Frame Selection in a Connectionist Model Of High-Level Inferencing

Trent E. Lange
Michael G. Dyer

Computer Science Department
University of California, Los Angeles

ABSTRACT

Frame selection is a fundamental problem in high-level reasoning. Connectionist models have been unable to approach this problem because of their inability to represent multiple dynamic variable bindings and use them by applying general knowledge rules. These deficits have barred them from performing the high-level inferencing necessary for planning, reasoning, and natural language understanding.

This paper describes a localist spreading-activation model, ROBIN, which solves a significant subset of these problems. ROBIN incorporates the normal semantic network structure of previous localist networks, but has additional structure to handle variables and dynamic role-binding. Each concept in the network has a uniquely-identifying activation value, called its *signature*. A dynamic binding is created when a binding node receives the activation of a concept's signature. Signatures propagates across paths of binding nodes to dynamically instantiate candidate inference paths, which are selected by the *evidential* activation on the network's semantic structure. ROBIN is thus able to approach many of the high-level inferencing and frame selection tasks not handled by previous connectionist models.

INTRODUCTION

High-level cognitive tasks, such as planning, reasoning, and natural language understanding, require the ability to perform inferencing to make explanations of and/or predictions from known states and actions. In natural language understanding, for example, a reader must often make multiple inferences to understand the motives of actors and to connect actions that are unrelated on the basis of surface semantics alone. Complicating the understanding process is the fact that language is often ambiguous on both the lexical and conceptual level. Consider the phrase:

P1: *"John put the pot inside the dishwasher"*

Most people will infer that John transferred a Cooking-Pot inside of a dishwasher in an attempt to get it clean. However, suppose P1 is followed by:

P2: *"because the police were coming."*

Suddenly, the interpretation selected for the word *"pot"* in P1 changes to Marijuana, and his Transfer-Inside action becomes a plan for hiding the Marijuana from the police.

The inferences needed to understand these two phrases (**Hiding Pot**) illustrate one of the fundamental problems in high-level inferencing, that of *frame selection*. When should a system make inferences from a given frame instantiation? Which of its related frames should it instantiate to make these inferences? Without being able to cope with these problems, a system will not be able to handle the following crucial tasks:

Word-Sense Disambiguation: Choosing the meaning of a word in a given piece of text. In P1, the word *"pot"* refers to a Cooking-Pot, but when P2 is presented, the evidence is that the interpretation should change to Marijuana.

Inferencing: Making inferences to understand the results of actions and the motives of actors. Nothing in **Hiding Pot** explicitly states that the police might see the pot, or even that the police will be in proximity to it and John. Nor is it explicitly stated what the police will do if they see he possesses Marijuana. All must be inferred from seemingly innocuous phrases P1 and P2.

Concept Refinement: Inferring a specific frame from a more general one. In P1, the fact that the pot was inside of a dishwasher told us much more than the simple knowledge that it was inside of a container. In **Hiding Pot**, however, the salient point is that it is inside of an opaque object, which allows us to infer that the police will not be able to see it.

Plan/Goal Analysis and Schema Instantiation: Recognizing the plan an actor is using to fulfill his goals. In P1, it appears that John put the pot into the dishwasher as part of the $Dishwasher-Cleaning script to satisfy his goal of getting it clean. In **Hiding Pot**, however, it appears that it is part of his plan to satisfy his sub-goal of hiding it from the police, which is part of his overall goal to avoid arrest.

Frame selection is complicated by the effect of additional context, which often causes reinterpretation to competing frames. The contextual evidence in **Hiding Pot** can conflict even more, and the explanation change again, if, for example, the next phrase is:

P3: *"They were coming over for dinner."*

As a result of P3, the word *"pot"* might be reinterpreted back to Cooking-Pot. These examples clearly point out two sub-problems of frame selection, *frame commitment* and *reinterpretation*. When should a system commit to one interpretation over another? And if it does commit to one interpretation, how does new context cause that interpretation to change?

PREVIOUS APPROACHES

Symbolic rule-based systems, such as BORIS [Dyer, 1983] and MOPTRANS [Lytinen, 1984], have had some success at performing the inferencing and frame selection necessary for high-level cognitive tasks. Their processing mechanisms, however, are often extraordinarily complex, being governed by large collections of brittle and sometimes ad-hoc rules that usually change with each type of knowledge structure modelled. Ambiguous input, such as that of **Hiding Pot**, has proven especially difficult for rule-based approaches, often requiring complicated and expensive backtracking rules when reinterpretation is required.

Distributed Spreading-Activation Networks

Distributed connectionist models, such as those of [McClelland & Kawamoto, 1986] and [Touretzky & Hinton, 1988], have lately been receiving much interest, mainly because of the learning algorithms available for their massively parallel networks of simple processing elements. Despite this attention, no distributed network model has yet exhibited the ability to handle inferencing having complexity even near that of **Hiding Pot**. The primary reason for this current lack of success is their inability to represent dynamic role-bindings and to propagate these binding constraints during inferencing. Distributed networks, furthermore, are sequential at the knowledge level and lack the representation of structure needed to handle complex conceptual relationships [Feldman, 1989].

Localist Spreading-Activation Networks

Localist spreading-activation models, such as those of [Cottrell & Small, 1983], [Waltz & Pollack, 1985], and [Shastri, 1988], also use massively parallel networks of simple processing units. Localist networks represent knowledge by simple nodes and their interconnections, with each node standing for a distinct concept. Activation on a conceptual node represents the amount of *evidence* available for that concept in the current context.

Unlike distributed networks, localist networks are parallel at the knowledge level and can represent structural relationships between concepts. Because of this, multiple inference paths are pursued simultaneously; a necessity to account for the understanding speed exhibited by people. Disambiguation is achieved automatically as related concepts under consideration provide evidence for and feedback to one another.

The main problem with previous localist models is that the evidential activation on their conceptual nodes gives no clue as to *where* that evidence came from. Because of this, previous localist models have had no more success than distributed models at handling dynamic non-local bindings — and thus remain unsuited to tasks requiring high-level inferencing.

Marker Passing Networks

Marker-passing models, such as those of [Granger *et al.*, 1986] and [Hendler, 1988], operate by spreading symbolic markers across semantic networks. Role-bindings are trivially represented using the symbolic pointers stored in their markers, whose propagation is used to generate plausible inference paths. Unfortunately, the logic and lisp-based symbolic mechanisms of existing marker-passing systems are far more complex than the simple processing units of spreading-activation networks. More importantly, marker-passing systems lack the natural constraint satisfaction abilities that allow localist networks to implicitly weigh contextual evidence in choosing a most highly-activated interpretation. They must therefore use a symbolic mechanism separate from the marker-passing process to apply a theorem prover and/or a heuristic path evaluator for path selection.

ROBIN

ROBIN (ROle Binding and Inferencing Network), is a localist spreading-activation model that has all of the advantages of previous localist approaches but, in addition, handles the problems of dynamic role-binding, inferencing, and frame selection. The localist networks in which ROBIN encodes its semantic networks consist entirely of connectionist units [Feldman & Ballard, 1982] that perform simple computations on their inputs: summation, summation with thresholding and decay, or maximization. Connections between units are weighted, and either excitatory or inhibitory.

ROBIN uses structured connections of nodes to encode frames [Minsky, 1975]. Each frame has one or more roles, with each role having expectations and logical constraints on its fillers. Every frame can be related to one or more other frames, with pathways between corresponding roles for inferencing. Activation spreads from frame to related frame when the constraints on their role fillers are met, thus automatically instantiating other frames and performing the processes of inferencing and frame selection.

As in previous localist models, ROBIN's networks have a node for every known conceptual frame in the network. Relations between concepts are represented by weighted connections between nodes. Activation on a conceptual node is *evidential*, corresponding to the amount of evidence available for the concept and the likelihood that it is selected in the current context.

Simply representing the amount of evidence available for a concept, however, is not sufficient for complex inferencing tasks. Role-binding requires that some means exist for *identifying* a concept that is being dynamically bound to a role in distant areas of the network. A network may have never heard about John having the goal of Avoid-Detection of his Marijuana, but it must be able to quickly infer just such a possibility to understand **Hiding Pot**.

Dynamic Role-Bindings With *Signature Activation*

To handle the problem of dynamic role-binding, every conceptual node in the network has associated with it a node outputting a constant, uniquely-identifying activation, called its *signature* [Lange & Dyer, 1989]. A dynamic binding is created when a role's *binding node* has an activation matching the activation of the bound concept's signature.

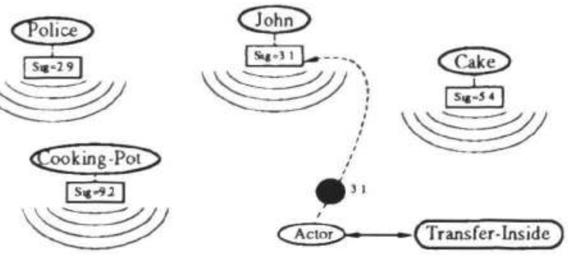

Figure 1. Several concepts and their uniquely-identifying signature nodes are shown, along with the Actor role of the Transfer-Inside frame. The dotted arrow from the binding node (black circle) to the signature node of John represents the virtual binding indicated by the shared signature activation, and does not exist as an actual connection.

In Figure 1, the virtual binding of the Actor role node of action Transfer-Inside to John is represented by the fact that its binding node, the solid black circle, has the same activation (3.1) as John's signature node. The complete Transfer-Inside frame is represented by the group of nodes that include the conceptual node Transfer-Inside, a conceptual node for each of its roles (only the Actor role shown), and the binding nodes for each of its roles.

Propagation of Signatures For Inferencing

The most important feature of signature activation is that it is spread across paths of binding nodes to generate candidate inferences. Figures 2a thru 2c illustrate how the network's structure automatically accomplishes this.

Evidential activation is spread through the paths between conceptual nodes on the bottom plane (i.e. Transfer-Inside and its Object role), while signature activation for dynamic role-bindings is spread across the parallel paths of corresponding binding nodes on the top plane. Nodes and connections for the Actor, Planner, and Location roles are not shown. Initially there is no activation on any of the conceptual or binding nodes in the network.

When input for P1 is presented, the lexical concept nodes for each of the words in the phrase are clamped to a high level of evidential activation, directly providing activation for concepts John, Transfer-Inside, Cooking-Pot, Marijuana, and Dishwasher.

To represent the role-bindings given by phrase P1, the binding nodes of each of Transfer-Inside's roles are clamped to the signatures of the concepts bound to them[1]. For example, the binding nodes of Transfer-Inside's Object are clamped to the activations (6.8 and 9.2) of the signatures for objects Marijuana and Cooking-Pot, representing the candidate bindings from the word *"pot"* (Figure 2a)[2].

The activation of the network's conceptual nodes is equal to the weighted sum of their inputs plus their previous activation times a decay rate, similar to the activation function of previous localist networks. The activation of the binding nodes, however, is equal

[1] ROBIN does not currently address the problem of deciding upon the original syntactic bindings, i.e. that *"pot"* is bound to the Object role of the phrase. Rather, ROBIN's networks are given these initial bindings and use them for high-level inferencing.

[2] An alternative input, such as *"John put the cake inside the oven"*, would be done simply by clamping the signatures of its bindings instead. A completely different set of inferences would then ensue. This is unlike previous localist models, where all instantiations must be hard-wired into the network.

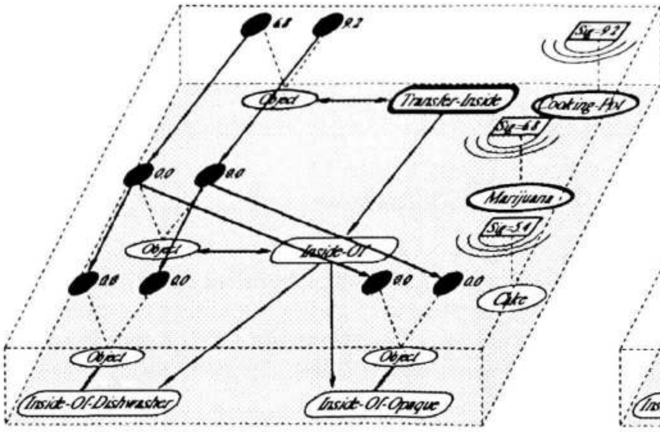

Figure 2a. Initial activation for P1.

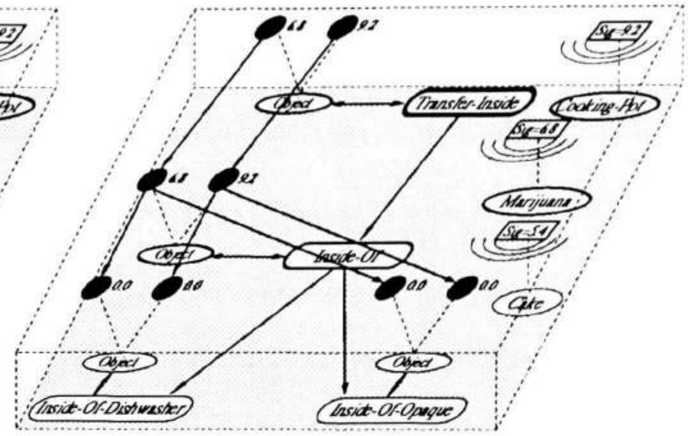

Figure 2b. After activation has reached Inside-Of.

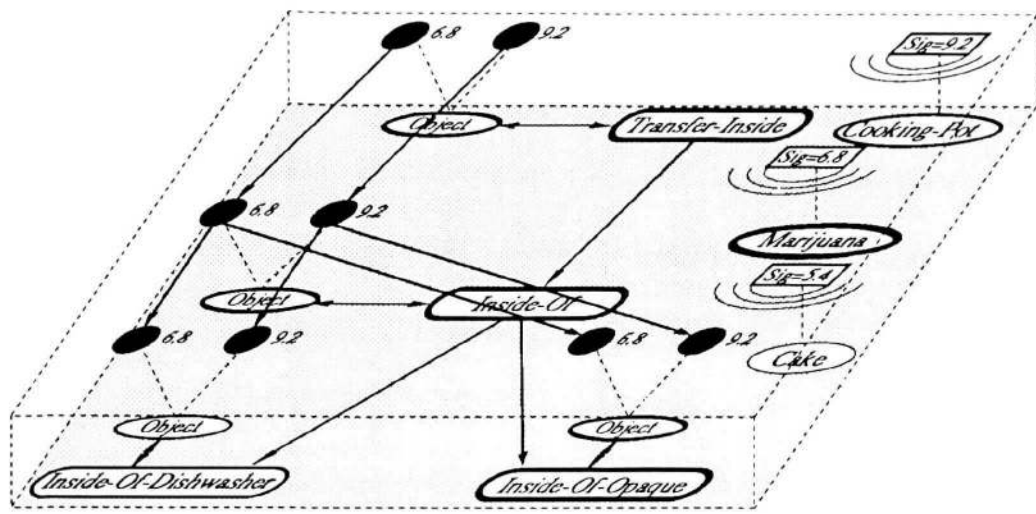

Figure 2c. Activation after quiescence has been reached in processing for **Hiding Pot**.

Figure 2. Simplified ROBIN network segment at three different cycles during processing of **Hiding Pot**. Each figure shows the parallel paths over which evidential activation (bottom plane) and signature activation (top plane) are spread for inferencing. Signature nodes (outlined rectangles) and binding nodes (solid black circles) are in the top planes. Thickness of conceptual node boundaries (ovals) represents their levels of evidential activation. (Node names do not affect the spread of activation in any way. They are simply used to initially set up the network's structure and to aid in analysis.)

to the maximum of their unit weighted inputs, allowing signatures to be propagated without alteration.

As activation starts to spread after the initial clamped activation values in Figure 2a, Inside-Of receives evidential activation from Transfer-Inside, representing the strong evidence that something is now inside of something else. Concurrently, the signature activations on the binding nodes of Transfer-Inside's Object propagate to the corresponding binding nodes of Inside-Of's Object (Figure 2b), since each of the binding nodes calculates its activation as the maximum of its inputs. The network has thus made the crucial inference of exactly which thing is inside of the other. Similarly, as time goes on, Inside-Of-Dishwasher and Inside-Of-Opaque receive evidential activation, with inferencing continuing by the propagation of signature activation to their corresponding binding nodes (Figure 2c).

Note that the actual activation values of signatures do not affect the network's processing. The signatures of Marijuana and Cooking-Pot were arbitrarily chosen to be 6.8 and 9.2 when the network was created, but could just as easily have been any other values. It is only necessary that each signature be different from all others — and so uniquely identify the concept bound to a role.

Frame	Binding Constraints	Used In
Inside-Of-Stove	(a Cooking-Pot is inside of a Stove)	$Stove-Cooking
Inside-Of-Dishwasher	(a Utensil is inside of a Dishwasher)	$Dishwasher-Cleaning
Inside-Of-Opaque	(a Phys-Obj is inside of an Opaque-Object)	Avoid-Detection

Figure 3. Three of the competing *refinements* of state Inside-Of.

FRAME SELECTION

Several paths of candidate inference chains are instantiated by the parallel propagation of signature and evidential activation. The path chosen as the network's interpretation at any given time is simply the one with the greatest evidential activation.

Consider how this process handles the problem of frame selection. Every frame in ROBIN's semantic knowledge base is related, through its roles, to one or more other frames. Some of those related frames compete, while others do not. The state Inside-Of, for example, has multiple concept *refinements*, three of which are described in Figure 3. No more than one of those refinements can be selected as the active *refinement* of a given instantiation of Inside-Of.

The mechanism described previously is sufficient for most examples of one or two phrases. Because of potential crosstalk from logically unrelated inferences[1], however, the network's structure is actually more complicated. Because of this, frame selection is a four part problem, controlled entirely by ROBIN'S structure of simple spreading-activation nodes:

1) *Choosing candidate frames:* When the role bindings of a frame match the logical binding constraints on the roles of a related frame, then that related frame becomes a *candidate frame* for instantiation. Related frames whose binding constraints are violated are rejected.

2) *Propagating bindings to candidate frames:* Candidate frames receive signature activation (representing role-bindings) from their instantiating frame. New candidate inferences can then propagate from each of the candidate frames to explore their respective inference paths.

3) *Propagating evidential activation to candidate frames:* Candidate frames receive weighted evidential activation from their instantiating frame. Candidates whose binding constraints are only partially matched receive proportionately less evidential activation than if their constraints were matched perfectly.

4) *Selection between candidate instantiation frames:* At any given time, the candidate frame with the most evidential activation represents the preferred interpretation. Commitments may change if new context gives more evidence to a competing frame.

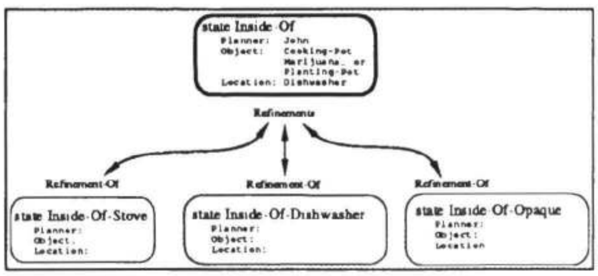

Figure 4. Overview of bindings instantiated with signature activation in Figure 2b.

As an example of how the frame selection process proceeds in ROBIN, consider Figure 4, which shows Inside-Of and three of its refinements. Evidential activation and signature role-bindings have reached Inside-Of (as in Figure 2b), so the candidates for its concept *refinement* need to be chosen. Inside-Of-Stove is rejected since a Dishwasher does not match the Stove constraint on its Location slot. Inside-Of-Dishwasher, however, is chosen as a candidate *refinement* frame, since its constraints are matched. Inside-Of-Opaque is also chosen as a candidate, since a Dishwasher *is-a* Opaque-Object.

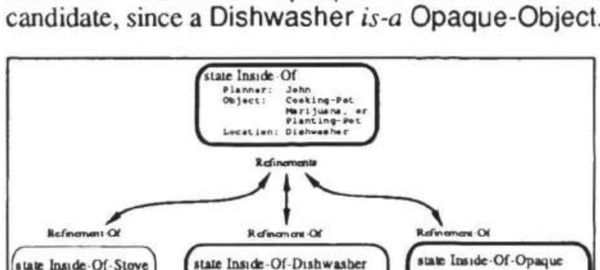

Figure 5. Overview after Inside-Of-Dishwasher and Inside-Of-Opaque become candidate *refinements* of Inside-Of (Figure 2c).

To implement this, the links allowing propagation of signature and evidential activation from one frame to another are gated by nodes that implement the frame selection process. Activation is only allowed to pass

[1]A problem not handled well by previous localist or marker-passing models.

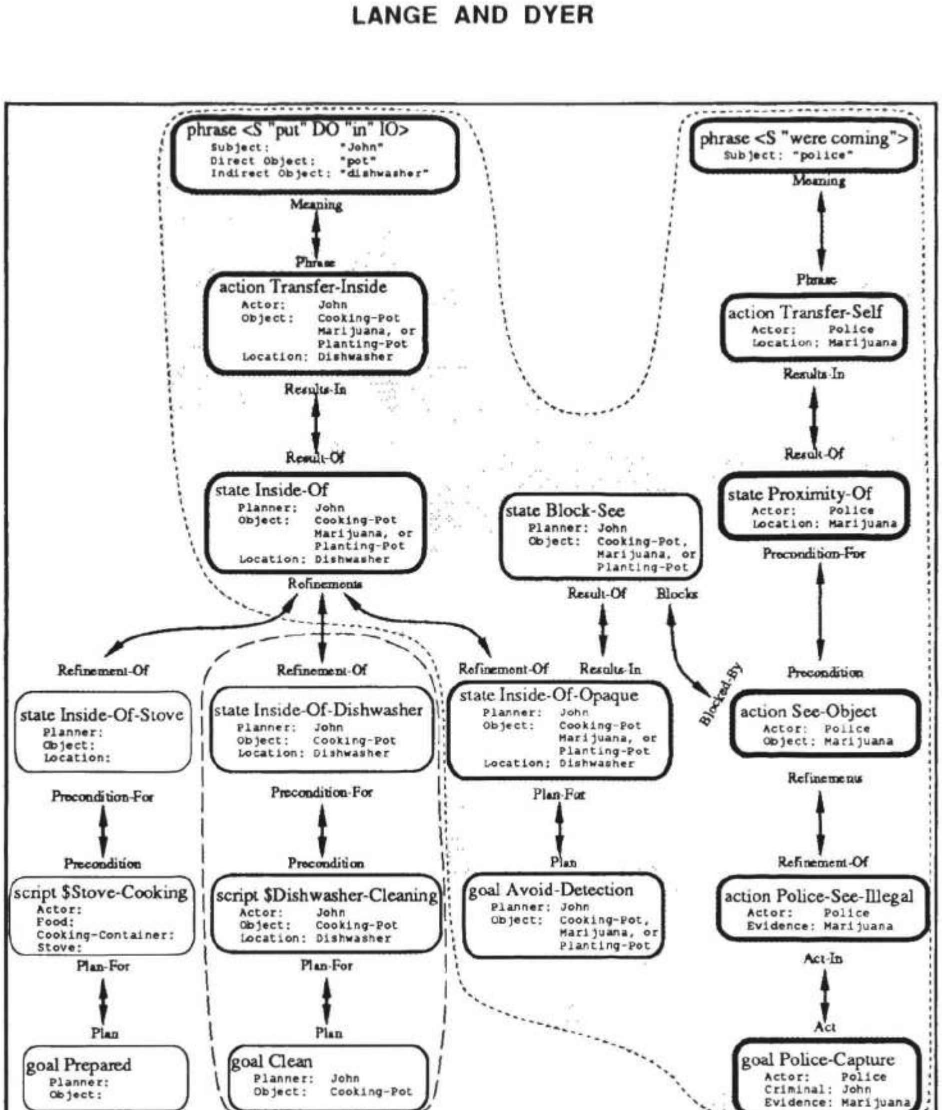

Figure 6. Overview of a small portion of a ROBIN semantic network showing inferences dynamically made after (syntactically pre-processed) input for phrases P1 and P2 of **Hiding Pot** have been presented. Thickness of frame boundaries shows the amount of evidential activation for the frames. Role fillers have been dynamically instantiated with signature activation. Darkly shaded area indicates the most highly-activated path of nodes representing the most probable plan/goal analysis of the input. Dashed area shows the discarded dishwasher-cleaning interpretation. Nodes outside of both areas show a very small portion of the rest of the network. These nodes received no evidential or signature activation from either phrase.

from a frame to one of its related frames when its role-bindings match the candidate frame's binding constraints. These logical binding constraints are calculated by groups of nodes that compare the frame's signature bindings to the candidate's binding constraints, and are described in [Lange, 1989].

As soon as Inside-Of-Dishwasher and Inside-Of-Opaque are chosen as candidate refinement frames, inhibitory gates (that disabled them from receiving signature or evidential activation from Inside-Of) are opened, performing steps 2 and 3 of the frame selection process. The result can be seen in Figure 5, where both have been instantiated. After activation has settled for **Hiding Pot**, Inside-Of-Opaque has the greater evidential activation (indicated by its thicker oval), and so is selected as the *refinement-of* Inside-Of, serving as the plan for hiding his Marijuana from the Police.

Selection of Ambiguous Role-Bindings

Note that all ambiguous meanings of a word are bound to a role with signature activation (Figures 2 thru 5). The network's interpretation of which binding is selected at any given time is the one whose concept

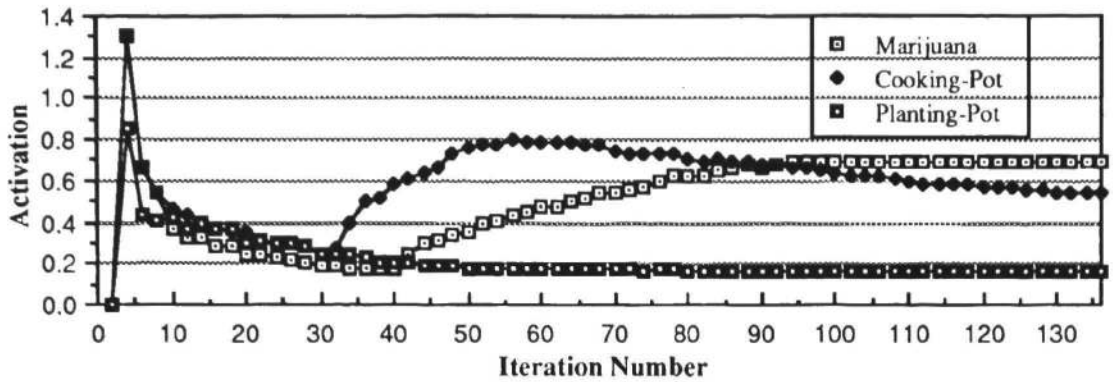

Figure 7. Evidential activations of the meanings of the word *"pot"* as activation spreads in **Hiding Pot**.

has greater evidential activation. Because all candidate bindings are propagated, with none being discarded until processing is completed, ROBIN is able to handle meaning reinterpretations without backtracking.

A DETAILED EXAMPLE

Figure 6 shows a segment of the semantic network embedded in ROBIN after input for both P1 and P2 have been presented and the network has reached stability, making the inferences needed to understand **Hiding Pot**. For example, the inference that the Marijuana is inside of an opaque object is represented by the instantiation of Inside-Of-Opaque. The role-bindings of the frames shown were instantiated dynamically with signature activation, with the final interpretation selected being the most highly-activated evidential path of frames inside the darkly shaded area.

During the interpretation of **Hiding Pot**, Cooking-Pot initially receives more evidential activation (Figure 7, cycles 40-70) than Marijuana by connections from the highly stereotypical usage of the Dishwasher for the Clean goal. The network's decision between the two candidate bindings at that point would be that it was a Cooking-Pot that was Inside-Of the Dishwasher. However, reinforcement and feedback from the inference paths generated by the Police's Transfer-Self eventually causes Marijuana to win out. The final selection of the Marijuana bindings over Cooking-Pot is represented simply by the fact that Marijuana has greater evidential activation. The resulting most highly-activated evidential path of frame nodes and propagated virtual bindings represents the interpretation of John hiding his Marijuana from the police (Figure 6).

Note that evidential activation remains on Cooking-Pot and Planting-Pot, available for possible reinterpretation given new input, such as phrase P3.

CURRENT STATUS AND FUTURE WORK

ROBIN has been fully implemented in the DESCARTES connectionist simulator[1] [Lange *et al.*, 1989]. ROBIN's inferencing, plan/goal analysis, schema instantiation, disambiguation, and reinterpretation abilities have been successfully tested on **Hiding Pot** and a number of other episodes in two domains, using syntactically preprocessed inputs of one or two sentences in length.

There are several directions for future work, including:

Signature dynamics: Currently, the identifying signatures are single arbitrary activations; instead, signatures should be distributed patterns of activation that are learned adaptively over time.

Embedded role-binding: Using signatures of pre-existing concepts, ROBIN can create and infer novel network instances. However, ROBIN currently cannot dynamically generate and propagate *new* signatures for one these instances. This ability is crucial for recursive structures, such as in: *"John told Bill that Fred told Mary that..."* Here each Object of the telling is itself a novel frame instance not having a pre-existing signature.

Network structure acquisition: Signatures allow ROBIN to create novel network instances over its pre-existing structure. The activation of these instances is transient. Over time, repeated instantiations should cause modification of weights and recruitment of underutilized units to alter network structure and create long-term memories.

---

[1] DESCARTES is a development environment that allows the flexible simulation of large-scale heterogeneous connectionist networks.

CONCLUSIONS

Inferencing and frame selection are fundamental problems in high-level reasoning. Unfortunately, previous connectionist models have been unable to approach these problems because of their inability to handle dynamic variable bindings and use them by applying general knowledge rules.

Although not completely solving the problems of role-binding and inferencing, we have presented a localist spreading-activation model that solves a significant subset of them. Using structure that holds *signature* activation, ROBIN is able to dynamically create novel frame instances by binding a role with any previously known concept in the network.

Since each signature is simply an activation value that uniquely identifies the concept bound to a role, it can be propagated across paths of binding nodes that preserve its activation, thus performing inferencing. This allows the encoding and "firing" of any general knowledge rule that states that the filler of one frame's roles can be inferred directly from the fillers of another. ROBIN's extra structure to handle dynamic variable-binding and rule-firing actually allows its networks to be smaller than other purely connectionist (non-marker passing) models, where all possible instantiations must be hard-wired into the network.

On the other hand, ROBIN's networks are not yet able to dynamically create new signatures, and thus cannot bind newly created recursive structures. This somewhat limits the model's inferencing capabilities in comparison to symbolic rule-based systems.

For the large portion of the inferencing process that it is able to handle, however, ROBIN has significant advantages over symbolic rule-based and marker-passing systems. The inherent constraint-satisfaction of ROBIN's normal *evidential* semantic network structure allows it to select the most plausible of the candidate frames and inference paths generated by the propagation of signature and evidential activation.

ROBIN is thus able to handle many of the high-level inferencing and frame selection tasks not approached by previous connectionist models, while at the same time perform disambiguation and semantic reinterpretation often difficult for symbolic systems.

Acknowledgements

This research has been supported in part by a contract with the JTF program of the DOD, and has been implemented on an Apollo DN4000 donated to UCLA by Apollo Computer Inc. Thanks to Jack Hodges, Eduard Hoenkamp, and the anonymous reviewers for their comments on previous drafts of this paper.

References

Cottrell, G. & Small, S. (1982): A Connectionist Scheme for Modelling Word-Sense Disambiguation. *Cognition and Brain Theory, 6*, p. 89-120.

Dyer, M. G. (1983): *In-Depth Understanding*. The MIT Press, Cambridge, MA.

Feldman, J. A., and Ballard, D. H. (1982): Connectionist Models and Their Properties. *Cognitive Science, 6:1*, p. 205-254.

Feldman, J. A. (1989): Neural Representation of Conceptual Knowledge. In L. Nadel, L. A. Cooper, P. Culicover, and R. M. Harnish (eds), *Neural Connections, Mental Computation*. The MIT Press, Cambridge, MA.

Granger, Eiselt, & Holbrook (1986): Parsing with Parallelism: A Spreading Activation Model of Inference Processing During Text Understanding. In *Experience, Memory, and Reasoning*, LEA Publishers, Hillsdale, New Jersey, p. 227-246.

Hendler, J. (1988): *Integrating Marker-Passing and Problem Solving*. LEA Publishers, Hillsdale NJ.

Lange, T. (1989): *High-Level Inferencing in a Localist Network*, Master's Thesis, Computer Science Dept., University of California, Los Angeles.

Lange, T. & Dyer, M. G. (1989): Dynamic, Non-Local Role-Bindings and Inferencing in a Localist Network for Natural Language Understanding. In David S. Touretzky (ed.), *Advances in Neural Information Processing Systems I*, p. 545-552, Morgan Kaufmann, San Mateo, CA.

Lange, T., Hodges J., Fuenmayor, M., & Belyaev, L. (1989): DESCARTES: Development Environment For Simulating Hybrid Connectionist Architectures. *Proceedings of the Eleventh Annual Conference of the Cognitive Science Society (CogSci-89)*, Ann Arbor, MI, August 1989.

Lytinen, S. (1984): Frame Selection in Parsing. *Proceedings of the National Conference on Artificial Intelligence (AAAI-84)*, August 1984.

McClelland, J. L. & Kawamoto, A. H. (1986): Mechanisms of Sentence Processing. In McClelland & Rumelhart (eds.) *Parallel Distributed Processing: Vol 2*. The MIT Press, Cambridge, MA.

Minsky, M. (1975): A Framework for Representing Knowledge. In P.H. Winston (ed.), *The Psychology of Computer Vision*. New York: McGraw-Hill, p. 211-277.

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

Touretzky, D. S. & Hinton, G. E. (1988): A Distributed Connectionist Production System. *Cognitive Science, 12*, p. 423-466.

Waltz, D. & Pollack, J. (1985): Massively Parallel Parsing. *Cognitive Science, 9:1*, p. 51-74.