Title
REMIND: Integrating Language Understanding and Episodic Memory Retrieval in a Connectionist Network

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 14(0)

Authors
Lange, Trent E.
Wharton, Charles M.

Publication Date
1992

Peer reviewed

REMIND: Integrating Language Understanding and Episodic Memory Retrieval in a Connectionist Network*

Trent E. Lange
Artificial Intelligence Lab, Department of Computer Science
University of California, Los Angeles

Charles M. Wharton
Department of Psychology
University of California, Los Angeles

Abstract

Most AI simulations have modeled memory retrieval separately from language understanding, even though both activities seem to use many of the same processes. This paper describes REMIND, a structured spreading-activation model of integrated text comprehension and episodic reminding. In REMIND, activation is spread through a semantic network that performs dynamic inferencing and disambiguation to infer a conceptual representation of an input cue. Because stored episodes are associated with the concepts used to understand them, the spreading-activation process also activates any memory episodes that share features or knowledge structures with the cue. After a conceptual representation is formed of the cue, the episode in the network with the highest activation is recalled from memory. Since the inferences made from a cue often include actors' plans and goals only implied in its text, REMIND is able to get abstract remindings that would not be possible without an integrated understanding and retrieval model.

Introduction

The most parsimonious account of comprehension and reminding is that they "amount to different views of the same mechanism" (Schank, 1982). Consider:

There were sightings of Great Whites off Newport, but Jeff wasn't concerned. The surfer was eaten by the fish. They found his board with a big chunk cut out.

When reading this passage, we may think of other stories of people being eaten by sharks, or, more abstractly, of others who knowingly ventured into mortal danger and died (e.g., skiers buried under avalanches). Why? In order to comprehend stories, a reader must find memory structures that provide inferences such as the goals and plans of story characters and the characteristic features of events and locations. Thus, while understanding a text, we may be reminded of analogous past episodes because they were understood with (and remembered with) the same knowledge structures.

In spite of the interweaving of comprehension and memory, AI simulations of memory have usually modeled reminding separately from language understanding. While this makes accounts of the phenomena more manageable, it is undeniable that real-world retrieval results from comprehension processes. Further, how an elaborated interpretation is constructed from a text will influence what is retrieved from memory. Consider:

John put the pot inside the dishwasher because the police were coming. **(Hiding Pot)**.

First appears John is cleaning a cooking pot, but later it seems he was hiding marijuana from the police to avoid being arrested. **Hiding Pot** might remind a person of superficially-similar stories involving police and marijuana. Or it might lead to more abstract remindings of hiding something to avoid punishment, such as *Billy put the Playboy under his bed so his mother wouldn't see it and spank him.* **(Dirty Magazine)**.

To retrieve episodes only related by similar plans and goals, a model must be able to infer them in the first place. As **Hiding Pot** shows, such interpretations often require both the ability to make multiple inferences and resolve ambiguities. Only with such language understanding capabilities can a retrieval model go directly from input texts to remindings of episodes that are analogous in terms of inferred plans, goals, and abstract relationships. Thus, a model that integrates the process by which a retrieval cue is understood with the process by which it is used to recall information can make an important contribution to the understanding of episodic reminding. In this paper we describe REMIND (Retrieval from Episodic Memory through INferencing and Disambiguation), a structured connectionist model that integrates language understanding and memory retrieval.

Previous Memory Retrieval Models

Memory retrieval has generally been explored in isolation from the process of language understanding. In case-based reasoning (CBR) models (cf. Hammond, 1989; Riesbeck & Schank, 1989), memory access is performed by recognition of meaningful *index patterns* that allow retrieval of episodes (or cases) most likely to aid their current task. CBR models are therefore generally models of expert reasoning within a given domain, rather than models of general human reminding. Whereas expert memory retrieval may be satisfactorily modeled by only retrieving cases matching expected indices within the domain of interest, general reminding

* This research was supported by NSF grant DIR-9024251, Army Research Institute contract MDA 903-89-K-0179, and by a grant from the Keck Foundation.

seems to be substantially "messier", being affected not only by the sort of useful abstract indices used in CBR models, but also by superficial semantic similarities that often lead to quite inexpert remindings. Further, the problem of selecting and recognizing appropriate indices becomes substantially more difficult when reading ambiguous texts outside of limited expert domains.

General, non-expert reminding has been modeled in systems such as MAC/FAC (Gentner & Forbus, 1991) and ARCS (Thagard et al., 1990). These systems model retrieval without using specific indexing methods. Instead they retrieve episodes whose representations share superficial similarities with cues, with varying degrees of preference towards episodes that are also analogically similar. However, unlike CBR models, these systems do not specify how they construct the representation of cues from a source input or text, and so cannot explain how inferences and comprehension affect reminding.

Previous Language Understanding Models

In REMIND, the understanding mechanism constructs an interpretation from its input that not only serves as the model's representation of the meaning of the text, but also serves as an elaborated cue for episodic memory retrieval. Symbolic, rule-based systems have had some success performing the inferencing necessary for this, but have substantial difficulties with ambiguous texts.

Distributed connectionist models can be trained to perform disambiguation and understand script-based stories (c.f. Miikkulainen & Dyer, 1991; St. John, in press). However, it is unclear whether they can be scaled up to handle language that requires the inference of causal relationships between events for completely novel stories. This requires chains of *dynamic inferences* over simple known rules, with each inference resulting in a potentially novel intermediate state. Other distributed connectionist models are able to partially handle this problem by explicitly encoding variables and rules in the network (c.f. Touretzky & Hinton, 1988). Unfortunately, these models are *serial at the knowledge level*, i.e. they can only select and fire one rule at a time, a serious drawback for language understanding, in which multiple alternative interpretations must often be explored in parallel (Lange & Dyer, 1989).

Marker-passing models (c.f. Riesbeck & Martin, 1986; Norvig, 1989) solve many of these problems by spreading symbolic markers across semantic networks in which concepts are represented by individual nodes. Because of this, they can perform dynamic inferencing and pursue multiple candidate interpretations of input in parallel as markers propagate across different parts of the network. A drawback of marker-passing models is that they must use a separate serial path evaluator to select the best interpretation path among the often large number of alternative paths generated. This is a particularly serious problem for ambiguous text in which many alternative paths must be evaluated (Lange, 1992).

Structured connectionist networks (c.f. Waltz & Pollack, 1985) are well-suited to disambiguation because it is achieved automatically as related concepts provide graded activation evidence and feedback to one another in a form of constraint relaxation. Like marker-passing models, they have the potential to pursue multiple candidate interpretations in parallel, since each interpretation is represented by activation in different local areas of the network. A number of researchers have recently shown how structured connectionist models can perform dynamic inferencing (c.f. Ajjanagadde & Shastri, 1989; Barnden, 1990). However, most of these new models no longer perform disambiguation. An exception is ROBIN (Lange & Dyer, 1989; Lange, 1992), a structured spreading-activation model that propagates *signature* activation patterns to perform dynamic inferencing and generate multiple possible interpretations of an input in parallel. At the same time, ROBIN uses the network's evidential constraint satisfaction to perform lexical disambiguation and select the contextually most plausible interpretation. This makes ROBIN a promising start for an integrated understanding and memory retrieval model.

Language Understanding in REMIND

REMIND is a structured spreading-activation model that integrates aspects of the language understanding and memory retrieval problems. REMIND is an extension of ROBIN, whose capabilities allow it to perform the high-level inferencing and disambiguation necessary to build interpretations of syntactically-parsed input for short texts such as **Hiding Pot** and **Dirty Magazine**. These interpretations are added to the network to represent the model's long-term memory episodes.

In REMIND, memory retrieval is a natural side-effect of the spreading-activation understanding process. The knowledge structures used to understand an input cue activate similar episodes that were understood and stored in the network earlier. For example, **Dirty Magazine** becomes active when **Hiding Pot** is being understood. An episode is retrieved from memory when there are enough similarities between it and a cue's interpretation to make it the most highly-active episode. Because inferencing and retrieval occur within a single spreading-activation network, both processes strongly interact and affect each other, as appears to be the case in human memory. In the following section, we give an overview of the language understanding portion of the model.

Knowledge Given To REMIND

As with ROBIN, REMIND uses structured networks of simple connectionist units to encode semantic networks of frames and rules representing world knowledge, such as the scripts, plans, and goals (Schank, 1982) necessary for understanding stories in a limited domain. Its knowledge base is hand-built, as in most structured models. However, it is given no information about specific episodes that the network is used to understand.

The knowledge given to REMIND is used to *construct* the actual structure of the network before any processing begins. As with other structured connectionist models, nodes in the network represent each frame or role. Rela-

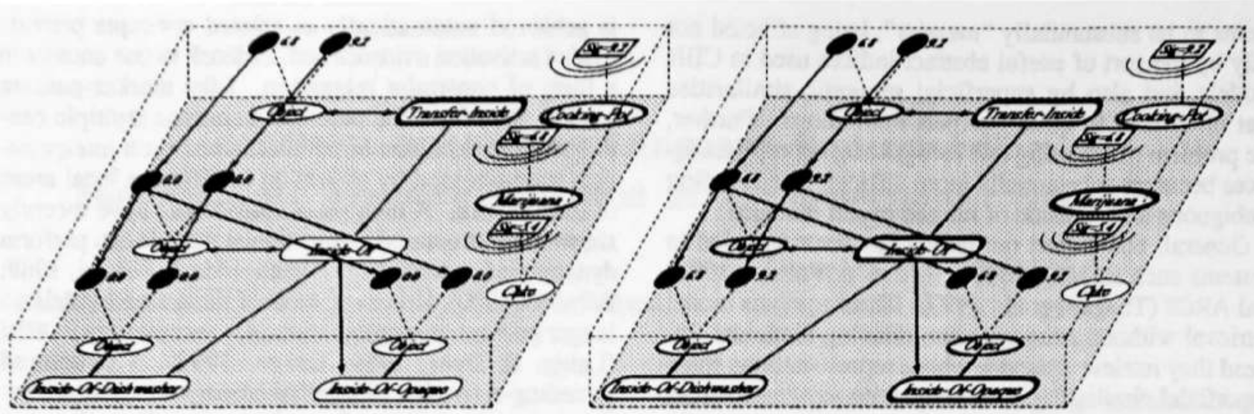

Figure 1. Simplified REMIND network segment in processing for **Hiding Pot**. **(a)** After initial clamping of signatures and evidential activation. **(b)** After network has settled. Figures show parallel paths over which evidential activation (bottom plane) and signature activation (top plane) spread. Signature units (outlined rectangles) and binding units (solid black circles) are in the top plane. Thickness of concept nodes (ovals) represents their evidential activation. Shown are two binding units per role; actual network has enough to hold meanings of network's most ambiguous word.

tions between concepts are represented by weighted connections between the nodes. Activation on concept nodes is *evidential*, corresponding to the amount of evidence available for them in the current context. However, as described earlier, simply representing the amount of evidence available for concepts is insufficient for language understanding. Solving the variable binding problem requires a means for *identifying* the concept dynamically bound to a role. Furthermore, the network's structure must allow these bindings to propagate across the network to dynamically instantiate inference paths and form an elaborated representation of the input.

Dynamic Inferencing With Signatures

Variable bindings are handled in REMIND by network structure holding *signatures* — activation patterns that uniquely identify the concept bound to a role (Lange & Dyer, 1989). Every concept has a set of *signature units* that output its signature, a constant activation pattern different from all other signatures. A dynamic binding exists when a role's *binding units* hold an activation pattern matching the bound concept's signature.

Figure 1a shows a simplified portion of REMIND after **Hiding Pot** has been input. The nodes in the lower layer of the network form a normal semantic network whose weighted connections represent world knowledge. The knowledge represented here is that: (a) transferring an object inside of another (Transfer-Inside) results in it being inside it (Inside-Of), and (b) that two possible concept refinements (or reasons) for it being inside are (1) because it is inside of a dishwasher (Inside-Of-Dishwasher), which will lead to further inferences about it being cleaned, or (2) because it is inside of an opaque object (Inside-Of-Opaque), which will lead to inferences about it being hidden.

Signature activations for variable binding and inferencing are held by the black binding units in the top plane of Figure 1a. In this simplified example, signatures are arbitrary scalar activation values. Here Marijuana is signified by 6.8, Cooking-Pot by 9.2, and Cake by 5.4. As shown, unit-weighted connections between binding units allow signatures to be propagated to other roles defined by general knowledge rules. For example, there are connections from the binding units of Transfer-Inside's Object to the respective binding units of Inside-Of's Object, since the object transferred inside is always the object that ends up inside.

To represent *John put the pot inside the dishwasher*, Transfer-Inside is clamped to a high level (dark oval in Figure 1a). Since Transfer-Inside's Object is either Marijuana or a Cooking-Pot, its Object's binding units is clamped to their signatures' activations, 6.8 and 9.2, respectively[1]. Similarly, one of the binding units of its Actor role is clamped to John's signature and of its Location role clamped to Dishwasher's signature.

Once the activations of the initial signature bindings and conceptual nodes are clamped, both types of activation spread through the network. Figure 1b shows the result after the network has settled from the inputs of Figure 1a and the rest of **Hiding Pot**. The signature activations representing the bindings have propagated along paths of corresponding binding units, so that the network has inferred that the Cooking-Pot or Marijuana is Inside-Of the Dishwasher. This is shown by the fact that their signatures are on the appropriate binding units. As can be seen, the propagation of signatures has also instantiated two different candidate interpretation paths. One path goes through Inside-Of-Dishwasher and continues through other cleaning frames such as $Dishwasher-Cleaning and Clean. Another path goes through Inside-Of-Opaque and continues through frames representing the object being blocked from sight (Block-See), the goal of avoiding detection (Avoid-Detection), and so on. Figure 2 shows a partial overview of the rest of the network.

[1] Other bindings can be presented by simply clamping the binding units to the activations of different signatures. REMIND does not currently address the problem of deciding upon the original syntactic bindings, e.g. that the *"pot"* is bound to Transfer-Inside's Object role.

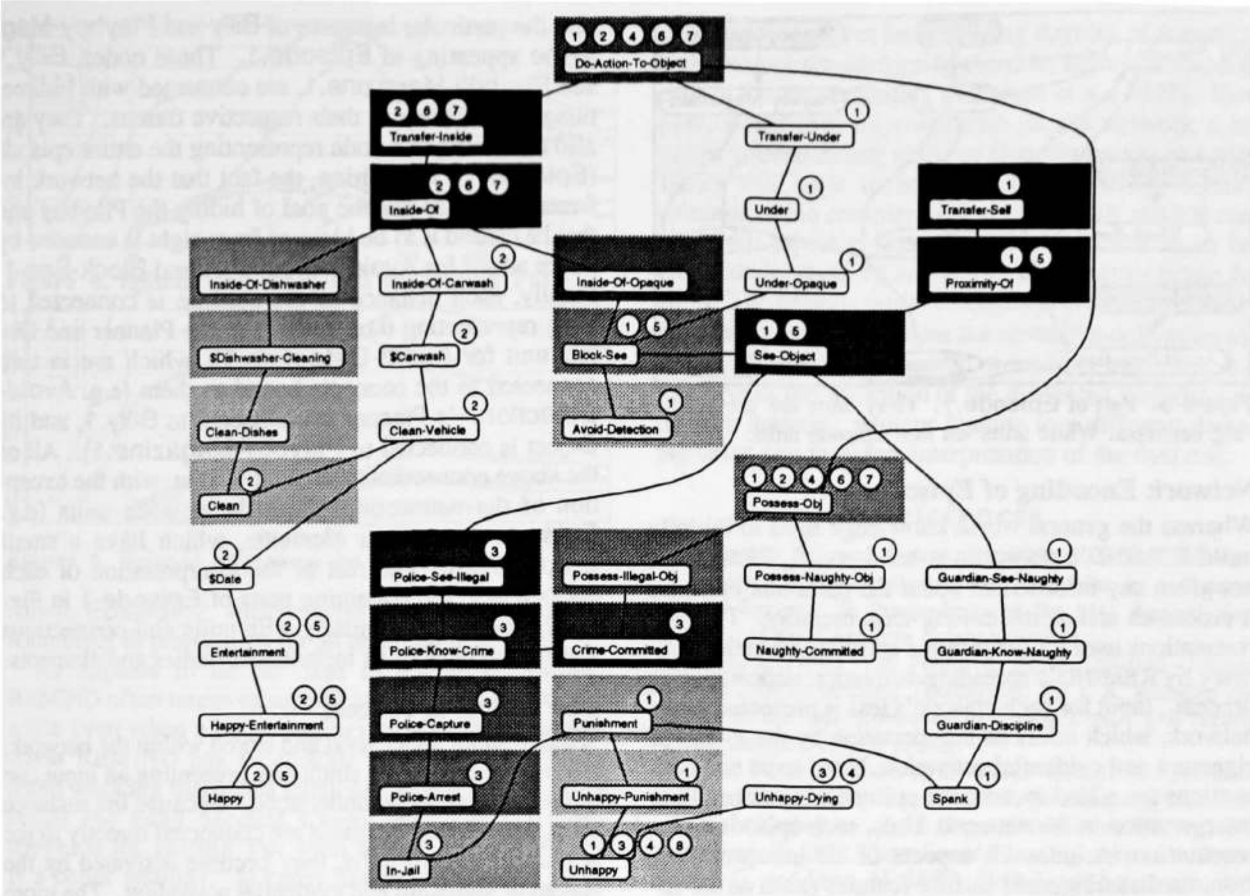

Figure 2. Overview of part of the network after activation has settled in processing **Hiding Pot**. Gray boxes represent level of evidential activation on the frames (darker = higher activation). Circles above frames indicate a long-term instance of that frame in an episode. Episodes understood and stored here: 1: **Dirty Magazine**. 2: *Fred put his car inside the car wash before his date with Wilma* (**Car Wash**). 3: *Jane shot Mark with a Colt-45. He died.* 4: *Betty wanted to smoke a cigarette, so she put it on top of the stove and lit it.* 5: *The pleasure boat followed the whales to watch them.* 6: *Barney put the flower in the pot, and then watered it.* 7: *Mike was hungry. He ate some fish.* 8: *Suzie loved George, but he died. Then Bill proposed to her. She became sad.*

This view includes the other frames instantiated by the propagation of signatures from Figure 1a and the clamped input for the remainder of **Hiding Pot** starting from Transfer-Self (*the police were coming*).

At the same time as signatures propagate to perform inferencing, activation spreads and accumulates along the bottom layer of conceptual nodes to disambiguate between those inferences. Initially the Inside-Of-Dishwasher path receives the most evidential activation because of feedback between it and its strong stereotypical connections to Cooking-Pot and Dishwasher. However, activation feedback between Inside-Of-Opaque and inferences from the police coming (Transfer-Self...Block-See) and the Police-Capture frames causes Inside-Of-Opaque to end up with more activation than Inside-Of-Dishwasher and Marijuana to end up with more activation than Cooking-Pot.

The network's final interpretation of **Hiding Pot** includes the most highly-activated path of frames in Figure 2 and their signature bindings. This interpretation includes the inferences that (a) Marijuana was inside of an opaque dishwasher (Inside-Of-Opaque) and has been blocked from sight (Block-See), (b) John possesses illegal marijuana (Possess-Illegal-Obj), and (c) John is in danger of being arrested by the police (Police-Arrest). Note that alternative interpretation paths retain activation for future possible reinterpretation, since REMIND uses a form of inhibition that normalizes activations rather than driving losers to zero. See Lange & Dyer (1989) and Lange (1992) for further details on how the network performs such inferencing and disambiguation for **Hiding Pot** and other inputs.

Memory Retrieval

In REMIND, memory retrieval occurs automatically as a side-effect of the spreading-activation understanding process. Representations of previously-understood episodes are connected directly to the semantic network that understood them in the first place. This direct form of "indexing" causes episodes that share many conceptual similarities with the cue to become active as REMIND interprets it. The most active (and hence most similar) episode gets chosen as the retrieved episode.

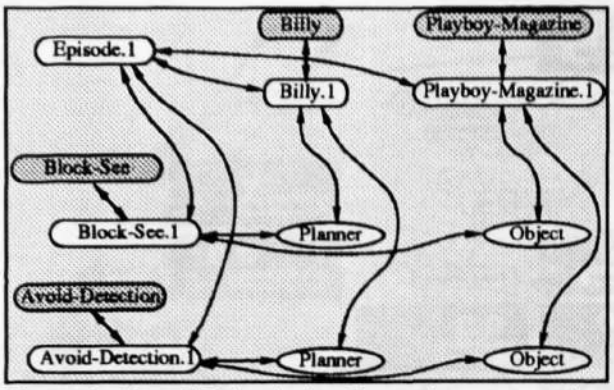

Figure 3. Part of Episode.1. Gray units are pre-existing concepts. White units are new episode units.

Network Encoding of Episodes

Whereas the general world knowledge used to initially build REMIND's networks is hand-coded, REMIND is not given any information about the particular episodes it processes and stores in long-term memory. The representations used for these target episodes are created entirely by REMIND's spreading-activation understanding process. Input for each episode's text is presented to the network, which infers an interpretation by the spread of signature and evidential activation. Next, units and connections are added to store the episode's entire resulting interpretation in the network. Thus, each episode's representation includes all aspects of its interpretation, from its disambiguated surface features (such as the actors and objects in the story) to the plans and goals that the network inferred that the actors were using.

As an example, consider how **Dirty Magazine** (*Billy put the Playboy under his bed so his mother wouldn't see it and spank him)* is processed and stored in the network as a memory episode. First, input for its phrases is clamped and an interpretation inferred, as described for **Hiding Pot**. As in **Hiding Pot,** the network infers that somebody is hiding something (Avoid-Detection) and that it is blocked from sight (Block-See). Here, however, the inferred signatures show that it is Billy hiding a Playboy-Magazine rather than John hiding Marijuana. Several other knowledge structures involved in **Hiding Pot** (e.g. Proximity-Of, Possess-Obj, Punishment) are also activated by **Dirty Magazine.** However, there are a number of differences, e.g. frames of the Guardian-Discipline structure are part of **Dirty Magazine**'s interpretation, but the Police-Capture frames are not. The rest of the frames activated as part of **Dirty Magazine**'s interpretation are shown by nodes that have a circled "1" above them in Figure 2. Other circled numbers represent elements of other stored episodes' interpretations.

To encode an episode after interpreting it, units are added to the network (by hand) for each of its interpretation's elements. Figure 3 shows a simplified part of the network's evidential layer after **Dirty Magazine** (Episode.1)'s interpretation has been added to the network. As can be seen, nodes have been added to represent the particular instances of Billy and Playboy-Magazine appearing in Episode.1. These nodes, Billy.1 and Playboy-Magazine.1, are connected with bidirectional connections to their respective frames. They are also connected to a node representing the entire episode (Episode.1). In addition, the fact that the network inferred that Billy had the goal of hiding the Playboy and that he caused it to be blocked from sight is encoded by nodes added for Avoid-Detection.1 and Block-See.1. Finally, each instance in the episode is connected to units representing their roles (e.g. the Planner and Object unit for Avoid-Detection.1), which are in turn connected to the concepts bound to them (e.g. Avoid-Detection.1's Planner is connected to Billy.1, and its Object is connected to Playboy-Magazine.1). All of the above connections have unit weight, with the exception of the connections from the episode units (e.g. Episode.1) to their elements, which have a small weight (0.05). The rest of the interpretation of each episode (e.g. the remaining parts of Episode.1 in Figure 2) is encoded similarly with units and connections that represent all of its instantiated frames and elements.

The Retrieval Process

With episodes understood and stored within the network, retrieval is performed simply by presenting an input cue to the network to be understood. Because the instance units representing episodes are connected directly to the normal evidential units, they become activated by the spread of signature and evidential activation. The more similarities an episode shares with the inferred interpretation of a cue, the more of its instances become active and the more activation its episode unit receives.

Figure 4 shows activations of the eight episodes from Figure 2 during understanding of **Hiding Pot**. Episode.6 (*Barney put the flower in the pot, and then watered it*) initially becomes highly active because it shares a number of surface features with **Hiding Pot** — e.g. both involve a Transfer-Inside, both have humans, and Planting-Pot is activated from the word *pot*. Similarly, Episode.2 and other episodes having varying degrees of shared features become active. However, as time goes on, the hiding and punishment frames are inferred and become active. Because of this, Episode.1 (**Dirty Magazine**)'s activation climbs and eventually wins, because it shares the most surface *and* abstract features of any episode with **Hiding Pot**'s interpretation (see Figure 2). **Dirty Magazine** is therefore retrieved as the episode most similar to **Hiding Pot**.

An example of how strongly the inferencing and disambiguation of the model affects retrieval is shown in Figure 5, which shows activations after presentation of input for *John put the pot inside the dishwasher because company was coming* (**Dinner Party**). Note that although this cue differs from **Hiding Pot** by only a single word (*company* instead of *police*), the interpretation REMIND reaches is completely different (i.e. that he was trying to clean a cooking pot to prepare for a dinner party). This causes a different episode to be re-

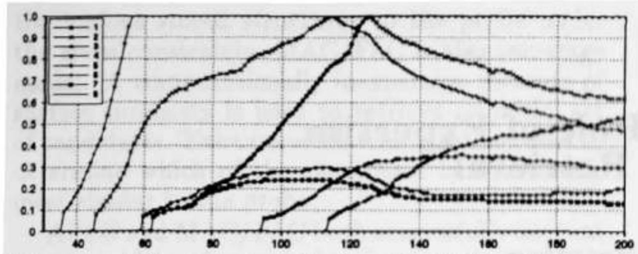

Figure 4. Episode unit activations for **Hiding Pot**.

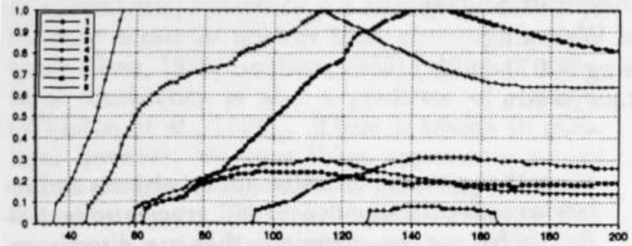

Figure 5. Episode activations for **Dinner Party**.

called, Episode.2, which shares the goals of cleaning something in preparation for an entertainment event.

As appears to be the case in human reminding, REMIND often retrieves superficially similar episodes to a cue even when a better analogy exists. For example, when REMIND has an episode in memory explicitly about smoking marijuana (such as *Cheech put the grass inside the bong because Chong was coming*), it retrieves it for **Hiding Pot** even though **Dinner Party** is a better analogy. In cases like this, REMIND often retrieves a superficially similar episode even though it has a different goal structure than the cue's interpretation and a better analogy exists in memory.

Discussion

Theoretically, REMIND lies somewhere between case-based reasoning models and general analogical retrieval models such as ARCS and MAC/FAC. Like ARCS and MAC/FAC, REMIND is meant to be a psychologically-plausible model of general human reminding, and therefore takes into account the prevalence of superficial feature similarities in remindings. However, we believe that many of the types of high-level planning and thematic knowledge structures used as indices in case-based reasoning systems also have an important effect on reminding. REMIND is thus partially an attempt to bridge the gap between case-based and analogical retrieval models. As it turns out, this gap is naturally bridged when the same spreading-activation mechanism is used to both understand cues and retrieve episodes from memory. Using the same mechanism for both processes causes retrieval to be affected by all levels that a text was understood with. This is the case in REMIND, in which the understanding mechanism is given the superficial features and actions of a text and attempts to explain them by inferring the plans and goals being used — causing memory episodes to be activated by both. This seems to give a more psychologically-plausible form of reminding than previous models, because the episodes it retrieves have varying degrees of superficial and abstract similarities to the cue, as seems to be the case in human reminding (Wharton et al., 1992). However, significant improvements in the network's language understanding abilities (see discussion in Lange, 1992) will have to be made before it can retrieve episodes of the complexity that some CBR models can.

A final aspect to note about REMIND is how its language understanding and retrieval processes come full circle. The episode retrieved depends crucially on the interpretation of the cue from the spreading-activation network's inferences. Once an episode is retrieved, it in turn primes the activation of the evidential spreading-activation network, perhaps leading to a different disambiguation and therefore interpretation of the next cue.

References

Ajjanagadde, V., and Shastri, L. 1989. Efficient Inference with Multi-Place Predicates and Variables in a Connectionist System. In Proceedings of the 11th Annual Conference of the Cognitive Science Society.

Barnden, J. 1990. The Power of Some Unusual Connectionist Data-Structuring Techniques. In J. A. Barnden and J. B. Pollack eds. *Advances in Connectionist and Neural Computation Theory*, Ablex.

Gentner, D., and Forbus, K. D. 1991. MAC/FAC: A Model of Similarity-Based Retrieval. *Proceedings of the 13th Annual Conference of the Cognitive Science Society*.

Hammond, K. 1989. *Case-Based Planning*. Academic Press.

Lange, T., and Dyer, M. 1989. High-Level Inferencing in a Connectionist Network. *Connection Science* 1:181-217.

Lange, T. 1992. Lexical and Pragmatic Disambiguation and Reinterpretation in Connectionist Networks. *International Journal of Man-Machine Studies* 36:191-220.

Miikkulainen, R., and Dyer, M. G. 1991. Natural Language Processing with Modular PDP Networks and Distributed Lexicon. *Cognitive Science* 15: 343-399.

Norvig, P. 1989. Marker Passing as a Weak Method for Text Inferencing. *Cognitive Science* 13: 569-620.

Riesbeck, C. K., and Schank, R. C. 1989. *Inside Case-Based Reasoning*. Hillsdale, NJ: Lawrence Erlbaum.

Riesbeck, C. and Martin, C. 1986. Direct Memory Access Parsing. In Kolodner and Riesbeck eds. *Experience, Memory, and Reasoning*, 209-226. Lawrence Erlbaum.

Schank, R. C. 1982. *Dynamic Memory*. New York: Cambridge University Press.

St. John, M. in press. The Story Gestalt: A Model of Knowledge Intensive Processes in Text Comprehension. *Cognitive Science*.

Thagard, P., Holyoak, K. J., Nelson, G., and Gochfeld, D. 1990. Analog Retrieval by Constraint Satisfaction. *Artificial Intelligence* 46:259-310.

Touretzky, D. and Hinton, G. 1988. A Distributed Connectionist Production System.*Cognitive Science* 12:423-466

Waltz, D., and Pollack, J. 1985. Massively Parallel Parsing: A Strongly Interactive Model of Natural Language Interpretation. *Cognitive Science* 9: 51-74.

Wharton, C., Holyoak, K., Downing, P., Lange, T., and Wickens, T. 1992. The Story with Reminding: Memory Retrieval is Influenced by Analogical Similarity. In Proceedings of the 14th Annual Conference of the Cognitive Science Society. Hillsdale, NJ: Lawrence Erlbaum.