawn only when there is no information to the contrary (i.e. no cancellation).

Existing connectionist models, or any computational models for that matter, so far cannot deal very well with these above patterns in a single unified model.

The Brittleness Problem

Those above examples are actually instances of a general problem that has been plaguing symbolic AI for long, namely the brittleness problem.

The brittleness problem can be defined, for the purpose of this research, as the inability of a system to deal with the following aspects of reasoning in a systematic way within a unified framework: (1) partial information (for example, the first protocol), (2) uncertain or fuzzy information, (3) no matching rules (for example, the second protocol), (4) rule interactions, i.e. lack of consistency and completeness in a fragmented rule base, (5) generalization, (6) bottom-up inheritance, (7) top-down inheritance (for example, the third protocol), and (8) learning new rules.

Detailed analysis of these aspects (with lots of examples) shows that, while they look like a disparate set of problems, they can all be characterized as reasoning with rules supplemented with (feature) similarity-related inferences. We define a measure of **conceptual similarity** (cf. Tversky 1977) as

$$(A \sim B) = \frac{|F_A \cap F_B|}{|F_A|} \in [-1,1]$$

such that [3]

$$if\ ACT_A = a,\ then\ ACT_B = a * (A \sim B)$$

where F_i is the feature representation of node "i", and ACT_i is the activation of node "i". We define a measure of **knowledge links** (i.e. rules) as

$$(A \longrightarrow B) = r \in [-1,1]$$

[3] We assume there is nothing else affecting ACT_B

such that

$$if\ ACT_A = a,\ then\ ACT_B = a * (A \longrightarrow B)$$

where ACT_i is the activation of node "i", and r is the knowledge link (rule) strength [4] between A and B. Each of these above cases can be analyzed and dealt with utilizing these two concepts, for example, the no-matching-rule situation can be described as:

$$A \sim B$$

$$B \longrightarrow C$$

and A is activated ($ACT_A \neq 0$). So we have

$$ACT_C = (B \longrightarrow C) * ACT_B$$

$$= (B \longrightarrow C) * ACT_A * (A \sim B)$$

Other cases can be described similarly, except learning new rules, which is a separate issue (see Sun & Waltz 1991).

These two mechanisms are embedded in our new architecture: CONSYDERR [5], so each of these aspects of brittleness can be handled by our system.

A Sketch of the Model

The CONSYDERR architecture consists of two levels: CL and CD. CL is a connectionist network with localist representation, or roughly reasoning at the conceptual level (cf. Smolensky 1988). Rules are represented in CL as links between two nodes representing the condition and the conclusion respectively. The scheme proposed, FEL or Fuzzy Evidential Logic, can handle a superset of Horn clause logic and Shoham's modal logic (or Causal Theories, cf. Shoham 1990), so that it can fully accommodate traditional rule-based reasoning and capture commonsense causal knowledge. Moreover, it is capable of approximate and cumulative evidential reasoning and works with partial and uncertain information. Unlike Horn clause logic, it can deal with negative as well as positive evidence. It can handle variable bindings by utilizing the DN/PDN formalism as in Sun (1989) and Sun (1990). The basic operation of this scheme is simply weighted-sum computations, therefore this scheme can be implemented, with ease, in a connectionist network with weighted-sum node activation functions. Because of the limited space and the need to emphasize the main points in this short presentation, we will not discuss in detail the above points regarding rule representations (see Sun 1991 for details).

CD is a connectionist network with distributed representation, roughly corresponding to reasoning at the subconceptual level. Concepts and rules are diffusely

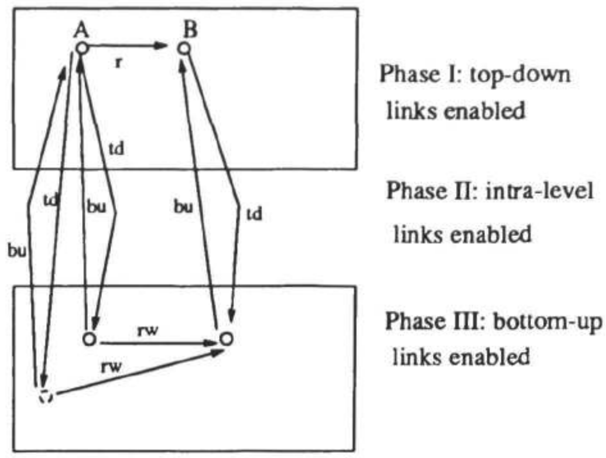

Figure 1: A two-level architecture

represented by sets of units overlapping each other. The amount of overlapping of two sets of units representing two different concepts is proportional to the degree of similarity between these two concepts. I call this a *similarity-based representation*, in which units can be features, perceptual primitives, internal goals or affect states. Concepts are "defined" in terms of their similarity to other concepts in these primitive representations. I will utilize these primitives only as a substratum for similarity-based representation of higher level concepts.

Now we can link the localist network (CL) with this distributed network (CD), by linking each node in CL representing one concept to all the nodes in CD representing the same concept, and assign them appropriate weights (see Fig.1). Those cross-component links are moderated by a latch mechanism. The rule links in CL are duplicated (diffusely) in CD. The interactions of the two components are in fixed cycles: first the latch opens to allow the activation of CL nodes to flow into corresponding CD nodes, and then the two parts start settling down on their own simultaneously, and finally the latch opens to allow the activation of nodes in CD to flow back into CL to be combined with the activation of corresponding CL nodes.

From the above description it is clear that the system is a combination of rule-based and similarity-based components, interwoven together. It implements naturally the functions defined above for knowledge links and conceptual similarity. The synergy of the two types of representation and reasoning helps to deal with the brittleness problem listed above and, therefore, to account for the aforementioned common reasoning patterns.

For example, in order to solve the no-matching-rule situation, we can explore the similarity between the current situation and the rule conditions as represented in the CD part of the system. Consider the following case:

Cars are for traveling on ground. Airplanes are for

[4] When there are multiple conditions in a rule, this measure becomes a vector, and the multiplication used here is generalized to inner-products.

[5] It stands for a CONnectionist SYstem with Dual representation for Evidential Robust Reasoning

traveling in air. Are buses for traveling on ground or in the air?

In the context of deciding modes of transportation, based on the similarity of relevant features [6] as represented with the amount of overlapping in corresponding sets of units in CD, buses are closer to cars than to airplanes. So "traveling on ground" is activated more strongly. One concludes that buses are for traveling on ground.

Another example concerns the problem of inconsistent/incoherent rule bases (rule interaction), we can utilize the rule interactions in CD:

If carrying cargo, buy utility vehicles. If carrying passengers, buy passenger vehicles. If carrying both cargo and passengers, what shall one buy?

Different types of vehicles are represented as features in CD. When the above two rules are both activated (in response to the question), all features corresponding to both utility and passenger vehicles will be activated in CD. All this information will go up to CL, and the things corresponding to the intersection of utility and passenger vehicles will be activated strongly (because they have all the features). So something like "van" will win.

Other aspects of the brittleness problem can be solved in a similar fashion, including the common patterns identified by Collins & Michalski (1990) [7]. This solution is quite different from Collins & Michalski. My contention is that this model is conceptually simpler and computationally more efficient (by combining and eliminating many parameters).

Formal mathematical analyses were performed for each of these aspects, and we came up with a set of requirements and constraints for each aspect regarding the parameters of a system that can deal with that particular aspect. After analyzing how these requirements and constraints imposed by each of these aspects interact with one another, a synthesis is achieved, so that a unified system is formed with a unique set of parameter settings satisfying all requirements. Based on that, a large-scale system consisting of about two hundred nodes was built to test, in a realistic setting, how these fragments combine. The system utilizes geographical knowledge extracted from encyclopedias and performs commonsense reasoning based on that knowledge (see Sun 1991).

A Detailed Example

The Problem

Look at the "Chaco" example (Collins & Michalski 1990).

Q: Is the Chaco the cattle country?

R: It is like Western Texas, so in some sense I guess it's cattle country.

We can put it another way to straighten out the reasoning:

Western Texas is cattle country.
Chaco is similar to western Texas (in some relevant aspects).
So Chaco is cattle country.

An analysis

In this example, because there is no known knowledge (or no applicable rules), an uncertain conclusion is drawn based on similarity with known knowledge (rules). Using the formalism we developed, it can be described as:

$$Chaco \sim WesternTexas$$

$$WesternTexas \longrightarrow cattlecountry$$

Given "Chaco" with $ACT_{Chaco} = 1$, "cattlecountry" is concluded with $ACT_{cattlecountry}$ calculated as follows:

$$ACT_{cattlecountry} = (WesternTexas \longrightarrow cattlecountry)$$
$$*ACT_{Chaco} * (Chaco \sim WesternTexas)$$

where the similarity measure is chosen to facilitate later implementations in the CONSYDERR architecture:

$$Chaco \sim WesternTexas = p * \frac{|F_{Chaco} \cap F_{WesternTexas}|}{|F_{Chaco}|}$$

$p \in [0, 1]$ is a parameter used for adjusting the system's behavior, from absolute rigidity to free-floating thinking (see Sun 1991).

These equations can be readily translated into the CONSYDERR architecture: Links between nodes in both CL and CD represent rule strength measures (the link weights are defined to be the corresponding rule strengths), and similarity measures are implemented with CD representations (we use a set of nodes to represent all features in CD and the amount of overlapping between representations of two concepts expresses the conceptual similarity of these two concepts). See Fig.2.

The Working of the System

After starting to receive input data, the CONSYDERR system operates in fixed cycles: (1) Top-down phase, (2) Settling phase, and (3) Bottom-up phase. This cycle can be repeated to continuously track inputs.

In Top-down phase, the computation is as follows:

$$x_i(t+1) = \max_a ACT_a(t)$$

where a is any node in CL that has $x_i \in CD_a$.

In Settling phase, the computation is as follows:

$$\Delta ACT_a = \alpha \sum W_i I_i(t) - \beta ACT_a(t)$$

[6] Such as Having-wings, Having-tails, Wheels-on-both-sides, Aerodynamic-shapes, Landing-gears, etc.

[7] We are certainly not implying that we solved the brittleness problem completely. Rather, we are aiming for a simple and elegant model that can deal with some important and predominant aspects of the problem very effectively and efficiently.

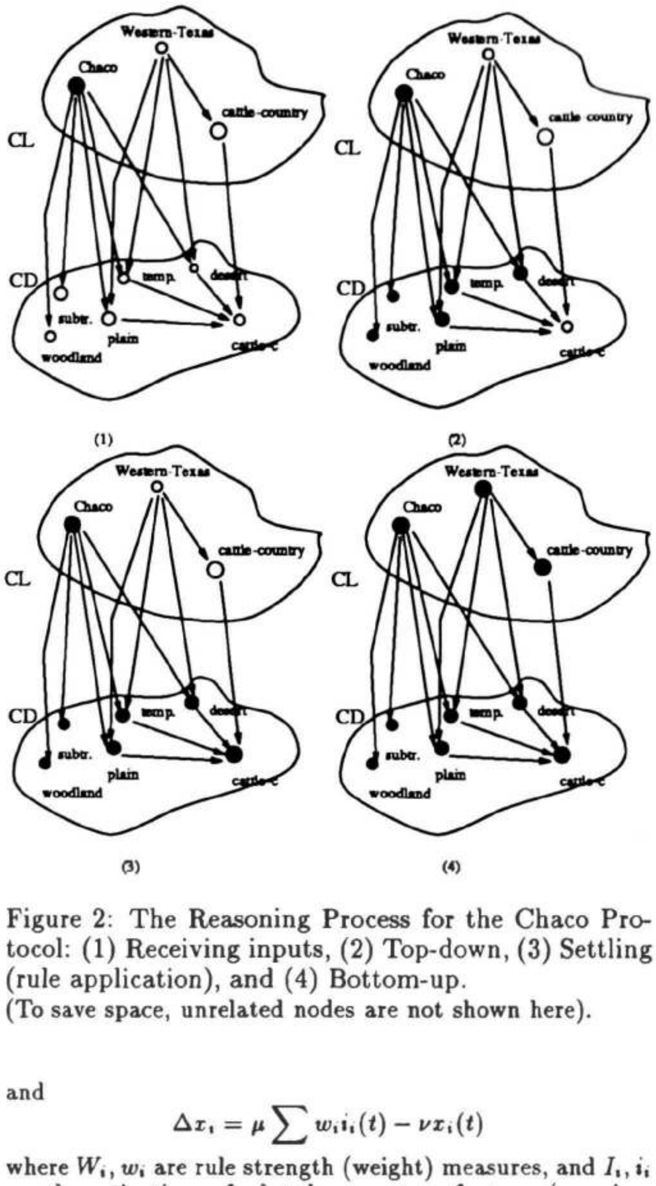

Figure 2: The Reasoning Process for the Chaco Protocol: (1) Receiving inputs, (2) Top-down, (3) Settling (rule application), and (4) Bottom-up.
(To save space, unrelated nodes are not shown here).

and

$$\Delta x_i = \mu \sum w_i i_i(t) - \nu x_i(t)$$

where W_i, w_i are rule strength (weight) measures, and I_i, i_i are the activations of related concepts or features (premises or logical predecessors).

In Bottom-up phase, the computation is as follows:

$$ACT_b(t+1) = \max(ACT_b(t), \sum_{x_i \in CD_b} \frac{x_i(t)}{|CD_b|})$$

where b is any node in CL.

Applying this cycle to the example: Top-down phase will activate the CD representation of "Chaco" and activate partially the CD representation of "WesternTexas" based on their similarity; then in Settling phase, rules (links) take effect and this amounts to applying in CD the rule: *WesternTexas is cattle-country*, so the CD representation of "cattle country" is partially activated; finally in Bottom-up phase, the partially activated CD representation of "cattle country" will percolate up to activate the "cattle country" node in CL. The result can be read off from CL. See Figure 2.

Comparing it with other systems: CONSYDERR utilizes parallelism inherent in the data to the maximum extent, especially when compared with Touretzky & Hinton (1985) or Dolan & Smolensky (1989). While most other connectionist rule-based systems (Lange & Dyer 1989, Ajjanagadde & Shastri 1989, etc.) are functionally comparable to the CL part of CONSYDERR, the CD part is unique in that it provides an efficient way for similarity matching to supplement rule-based reasoning; the CL/CD dual representation scheme constitutes a principled way of accounting for the dichotomy of conceptual level and subconceptual (intuitive) level reasoning (Smolensky 1988, Sun 1991). More recently Barnden & Srinivas (1990) utilize connectionist rule-based systems to explore similarity in reasoning (i.e. connectionist case-based reasoning); while the idea is very similar to ours (see Sun 1991 for details), their system requires a complex retrieval/matching process.

The Initial Setup

The question of how we can gather data and set up a large system can be divided into two questions: how do we obtain rule weights and how do we obtain similarity measures?

In our large-scale system (Sun 1991), rules are obtained by going through geography sourcebooks, picking out the relevant information and integrating it into the network with the CFRDN procedure (see Sun 1991). The rules being put into the system include *WesternTexas is cattlecountry*, etc. [8]

Similarity measures are obtained by an indirect means: we first obtain all the relevant features needed for representing the concepts involved, and then naturally the amount of feature overlap is the similarity between concepts involved. In order to come up with detailed feature representations for concepts, we preestablish a set of feature nodes, and we then go through sourcebooks, establishing links (cross-level links) between a concept in CL and its features in CD, based on what we read in the sourcebooks. The features include: altitude, rainfall, vegetation, population, temperature, terrain, etc. with various ranges. [9]

One important issue is how the system focuses on relevant features and ignore or discount somehow the irrelevant ones, given the context (or the query, in the above-mentioned cases). This is done by the attention focusing module external to the system, in which a set of "context rules" are used to pick out all relevant features and suppress others when a certain context

[8] In general, weights representing rules can be obtained by reading textbooks, instructions, or by using learning algorithms through interactions with the environment. There is no universally applicable way to do this, or in other words, it is domain-specific.

[9] Another possible way of obtaining similarity measures is to conduct a test, asking a group of subjects to rate the similarity of concepts concerned and then construct CD representations based on the collected test scores with the STSIS procedure (see Sun 1991).

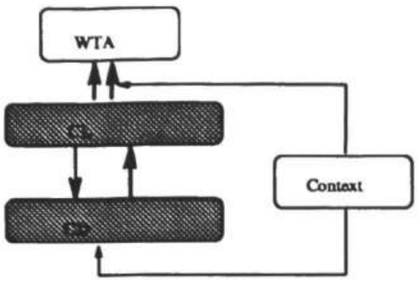

Figure 3: The Overall Architecture with Feature and Result Selections

is given (by activating the node representing the particular context, in ways described by Sun 1989). See Figure 3 for a sketch of these mechanisms. Because of the fact that these mechanisms are domain specific, they are not part of the CONSYDERR architecture, but add-on mechanisms (cf. Sun 1991).

Summary

We analyzed connectionist models for rule-based commonsense reasoning. A connectionist architecture is proposed to account for some common patterns found in commonsense reasoning and to remedy to a certain extent the brittleness problem found in typical symbolic systems. Different from other existing connectionist systems, a dual representational scheme is devised, which has extensional objects (localist representation) as well as intensional objects (distributed representation with features). By using feature-based distributed representation in addition to the localist representation, we are able to explore the synergy resulting from the interaction between these two types of representations and between rule-based reasoning and similarity-based reasoning. This synergy helps to deal with problems such as partial information, no exact matching, property inheritance, rule interaction. and therefore the CONSYDERR system is capable of accounting for many difficult reasoning patterns in one unified system. This architecture also demonstrates that connectionist models equipped with symbolic capabilities are powerful tools for modeling reasoning capacities as well as for constructing efficient practical systems (by utilizing massive parallelism), and they are not mere implementations of their symbolic counterparts.

Acknowledgments I wish to thank Dave Waltz, James Pustejovsky, and Tim Hickey for many helpful discussions, and Don Smith for commenting on the paper.

References

V. Ajjanagadde and L. Shastri, 1989. Efficient Inference with Multi-place Predicates and Variables in a Connectionist System. In Proc.11th Cognitive Science Society Conference, 396-403. Hillsdale, NJ: Erlbaum.

J. Barnden, 1988. The right of free association: relative-position encoding for connectionist data structures. In Proc.10th Conference of Cognitive Science Society, 503-509. Hillsdale, NJ: Erlbaum.

J. Barnden and K. Srinivas, 1990. Overcoming Rule-Based Rigidity and Connectionist Limitations Through Massively Parallel Case-based Reasoning, Technical Report MCCS-90-187, Computing Research Laboratory, New Mexico State University, Las Cruces, NM.

B. Buchanan and E. Shortliffe, 1984. Rule-Based Expert Systems. Reading, MA: Addison-Wesley.

A. Collins and R. Michalski, 1990. The logic of plausible reasoning, Cognitive Science, 13(1):1-49.

C. Dolan and P. Smolensky, 1988. Implementing a connectionist production system using tensor products. In D. Touretzky et al. eds. Proc.1988 Connectionist Summer School, 265-272. San Mateo, CA: Morgan Kaufman, 1989

J. Klahr et al, eds. 1989. Production System Models of Learning and Development. Cambridge, MA: MIT Press.

T. Lange and M. Dyer, 1989. Frame selection in a connectionist model. In Proc.11th Cognitive Science Conference, 706-713. Hillsdale, NJ: Erlbaum.

L. Shastri, 1988. A Connectionist Approach to Knowledge Representation and Limited Inference, Cognitive Science, 12:331-392.

Y. Shoham, 1990. Non-monotonic Reasoning and Causation, Cognitive Science, 14:213-252.

P. Smolensky, 1988. On the proper treatment of connectionism, Behavioral and Brain Sciences, 11:1-43.

R. Sun, 1989. A Discrete Neural Network Model for Conceptual Representation and Reasoning. In Proc.11th Cognitive Science Society Conference, 916-923. Hillsdale, NJ: Erlbaum.

R. Sun, 1990a. Rules and Connectionism. In Proc.INNC-Paris, p.545. Netherlands: Kluwer.

R. Sun, 1990. The Discrete Neuronal Models. In Proc.INNC-Paris, 902-907. Netherlands: Kluwer.

R. Sun and D. Waltz, 1991. A Neurally Inspired Massively Parallel Model of Rule-Based Reasoning. In B. Soucek ed. Neural and Intelligent System Integration. New York, NY: John Wiley & Sons.

R. Sun, 1991. Integrating Rules and Connectionism for Robust Reasoning: An Architecture with Dual Representation. Ph.D Thesis, Computer Science Department, Brandeis University, Waltham, MA.

D. Touretzky and G. Hinton, 1985. Symbol Among Neurons. In Proc.9th IJCAI, 238-243. San Mateo, CA: Morgan Kaufman.

A. Tversky, 1977. Features of Similarity, Psychological Review, 84(4):327-352.

D. Waltz and J. Pollack, 1985. Massive parallel parsing, Cognitive Science, 9:51-74.