Title
A Case for Symbolic/Sub-symbolic Hybrids

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 11(0)

Authors
Rose, Daniel E.
Belew, Richard K.

Publication Date
1989

Peer reviewed

A Case for Symbolic/Sub-symbolic Hybrids

Daniel E. Rose
Richard K. Belew

Cognitive Computer Science Research Group
Computer Science and Engineering Department
University of California, San Diego

Abstract

This paper considers the question of what qualities are necessary for an AI system to be a hybrid of symbolic and sub-symbolic approaches. Definitions of symbolic and sub-symbolic systems are given. SCALIR, a hybrid system for information retrieval, is presented, and then used to show how both symbolic and sub-symbolic processing can be combined. Arguments against SCALIR's hybrid nature are presented and rejected.

INTRODUCTION

Several times in the history of artificial intelligence (AI), researchers have divided up into rival camps arguing for or against certain approaches to various problems. The procedural vs. declarative knowledge dispute (Winograd, 1985) is one past instance of this. In many cases the disputes are resolved over time. One way this happens is that new conceptual frameworks arrive which encompass rival approaches, or simply show that one approach subsumes the other. Another way is for new techniques to be developed which merge the two sides.

Today, the rival camps are arguing for "symbolic" or "connectionist" AI systems. As supporters of the connectionist approach, we hope that the long-range solution to the debate will be the first one described above: symbolic processing will be shown to be an emergent property of large-scale sub-symbolic processing. In the mean time, however, we are investigating the second solution: developing a hybrid system which takes advantage of the strengths of both approaches.

SCALIR is a hybrid[1] system for legal informational retrieval currently being developed by one of the authors. When we first began to describe the system to others, we discovered an interesting phenomenon. Rather than arguing for or against the hybrid nature of the system, some members of our audience claimed that the system was not really a hybrid at all. Even more notably, some claimed that it was essentially just symbolic, while others said it was just connectionist.

In this paper we will outline the system in question, and then present and refute arguments challenging its dual symbolic and connectionist nature. In doing this, we will examine the question of what it means to be a symbolic or sub-symbolic system — and what it takes to be a hybrid.

THE SUB-SYMBOLIC/SYMBOLIC DICHOTOMY

The term "symbolic," as used in this paper, refers to the dominant approach to AI research for much of the past thirty years. From a cognitive perspective, the symbolic approach rests on Newell's Physical Symbol System Hypothesis (which says in essence that symbol manipulation is a necessary condition for intelligence) (Newell, 1980), but the use of symbolic AI systems predates this exposition of its premises.

Many traditional AI tools and techniques, such as expert systems, frames, and heuristic search do some form of symbol manipulation. We wish to characterize computational properties of symbolic systems, independent of accounts (such as Newell's) of their role in cognition.

[1] Throughout this paper, we use "hybrid" as shorthand for "symbolic/sub-symbolic hybrid."

We will focus on a narrow view of symbols rather than attempting to incorporate the vast philosophical tradition associated with them.

Our account begins with the following:

Definition 1 *A label is a unique identifier belonging to a previously enumerated, fixed set.*

With this definition in mind, we can say what it means to be a symbolic system:

Definition 2 *Symbolic systems are those in which the next state is selectively determined by labels associated with the objects of computation.*

This definition works for all systems generally considered symbolic, not just those with explicit logical rules. For example, the definition explains why a semantic network with spreading activation search is essentially a symbolic system. First, links as well as nodes in the network are labeled (and thus are symbols). Second, "activation" is typically a discrete marker (symbol), whose label affects its treatment. The decision to pass or not pass a marker depends solely on the links' labels.

For our other level of processing, we need another definition:[2]

Definition 3 *A class C is interchangeable if and only if all functions which operate on elements of C (i.e. whose domains are C^n for any integer n)*

1. *map only to range C, and*
2. *are infinitely many-to-one.*

Note that any computation using a member x of such a class will proceed identically regardless of which of the infinitely many possibilities which map to x were used, hence the name "interchangeable."

The term "connectionist" refers to the class of approaches based on neurally-inspired computer models. We will use the term to be representative of a broader class of sub-symbolic models which have the following property:

Definition 4 *Sub-symbolic systems are those in which the mechanism for mapping from input(s) to output(s)*

1. *is expressed in terms of interchangeable, continuous quantities, and*
2. *is modulated by continuous parameters determined by specific characteristics of the data (as opposed to general properties of the computational mechanism).*

Connectionist nets are fundamentally sub-symbolic; the interchangeable quantities is generally called "activations" and the modulating parameters are called "weights." But connectionist nets are not the only examples of sub-symbolic systems. The Classifier system (Holland et al., 1986) uses message "intensity" as its quantity and classifier "strength" as its modulating parameter, and so has sub-symbolic characteristics.

Armed with these definitions, we can now be clear about what we mean by "hybrid." A *hybrid system* is simply one which contains both symbolic and subsymbolic components. A degenerate case of a hybrid system is a connectionist net and an expert system put in the same box, each doing something different with the same input and producing its respective portion of the box's output. We will generally be interested in *tightly-coupled* systems, in which the components interact before the final output is produced.

OVERVIEW OF SCALIR

SCALIR (for Symbolic and Connectionist Approach to Legal Information Retrieval) is an AI system for full-text document retrieval. It uses techniques similar to Belew's AIR (Belew, 1986) but adds a symbolic mechanism useful for the legal domain. We claim that SCALIR is a tightly-coupled hybrid system.

The legal system has an interesting dual nature which makes it especially amenable to a hybrid approach. On the one hand, statutes are sets of rules to be applied in specific cases, and explicit symbolic relationships exist in almost every aspect of the law. On the other hand, the use of precedent to decide court cases means that the law is made in a parallel and distributed way; each case is decided on the basis of all relevant decisions in the past. Furthermore, statutes and court decisions are written in natural language, whose ambiguity makes it resistant to symbolic approaches. (These arguments are developed further in another paper (Rose and Belew, 1989)).

SCALIR's retrieval mechanism uses two interleaved networks, one connectionist and one symbolic. Nodes

[2] We use mathematical notation here for conciseness, not mathematical rigor.

representing terms, court cases, and statute sections are shared between the two networks. However, they are connected with two separate sets of links.

C-links are weighted, unlabeled connectionist links which use the microfeatures of the law (e.g. the individual words which appear in a court decision) to form associative relationships between legal documents. The resulting C-network is similar to the associative network in AIR.

S-links are labeled, unweighted symbolic links which use explicitly encoded knowledge about the law for inference. The S-links form a kind of semantic network which describes the relationships of the different types of nodes. For example, the well-known "key number" taxonomy of the law produced by West Publishing Co. provides a hierarchic structure for SCALIR's term nodes.

User queries can involve both associative and symbolic components. For example, the user could essentially ask the system to retrieve all cases which disputed a certain decision. Activity would propagate only through symbolic links which corresponded to negative citations, e.g. to an overruling case. From there it might spread to other cases generally related to the overruling ones.

Hybrid activity propagation is implemented by separating the activity into several components, which metaphorically can be viewed as different colored light. C-links are like grey filters which modify the intensity of the light. Each type of S-link is like a different colored filter, which allows only its corresponding type of light (i.e. activity) to pass. This means that the symbolic inference process — deciding whether or not to pass on activity depending on the type of S-link — can be done locally at each node. Figure 1 shows the process schematically.

Input to SCALIR takes the form of real-valued activations placed on a set of nodes. These activations spread through both C-links and S-links, and are combined numerically along the way. Finally, nodes which reach a high enough level of activity are considered outputs. Thus all processing debts are ultimately cashed in connectionist coin.

As in AIR, learning occurs as a result of negative and positive reinforcement from the user. Since there is no exact right answer in information retrieval, the user's browsing behavior becomes the feedback signal. In other words, when a user indicates that the search is to be pruned in a certain direction ("I don't want any more documents like this") or expanded ("I want to see more about this topic"), this results in negative or positive feedback to the system.

It is a relatively simple matter to use the feedback to train the weights on the C-links. S-links, on the other hand, do not have weights, since they represent explicit knowledge. Hybrid learning suffers from the traditional credit-assignment problem. When the system performs well (or poorly), how do we know whether the C-links or the S-links are primarily responsible? SCALIR's solution is to also learn the appropriate contributions of each component.

BUT IS IT CONNECTIONIST?

In this section we shall examine various versions of an argument that SCALIR is essentially a symbolic system. We will begin with general objections and move gradually to more specific ones:

> "Your system is simulated on a Von Neumann architecture using a program written in symbols. Therefore it is symbolic."

There is always an implementation level below the level of interest. One could equally well say that a theorem-prover was sub-symbolic because it depended completely on continuously varying electric fields in the circuits of the computer. But the most accurate description of the behavior of a theorem-prover is at the level of symbols. Similarly, SCALIR's C-network is best understood as sub-symbolic.

> "Even so, any processing done with symbolic nodes is symbolic processing."

If this were true, than essentially all connectionist systems would be symbolic. Inputs and outputs must represent something in the world in order to be useful, thus they are necessarily symbols. The designer of any connectionist net must explicitly code the meaning of input and output nodes. Having some nodes being symbolic does not make the system symbolic.

> "Yes, but the nodes in other connectionist systems form distributed representations, while SCALIR's are localist."

While distributed representations have many virtues in connectionist systems (see (Hinton et al., 1986), for example), this issue is a red herring with respect to the symbolic/sub-symbolic question. In fact, local is a subjective concept. For example, an ASCII code is a representation of a character distributed over seven bits. Yet some of the bits are localist representations of features of the character (such as case and printability).

> "Okay, but at least most connectionist systems have hidden nodes. SCALIR's nodes are all visible."

This claim presupposes a certain network architecture which SCALIR does not have. In networks with hidden units, such as layered feedforward nets or Boltzmann machines, these units are not accessible from the environment in any way. In this sense, it is true to say that SCALIR has no hidden units. However, in these systems, hidden units are usually defined as those which are neither inputs nor outputs. Input units and output units are all manipulated and examined every time the network is used. In SCALIR, only a fraction of the network is activated by a query as input, and only a fraction becomes active as output. But for the purposes of that query, all the remaining nodes in the network can serve as hidden units which do sub-symbolic processing.

> "But these so-called hidden units are still symbols. They represent features, rather than microfeatures."

Again, feature-hood is a subjective concept; one net's feature is another's microfeature. For example, a net trained to recognize handwriting might learn microfeatures corresponding to various arcs at various orientations, with letters as features to be detected. At the same time, a word-recognition net could use those letters as microfeatures. In SCALIR, terms are viewed as microfeatures of the law.

> "As long as the nodes have meaningful labels, *they are still symbols!*"

The question is not whether there are labels. The question is the labels are used in processing. SCALIR's connectionist component ignores the labels entirely.

It is a truism in connectionism that "the knowledge is in the weights." This being the case, it is the existence of these weights, communicating only interchangeable continuous activation, which should be our litmus test for sub-symbolic processing. The fact that the nodes are labeled is irrelevant. As Fodor and Pylyshyn explain:

> Strictly speaking, the labels play *no role at all* in determining the operation of the 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 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)

BUT IS IT SYMBOLIC?

As in the previous section, we will now consider more arguments which dispute the hybrid nature of SCALIR. This time, however, the claims are that SCALIR doesn't really do any symbolic processing. These arguments rest on some the previously noted observation that "all processing debts in SCALIR are ultimately cashed in connectionist coin."

We begin with what is essentially the complement of one of the arguments from the previous section:

> "Connectionism pervades even the allegedly symbolic parts of SCALIR. Therefore SCALIR does no symbolic processing."

To begin with, we will concede that connectionist "baggage" plays more of a role in SCALIR's symbolic component than the reverse. This is simply because activation is the *lingua franca* chosen for communication between the two components of the system. Nevertheless, the presence of one thing (traces of connectionism) does not prove the absence of another (symbolic processing).

> "But whatever the reason, it is real-valued activations, not symbols, which flow though S-links. What is symbolic about them?"

It is true that real-valued activations are passed along by S-links. In fact, to prevent activity from spreading to the whole network, S-links cause a slight attenuation of the activity, and can thus be considered to have a "weight" just like connectionist links. Despite this, two qualities make S-links symbolic.

First, *knowledge is not in the weights*. Not only are the weights unlearned, they are unrelated to the data. Weights in connectionist nets are either learned (e.g. via back-propagation) or are set to pre-computed values designed to produce a certain behavior with respect to the data. The "weights" on SCALIR's S-links are set at a fixed value designed only to prevent infinite spread of activity. This is similar to the constraints on the distance markers can be passed in some semantic network implementations.

Note that systems which use pre-computed weights may still be sub-symbolic. Two examples are Hopfield nets (in which weights are determined algorithmically) (Tank and Hopfield, 1987) and the interactive activation model of McClelland and Rumelhart (McClelland and Rumelhart, 1981; Rumelhart and McClelland, 1982) (in which the weights were set "by hand" in order to model certain empirical phenomena).

Second, *S-links respond selectively to symbols*. Specifically, the presence or absence of various types of activation determines whether those components are passed along those S-links. The filtering done by the S-links is a symbolic process, for it is exactly by virtue of having a specific label that an S-link allows or does not allow certain activity to pass. Returning to the comments of Fodor and Pylyshyn:

> ... [T]he state transitions of Classical [symbolic] machines are causally determined by *the structure ... of the symbol arrays that the machines transform*: change the symbols and the system behaves quite differently. (Fodor and Pylyshyn, 1988)

In SCALIR, the symbols being transformed are the symbolic components of activation at each node, and it is the S-links, by their filtering ability, that do the transformation.

THE LIMITS OF THE DICHOTOMY

While we have constructed our definitions of symbolic and sub-symbolic processing as robustly as possible, we do not believe the two approaches are mutually exclusive. In fact, there is a continuum from sub-symbolic to symbolic. In this section we will examine some of the harder cases which fall closer to the center of the continuum.

As explained in Section 2, semantic networks fall squarely into the realm of symbolic systems. What happens when we begin to add more "connectionist" attributes to a semantic network? (Systems with these attributes actually exist (Hendler, 1987).)

As a first step, suppose that the designer of a semantic network wishes to prevent too much of the network from being marked each time. We can imagine passing a number along with the marker. This number could be a counter for measuring path length. It could be incremented each time a link was traversed, and then used to terminate the search when it reaches a certain threshold.

Now, for computational simplicity, imagine that each node decrements this counter, rather than incrementing it, and stops if the value reaches zero. This way the parameter becomes easily tunable; the programmer can change the desired path length of searches by starting the initially marked nodes with various quantities in the counters.

Suppose that too many nodes are still being marked. The programmer might want to introduce a penalty for fan-out as well as path length. Each time markers leave a node, their counters can be set to the incoming marker's counter divided by the out-degree of the node.

One last modification: subtraction is too crude a control for path length; its effect is not proportional to the current magnitude of the counter. Instead of subtraction, we will multiply each counter by a value slightly less than one as it traverses a link. As an implementation detail to prevent roundoff errors, we will replace the integer counter with a real-valued one, and use real arithmetic for all our multiplications and divisions.

If we call the counters "activation" and the product of the divisors and the multipliers "weights", do we now have a connectionist system? Our claim is that we do not. As with SCALIR's S-links, the "weights" bear no relationship to the data; there is no knowledge in them. The currency of the system, markers, are not interchangeable, because the system responds selectively to them depending on the link labels. Symbolic inference remains the fundamental processing operation.

Now considering the other extreme, imagine a connectionist network in which each node in the input gets different kinds of activation — colored blue or yellow, perhaps. All computation is done in some standard connectionist fashion, except that active nodes become tagged with the color of their activation: blue, yellow, or green (where the colors have mixed). Nodes which become sufficiently active after a certain time (say, when the network reaches equilibrium) are considered outputs, with the following proviso: only green-tagged nodes are candidates for output. Do we now have a symbolic system?

In this case, we believe we have a tightly-coupled hybrid system. It still meets the conditions for sub-symbolic systems; its weights are either learned or constructed to produce a certain mapping on the data, using interchangeable activation. But it also meets the conditions for symbolic processing; the mapping from input to output depends on a differential response to labels.

There are many variations of this exercise, in which various attributes are added or removed to traditional symbolic or sub-symbolic systems. We will consider only one more case: a system which we claim lies at the boundary of the two approaches.

The system can be characterized in two ways. It is a semantic network in which there is only one kind of link (IS-A), and only one kind of marker. Alternatively, it is a connectionist network in which all nodes are localist and labeled, and all weights have the value one. Since there is only one type of link and marker, there can be no selective response on the basis of labels and the system is therefore not symbolic. Since all weights are equal and independent of the data, the network cannot do meaningful sub-symbolic processing. This illustrates that hybrid systems result from combining symbolic and sub-symbolic features, not by averaging them.

The Classifier system provides another example of a tightly-coupled symbolic/sub-symbolic hybrid. While this system shares many of the continuous, sub-symbolic qualities of connectionist nets, the fact that it also *broadcasts* messages globally (i.e., without the attenuation associated with path traversal) and typically performs a *discontinuous* match of messages with classifier conditions make the system have significant symbolic characteristics as well.

DISCUSSION

Along with its corresponding approaches to AI, the Symbolic/Sub-symbolic dichotomy is often described in terms of two views of cognition; Table 1 shows an informal characterization of the two views. Recently, many have suggested that both views are helpful in understanding cognition. As Norman explains:

> People interpret the world rapidly, effortlessly. But the development of new ideas, or evaluation of current thoughts proceeds slowly, serially, deliberately. People do seem to have at least two modes of operation, one rapid, efficient, subconscious, the other slow, serial, and conscious. (Norman, 1986, p. 542)

Since these two levels both have important roles to play, we believe it is useful (at least for the present) to design hybrid systems which take advantage of techniques designed for both levels. (Similar arguments have been made by other proponents of hybrid systems, such as Hendler (Hendler, 1989) and Dyer (Dyer, 1988).)

We have outlined some criteria for what it means to be a symbolic or sub-symbolic system, and what a hybrid of the two approaches might look like. Our work on SCALIR has given us an informal existence proof that such hybrids are feasible.

One practical benefit of a hybrid system is obvious: the techniques developed for the two paradigms have different strengths and weaknesses. In particular, connectionist systems are much better at learning, while it is much easier to store explicit knowledge in symbolic systems. In addition, we believe that hybrid systems will exhibit emergent properties not found in either of their single-paradigm components.

References

Belew, R. K. (1986). *Adaptive Information Retrieval: Machine Learning in Associative Networks.* PhD thesis, University of Michigan.

Dyer, M. G. (1988). Symbolic neuroengineering for natural language processing: A multilevel research approach. Technical Report UCLA-AI-88-14, Computer Science Department, University of California, Los Angeles.

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

Hendler, J. A. (1987). *Integrating Marker-passing and Problem Solving: A spreading activation approach* *to improved choice in planning*. Lawrence Erlbaum Associates, Hillsdale, NJ.

Hendler, J. A. (1989). Marker-passing over microfeatures. *Cognitive Science*. To appear.

Hinton, G. E., McClelland, J. L., and Rumelhart, D. E. (1986). Distributed representations. In Rumelhart, D. E. and McClelland, J. L., editors, *Parallel Distributed Processing: Explorations in the microstructure of cognition. Volume 1: Foundations*, pages 77–109. MIT Press, Cambridge, MA.

Holland, J. H., Holyoak, K. J., Nisbitt, R. E., and Thagard, P. (1986). *Induction: Processes of learning, inference and discovery*. Bradford Books, Cambridge, MA.

McClelland, J. L. and Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: Part 1. an account of basic findings. *Psychological Review*, 88:375–407.

Newell, A. (1980). Physical symbol systems. *Cognitive Science*, 2.

Norman, D. A. (1986). Reflections on cognition and parallel distributed processing. In McClelland, J. L. and Rumelhart, D. E., editors, *Parallel Distributed Processing: Explorations in the microstructure of cognition. Volume 2: Psychological and Biological Models*, pages 531–546. MIT Press, Cambridge, MA.

Rose, D. E. and Belew, R. K. (1989). Legal information retrieval: A hybrid approach. In *Proceedings of the Second International Conference on Artificial Intelligence and Law*, Vancouver, Canada. To appear.

Rumelhart, D. E. and McClelland, J. L. (1982). An interactive activation model of context effects in letter perception: Part 2. the contextual enhancement effect and some tests and extensions of the model. *Psychological Review*, 89:60–94.

Tank, D. W. and Hopfield, J. J. (1987). Collective computation in neuronlike circuits. *Scientific American*, 6(257).

Winograd, T. (1985). Frame representations and the declarative/procedural controversy. In Brachman, R. J. and Levesque, H. J., editors, *Readings in Knowledge Representation*, pages 357–370. Morgan Kaufmann, Los Altos, CA.

	SYMBOLIC	SUB-SYMBOLIC
AI Approach	Traditional	Connectionist/PDP
Inference	Rule-based	Statistical
Processing	Sequential	Parallel
Speed in brain	Slow (> 100ms)	Fast (< 100ms)
Robustness	Brittle	Graceful degradation
Precision	High	Low
Representation	Features	Microfeatures

Table 1: Comparison of two paradigms.

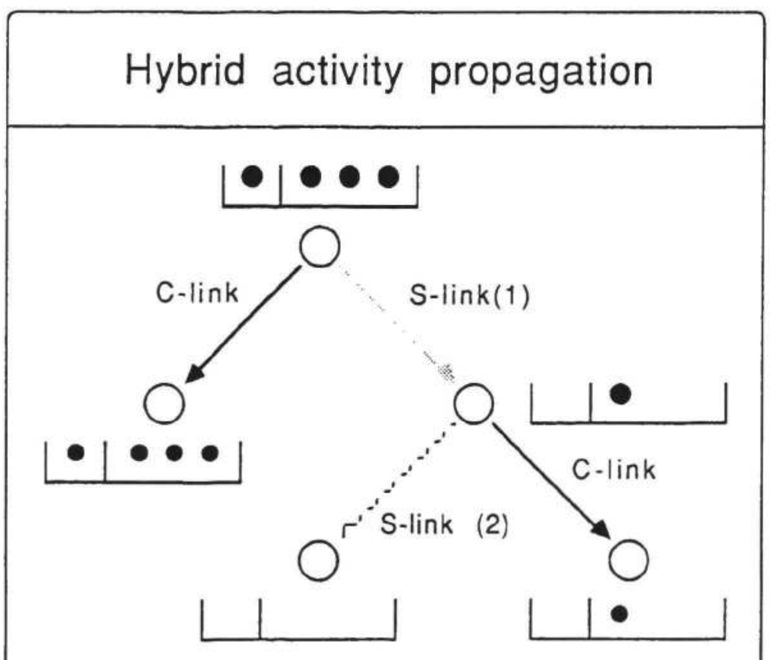

Figure 1: The row of filled dots represents different components of activation; the first is unspecified (i.e. strictly connectionist) while the others correspond to certain symbolic relationships. Larger dots indicate more activity. C-links allow all components to pass, attenuated by weight. S-links pass only the component of activity which corresponds to the link type.