Integrating Rules and Connectionism for Robust Reasoning

Ron Sun

Brandeis University
Computer Science Department
Waltham, MA 02254

(617) 736-2712
rsun@cs.brandeis.edu

Technical Report TR-CS-90-154

January 10, 1991

Abstract

A connectionist model for robust reasoning, CONSYDERR, is proposed to account for some common reasoning patterns found in commonsense reasoning and to remedy the brittleness problem. A dual representational scheme is devised, which utilizes both localist representation and distributed representation with features. We explore the synergy resulted from the interaction between these two types of representations, which helps to deal with problems such as partial information, no exact match, property inheritance, rule interaction, etc. Because of this, the CONSYDERR system is capable of accounting for many difficult patterns in commonsense reasoning. This work also shows that connectionist models of reasoning are not just an "implementation" of their symbolic counterparts, but better computational models of common sense reasoning, taking into consideration of the approximate, evidential and adaptive nature of reasoning, and accounting for the spontaneity and parallelism in reasoning processes.

It is quite clear now that connectionist models are capable of implementing rule based reasoning in a variety of ways (e.g. Hinton & Touretzky 1986, Barnden 1989, Sun 1989, Ajjanagada & Shastri 1989, Lange & Dyer 1989, etc), as well as carrying out other types of reasoning (such as similarity based reasoning, categorization, or domain-specific reasoning tasks). The questions are: Can connectionist models do more and better in terms of accounting for robust, flexible, and multi-facet human reasoning, especially in the types of reasoning that are most often dealt with by rule based paradigms? How can those data that are difficult to be accounted for by symbolic rule-based system paradigms (cf Klahr et al 1987 and Buchanan & Shortliffe 1984) be explained by connectionist reasoning models? The questions can also be asked the other way around: can we develop some computational mechanisms that can *better* explain existing data regarding some common reasoning patterns (such as in Collins & Michalski 1989) and deal with some nasty brittleness problems that have been plaguing symbolic reasoning systems for long, for the purpose of enhancing our understanding of reasoning in general and connectionist reasoning in particular?

This work tackles the above questions, by actually building and analyzing a connectionist model as a computational mechanism for accounting for the aforementioned data.

This work shows that connectionist models of reasoning are not just an "implementation" of their symbolic counterparts; Rather, they are better computational models of common sense reasoning, taking into consideration of the approximate, evidential and adaptive 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.

1 Common Reasoning Patterns

Allan Collins collected a number of protocols of commonsense reasoning, in the area of elementary geography and the like. 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 common sense reasoning tasks. He and his colleagues have done an impressive job in terms of analyzing the data and establishing a unifying framework for explaining them (for example, Collins 1975, Collins & Michalski 1989). What is needed is a computational mechanism from which various inference patterns contained in the data can emerge into existence. That 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 (1989). 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.

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 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 concellation).

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.

2 The Brittleness Problems

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 in dealing with the following aspects of reasoning in a systematic way within a unified framework:

- partial information (for example, the first protocol),
- uncertain or fuzzy information,
- no matching rules (for example, the second protocol),
- combinational rule interactions, i.e. lack of consistency and completeness in a fragmented rule base,
- generalization,
- bottom-up inheritance,
- top-down inheritance (for example, the third protocol),
- 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 inferencing. We define **conceptual similarity** as [1]

$$A\ sim\ B = \frac{|F_A \cap F_B|}{|F_A|}$$

such that

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

where F_i represents the feature representation of "i", and ACT_i is the activation of "i", and f is a multiplicative, monotone increasing function; and we define **knowledge links** (i.e. rules) as

$$A\ kl\ B = r \in [-1, 1]$$

such that

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

where ACT_i is the activation of "i", and r is the knowledge link (rule) strength 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:

$$A\ sim\ B$$

$$B\ kl\ C$$

[1] assume there is nothing else affecting ACT_B

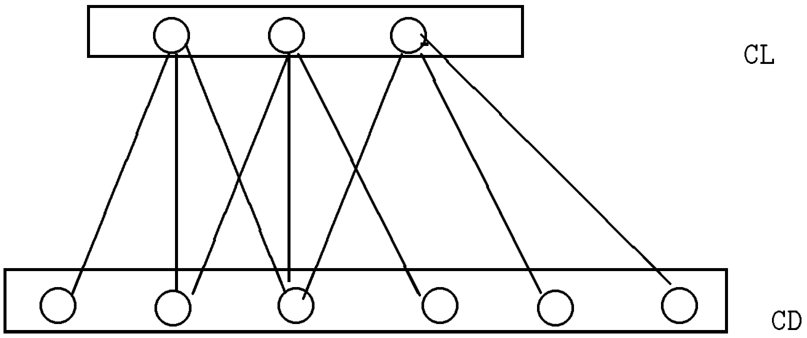

Figure 1: a two-layer architecture

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

$$ACT_C = (B\ kl\ C) * ACT_B = (B\ kl\ C) * f(ACT_A, A\ sim\ B)$$

Other cases can be described similarly, except learning new rules, which is a separate issue (see Sun 1990 for details).

These two mechanisms are embedded in the CONSYDERR [2] architecture, so each of these aspects can be handled by the CONSYDERR system.

3 A Sketch of the Model

The CONSYDERR system consists of two levels: CL and CD. CL is a connectionist network with localist representation, or roughly reasoning at conceptual level (cf. Smolensky 1989). 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. Moreover, it is capable of approximate and cumulative 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 computation, therefore this scheme can be implemented, with ease, in a connectionist network with weighted-sum node activation functions. Because of the page limit and the need to emphasize the main points in this short presentation, we will not discuss in detail the above points regarding rule representations.

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

[2] it stands for a COnnectionist SYstem with Dual representation for Evidential Robust Reasoning

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 patterns in reasoning.

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 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 [3] 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. We conclude that buses are for traveling on ground.

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

> If carrying cargo, buy utility vehicles. If carrying passenger, buy passenger vehicles. If carrying both cargo and passenger, 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 to a certain extent in a similar fashion, including the common patterns identified by Collins & Michalski (1989) [4]. The solution is quite different from Collins & Michalsky. My contention is that this model is conceptually simpler and computationally more efficient (by combining and eliminating many parameters).

The learning of new rules will be discussed in detail elsewhere.

Formal analyses are performed for each of these aspects, and we come up with a set of requirements and constraints for each aspect regarding the range of paramenter values 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 paramenter settings satisfying all requirements. Based on that, a system consisting of about two hundred nodes is 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.

4 A Detailed Example

4.1 The Problem

Look at the "Chaco" example (Collins & Michalski 1989).

> Q: Is the Chaco the cattle country?
>
> R: It is like western Texas, so in some sense I guess it's cattle country.

[3] such as Having-wings, Having-tails, Wheels-on-both-sides, Aerodynamic-shapes, landing-gears, etc.

[4] we are certainly not implying that we solved the brittlness problem. 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.

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.

4.2 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:

$$C \; sim \; WT$$

$$WT \; kl \; CC$$

Given C with $ACT_C = 1$, CC is concluded with ACT_{CC} calculated as follows:

$$ACT_{CC} = (WT \; kl \; CC) * f(ACT_C, \; C \; sim \; WT)$$

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

$$f(ACT_C, \; C \; sim \; WT) = p * ACT_C * \frac{|F_C \cap F_{WT}|}{|F_C|}$$

or equivalently,

$$f(ACT_C, \; C \; sim \; WT) = p * ACT_C * \frac{|CD_C \cap CD_{WT}|}{|CD_C|}$$

and $p \in [0, 1]$ are parameters used for adjusting the system's behavior, from absolute rigidity to free-floating thinking (see Sun 1989).

These equations can be readily translated into the CONSYDERR architecture: Links between nodes in both CL and CD represent *kl* measures (the link strength is defined to be the corresponding knowledge link measure), and *sim* 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 represents the conceptual similarity of these two concepts). See Fig.2.

4.3 The Working of the System

The CONSYDERR system operates in fixed cycles, after starting to receive input data:

1. Top-down phase
2. Activation and Settling phase
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 Activation and Settling phase, the computation is as follows:

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

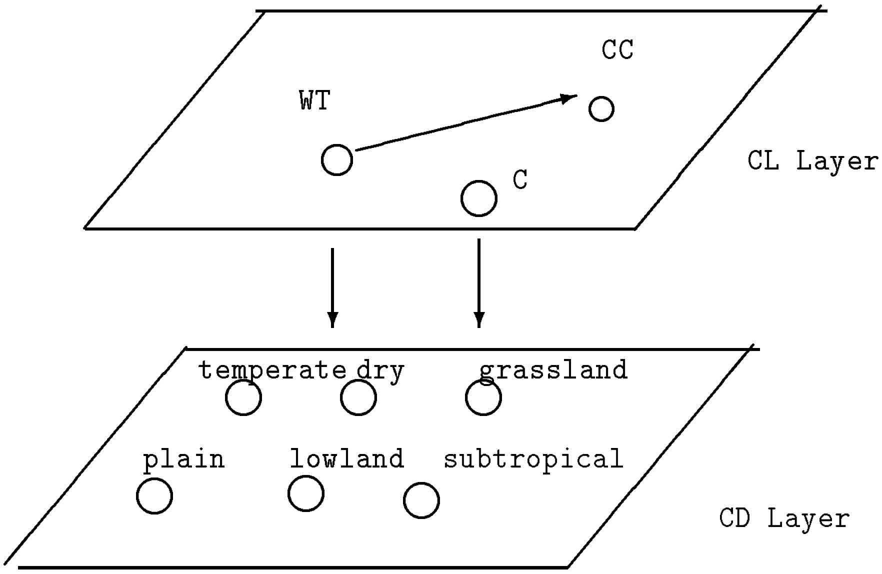

Figure 2: The Chaco example: C=Chaco, CC=cattle country, and WT=Western Texas. (There are actually a lot more features in CD than shown here).

and

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

where W_i, w_i are *kl* measures, and I_i, i_i are the firings of related concepts (premises or logical predecesors).

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: the top-down phase will activate the CD representation of C and activate partially the CD representation of WT based on the *sim* measure; then in the activation and settling phase, *kl* links take effect and this amounts to applying the rule: *WT is CC* in CD, so the CD representation of CC is partially activated; finally in the bottom-up phase, the partially activated CD representation of CC will percolate up to activate the CC node in CL. The result can be read off from CL.

4.4 The Initial Setup

Now the question is how we can come up with values of these parameters involved in representing information in the system in the first place, so that the whole system will work. There are basically two issues: how do we obtain *kl* measures and how do we obtain *sim* measures?

In this particular example, *kl* measures are obtained by going through geography sourcebooks, picking out the relevant information and integrating it into the network with the CFRDN procedure (see Sun 1990). The rules being put into the system include *WT is CC* [5].

sim measures are obtained by an indirect means. We first obtain all the relevant features for the concepts involved. Then the amount of feature overlap is the similarity between concepts involved. In order to

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

come up with detailed feature representations, we pre-establish a set of feature nodes [6]. And then we go through sourcebooks, and establish links (cross-layer 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. [7]

One important aspect 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 is given (by activating the node representing the particular context, in ways described by Sun 1989).

Everything here is done with a rather mechanical process. And the architecture is modular: new nodes can be added and the existing ones can be deleted at either layers and the effect is localized, without affecting the global structure.

5 Summary

In summary, a connectionist model for robust reasoning, CONSYDERR, is proposed to account for some common reasoning patterns found in commonsense reasoning and to remedy to a certain extent the brittleness problem found in typical symbolic systems (although, of course, it can not solve all of the brittleness problem). 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 resulted 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 match, property inheritance, rule interaction, etc. and therefore the CONSYDERR system is capable of accounting for many difficult patterns in reasoning. This architecture also demonstrates that connectionist models equiped with symbolic capabilities are powerful tools for modeling reasoning capacities and constructing efficient practical systems (by utilizing massive parallelism), and they are not mere implementation of symbolic counterparts.

ACKNOWLEDGEMENTS I wish to thank Dave Waltz, James Pustejovsky, and Tim Hickey for lots of discussions and comments on the manuscripts.

REFERENCES

V. Ajjanagadde and L. Shastri, Efficient Inference with Multi-place Predicates and Variables in a Connectionist System, 11th Cognitive Science Society Conference, 1989

J. Barnden, The right of free association: relative-position encoding for connectionist data structures, 10th conf. of Cognitive Science Society, 1988

L. Bookman, A Connectionist Scheme For Modeling Context, Proc. 1988 connectionist Summer School, Morgan Kaufman, 1988

B. Buchanan and E. Shortliffe, Rule-Based Expert System, addison-weslley, 1984

C. Dolan and P. Smolensky, Implementing a connectionist production system using tensor products, Proc. 1988 Connectionist Model Summer School, 1988

[6] for example, rainfall heavy (qualitative features), or rainful 3.5-4.5 inches (quantitative features), etc. The range each of the feature nodes covers overlaps each other to account for similarity. The granularity of the ranges varies.

[7] Another possible way of obtaining *sim* measures is to conduct a test, asking a group of subjects to rate the similarity of concepts concerned and then calculating the *sim* measures based on the collected test scores with the STSIS procedure (see Sun 1990).

J. Feldman and D. Ballard, Connectionist models and their properties, Cognitive Science, July, 1982

J. Hendler, Marker Passing and microfeature, 10th IJCAI, 1987

J. Hinton and D. Touretzky, Symbol Among Neurons, IJCAI, 1985

J. Holland, Escaping brittleness, Machine Learning, Vol.2, 1986

J. Klahr et al, Production System Model of Learning and Development, 1989

T. Lange and M. Dyer, Frame selection in a connectionist model, Proc. Cognitive Science Conference, 1989

J. Pearl, Probabilistic Reasoning in Intelligent Systems, Morgan Kaufman, 1988

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

T. Shultz et al, Managing Uncertainty in Rule-Based Reasoning, 11th Cognitive Science Conference, 1989

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

R. Sun, The logic for approximate reasoning combining probability and fuzziness, 1st Congress of International Fuzzy System Association, 1985

R. Sun, A discrete neural network model for conceptual representation and reasoning, 11th Cognitive Science Society Conference, 1989

R. Sun, Designing inference engines based on a discrete neural network model, IEA/AIE 1989

R. Sun and E. Schalit, Is the tag necessary? Brain and Behavior Sciences, 1989

R. Sun, Rules and Connectionism, Proc.INNC-Paris, 1990

R. Sun and D. Waltz, Neural networks and human intelligence, Journal of Mathematical Psychology, 1990

R. Sun and D. Waltz, Neurally inspired Massively Parallel Model of Rule Based Reasoning, in: Neural and Intelligent System Integration, John Wiley & Sons, 1990

R. Sun, The Discrete Neuronal Models, Proc. INNC-Paris, 1990

R. Sun, Reasoning at conceptual and subconceptual level: a connectionist model, TR 90-136, Brandeis University, Computer Science Dept., 1990

A. Tversky, Features of Simiarity, Psychological Review, 84(4), 1977

D. Waltz and J. Pollack, Massive parallel parsing, Cognitive Science, No 9, 1985