A Parallel Distributed Processing Model of Story Comprehension and Recall

Richard M. Golden
University of Texas at Dallas

and

David E. Rumelhart
Stanford University

Richard M. Golden
University of Texas at Dallas, School of Human Development
GR41, Box 830688, Richardson, Texas, 75083–0688
Phone: (214) 690–2423, Email: golden@utdallas.edu

Running head: Model of Story Comprehension and Recall

Abstract

An optimal control theory of story comprehension and recall is proposed within the framework of a "situation" state space. A point in situation state space is specified by a collection of propositions each of which can have the values of either "present" or "absent". Story comprehension is viewed as finding a temporally-ordered sequence of situations or "trajectory" consistent with story-imposed constraints. Story recall is viewed as finding a trajectory consistent with episodic memory constraints. A multi-state probabilistic (MSP) causal chain representation is then introduced for compactly representing the reader's belief in the likelihood of different trajectories in situation state space. A parallel distributed processing model is also introduced that searches for highly probable trajectories with respect to a set of constraints and an MSP causal chain representation. The recall performance of the model is then qualitatively compared to the recall performance of human subjects.

A mental model is a mental blueprint or abstraction of certain aspects of the physical world (Johnson-Laird, 1983). Many theorists (Rumelhart, 1977a; Schank and Abelson, 1977; Bower and Morrow, 1990; Bransford, Barclay, and Franks, 1972) have suggested that understanding a story involves a process of finding or constructing a mental model consistent with the constraints of the story.

Once an appropriate mental model has been *constructed*, this representation may be encoded with the constraints of the story into the reader's memory and *reconstructed* at some later time period. The types of recall errors made by subjects can then be used to provide insights into how the reader's world knowledge interacts with the constraints of the story during the constructive encoding and reconstructive retrieval processes. Under certain conditions, recall errors due to both constructive memory processes (e.g., Bransford and Johnson, 1972) and reconstructive memory processes (e.g., Spiro, 1977, 1980) may be observed.

Trabasso and his colleagues (Trabasso, Van Den Broek, and Suh, 1989) have become especially interested in causal knowledge structures of the type described by Mackie (1980). According to this viewpoint, the reader creates a *causal field* which is updated as each statement in the story is read. The causal field specifies the *circumstances* in which the story takes place, and so explicitly identifies states and conditions which may be only implicitly stated in the story. Causal relationships between pairs of events in the causal

field are identified using a *counterfactual* argument. In particular, event A causes event B means that if event A had not occurred in the circumstances, then event B would not have occurred. It also necessary that event A occurs before event B occurs.

In a series of studies (see for example, Trabasso, Secco, and van den Broek, 1984; Warren, Nicholas, & Trabasso, 1979), Trabasso and his colleagues analyzed some simple stories by (i) parsing the story into clauses roughly corresponding to events in the story, (ii) intuitively identifying critical causal relationships between clauses, and (iii) checking that these relationships satisfy the above counterfactual criteria. Once the causal relationships among pairs of events in the story were identified, they could be assembled into a causal chain representation of the story. Some experimental evidence supports the validity of constructing chains in this manner (Trabasso, Van Den Broek, & Suh, 1989). Given the causal chain representation of a story, predictions regarding which propositions in the story will be recalled from memory can be made. In particular, propositions on the main causal chain of the story are more likely to be recalled (Trabasso and Van Den Broek, 1985; Trabasso, Secco, and Van Den Broek, 1984) and rated as important (Trabasso and Sperry, 1985; Trabasso and Van Den Broek, 1985), and propositions with more causal connections are more likely to be recalled (Trabasso and Van Den Broek, 1985; Trabasso, Secco, and van den Broek, 1984). Additional experiments and analyses which have attempted to control for effects of shared referents, serial position, and imagery (Trabasso and Van Den Broek, 1985), and hierarchical structure (Van Den Broek and Trabasso, 1986; Van Den Broek, 1988) have supported the hypothesis that the causal chain representation is a good (but not necessarily sufficient) predictor of recall performance.

The approach to story comprehension and recall described here is similar to the general approach to story comprehension and recall pursued by Trabasso, Van Den Broek, and their colleagues, but differs substantially in its technical details. In particular, the fundamental building block of the Trabasso et al. causal chain representation is the event clause (i.e., a sentence fragment), while in our approach the event clause is a proposition. Second, the Trabasso et al. causal chain representation is not really self-contained in the sense that additional English text is often required to explain why certain links are attached, we have constructed a formal notation that attempts to make the causal chain representation more self-contained. Third, the machinery used to represent the changes to the causal field is essentially a *finite-state machine* (Johnson-Laird, 1983) which permits only one change at a time to the causal field. We introduce a more general graphical notation for representing the time evolution of the causal field that compactly represents and allows multiple state changes to the causal field to simultaneously occur at each instant of time. Fourth, we propose an explicit model of story comprehension and recall which, although similar in its general characteristics to the models proposed by Trabasso and his colleagues, is explicit enough to simulate on a computer and is also explicit enough to

suggest a new performance measure of the "causal coherence" of a temporally-ordered group of propositions. And fifth, we simulate the model and qualitatively compare the performance of the model to the performance of human subjects.

Finally, although the structural form of a good story is almost limitless, we restrict ourselves in this paper to stories describing a sequence of events (usually causally related) that follow one another in an approximately linear temporal sequence. We do not believe, however, that many of the ideas suggested here are only applicable to this restricted class of stories. Rather, we hope a better understanding of simple stories possessing a strong causal structure will eventually allow us to extend our theories to more general cases.

An Optimal Control Theory of Story Comprehension and Recall

World Knowledge is a Set of Highly Probable Trajectories

We have approached the problem of story comprehension and recall within the framework of a special high dimensional state space called "situation" state space (Golden, 1986c; Rumelhart, 1984). A point in situation state space consists of a collection of d facts or features about the world each of which can be classified as being either "present" or "absent." The reader's current mental state is therefore modelled as a single point in a d-dimensional situation state space. For example, consider the following "point"(four-dimensional vector) in a four-dimensional situation state space. Note that "1" specifies the presence of a feature, while "0" specifies the absence of a feature.

Insert Table 1 about here

At some later point in time, the reader's mental state would be modelled as another point in the same d-dimensional situation state space. It will be convenient, therefore, to order these points according to a time index. Thus, the "unfolding" or "evolution" of the reader's mental state as a function of time, may be represented as an ordered sequence of points or *trajectory* in situation state space.

Some trajectories in situation state space will be more likely to occur than others. For example, using the "Mary" example, a highly probable trajectory (i.e., a trajectory that typically will be assigned a high degree of belief by a reader's belief function) in situation state space might be specified by the following sequence of three 4–dimensional situation state vectors.

Insert Table 2 about here

World knowledge is therefore represented by a belief function which assigns a specific "degree of belief" to different trajectories in the situation state space. In particular, trajectories which are more likely are assigned values close to one, while trajectories which are less likely are assigned values close to zero. The specific form of belief function used in this paper is a *subjective probability function* (e.g., Cox, 1979) which

assigns a probability (i.e., a degree of belief) to all possible trajectories in situation state space. Note that a trajectory in situation state space is similar to the causal field discussed by Trabasso and his colleagues. Finally, rather than storing all possible trajectories and their probabilities in a look-up table, readers might possess this knowledge implicitly by "calculating" the likelihood of a given trajectory.

Stories are Constraints upon Trajectories

We will now consider the representation of stories within the theoretical framework we have been sketching. We view a story as a set of "hints" or constraints about some sequence of events in the world which the author of the story would like to share with the reader of the story. More formally, a story is a a trajectory through situation state space which contains many feature values which are not observable and need to be estimated.

Comprehension and Recall Processes

What does it mean to understand a story? One popular hypothesis (Schank and Abelson, 1977; Rumelhart, 1977b; Johnson and Laird, 1983) is that understanding a story involves the construction of some underlying representation of the story which is consistent with both the reader's and the author's knowledge of the world. The resulting representation can then be learned, and used to answer questions about the story, summarize the story, or recall the story from memory at some later time period. The particular theory of story comprehension proposed in this paper may be viewed as a special case of this theory. In particular, we propose the reader understands a story by searching his or her knowledge base for a sequence of situations which are likely to be consistent with the constraints imposed by the story.

Actually, our viewpoint allows us to exploit a well developed branch of engineering and mathematics known as optimal control theory (see Luenberger, 1979, for an introduction) to partially address some of the modelling problems posed within this theoretical framework. Consider the problem of flying an airplane. Because of the physical constraints imposed by the plane, the plane's environment, and the pilot, the plane can only fly along certain trajectories through the air. For example, it is physically impossible for an airplane to fly the same types of trajectories as a helicopter. We can think of the location of the aircraft in three-dimensional space as analogous to a point in situation state space, and the probability distribution of trajectories of the airplane as corresponding to the belief structure of the reader which is a probability distribution over trajectories in situation state space.

We might now ask the airplane's pilot to fly through certain "windows" in space and time. For example, we might ask the pilot to take-off in Pittsburgh in the morning, pass over Dallas, and land in San Francisco in the early evening. These space-time windows actually define a set of constraints which the pilot must achieve. But given these constraints are achieved, the pilot has considerable freedom in choosing a flight

pattern consistent with the physical constraints imposed by the world (i.e., the probability distribution of trajectories of the airplane). In a similar manner, the reader of the story must plot a course (i.e., a trajectory) through situation state space. The reader has considerable freedom in choosing a flight trajectory but this trajectory must satisfy the constraints of the story *and* be "highly probable" with respect to the reader's belief structure.

More formally, the problem of story comprehension is defined as finding a *most probable* trajectory in situation state space with respect to some belief function or probability distribution (see Rumelhart (1977a, 1977b, 1984) for closely related approaches). Learning in the model corresponds to using the constructed trajectory or mental model to update the reader's world knowledge (i.e., the parameters of the reader's belief function). Story recall is viewed as another process which attempts to retrieve the constructed trajectory from memory at some later time period using an appropriate retrieval cue, and then samples the reconstructed trajectory to generate a recalled sequence of events.

Related Models of Text Comprehension and Recall

The basic features of this approach are shared by many other models of text comprehension and recall in both the psychological (Graesser and Clark, 1985; Kintsch and Van Dijk, 1978) and artificial intelligence literature (see Schank and Riesbeck, 1981, for a review; also see Norvig, 1989, and Dyer, 1983, for more recent approaches). Our approach to the representational problems was influenced greatly by Schank and

Abelson (1977) as well as by the work of Trabasso and his colleagues (e.g., Trabasso, Secco, and Van Den Broek, 1984; Warren, Nicholas, and Trabasso, 1979). Our basic philosophy of formally viewing the comprehension process as a search for the most probable interpretation of a story is best stated in Rumelhart's (1977b) interactive reading model. It is also worth noting that this *two-stage model* of memory recall has important similarities with the SAM model of human memory (for a review of the SAM model see Raaijmakers, 1987).

Our approach is different from many other approaches in the psychological and artificial intelligence literature, in that multiple "states" in our model can be simultaneously active and the representation of world knowledge is a probability distribution over classes of permissable trajectories through situation state space. Thus, the foundations of our approach also have important connections with research in situation semantics (e.g., Barwise and Perry, 1983), hierarchical multi-state machines (e.g., Gabrielian and Iyer, 1990), and nonlinear optimal control theory (e.g., Luenberger, 1977).

Integration with Syntactic Subsystems

There are two major classes of approaches for linking a system of causal inference with syntactic processing systems. The first approach involves transforming a sentence clause into a proposition representation, and then constructing situations as sets of propositions as we have done in this paper. This approach is also characteristic of traditional artificial

intelligence approaches to text comprehension (e.g., Schank and Abelson, 1977) where the original text is transformed first into a propositional representation, and then inferences are are made with respect to the transformed text.

An alternative approach, following Rumelhart (1977a), is to take a much broader view where both syntactic and semantic constraints can freely interact to construct a highly probable interpretation of a story. Within this framework, a situation would consist of syntactic as well as semantic features. Relations or equivalently "constraints" defined over regularities in time-ordered sequences of situations could then be exploited to make both syntactic and semantic inferences. Fenstad, Halvorsen, Langholm, and Benthem (1987) provide an interesting example of how syntactic and semantic constraints might freely interact within a situation semantics framework.

Acquisition of World Knowledge

The acquisition of world knowledge is not explicitly addressed in this model, but one theory consistent with the model is that causal connections between events are constructed from spatio-temporal regularities in the world. Some recurrent connectionist network models acquire knowledge in just this way (Allen, 1990; McClelland, St. John, and Taraban, 1989; Miikkulainen, 1990; Miikkulainen and Dyer, 1989; St. John and McClelland, 1990). In addition, Nelson and her colleagues (Nelson, Fivush, Hudson, and Lucariello, 1983) have suggested that representations of event knowledge in children initially are highly structured, and then gradually become more flexible with experience which is consistent with the “regularity extraction” hypothesis.

A Model of Story Comprehension and Recall

In this section, we describe a notation for representing complex classes of trajectories in a high-dimensional situation state space. Second, we propose a psychological process model of how the story comprehension and recall processes might actually operate using causal chain representations. Finally, the performance of the model will be evaluated with respect to the relevant psychological data.

Multi-state Probabilistic Machines for Causal Knowledge Representation

Gabrielian and his colleagues (for a review see Gabrielian and Iyer, 1990) have been developing the theoretical foundations of a new class of machines known as *hierarchical multi-state machines* for specifying the performance requirements of complex high-dimensional systems. We have found that a probabilistic extension of the multi-state machines studied by Gabrielian and his colleagues provides a useful way of implicitly, yet precisely, representing complex belief functions that assign degrees of belief to specific trajectories in situation state space. We will call this special notation for representing complex probabilistic knowledge structures a MSP (multi-state probabilistic) machine.

A MSP machine consists of three distinct types of entities: local states, transitions, and controls. The *local states* of the MSP machine correspond to the set of d features

required to specify a point in a d-dimensional situation state space. The set of all states of the MSP machine is sometimes referred to as a *global state* of the machine. A *marking* of the MSP machine indicates the values of local states of the machine at an instant in time. Thus, a marking of the machine identifies a point in the d-dimensional situation state space. The *transitions* and the *controls* of the MSP machine specify how the global state of the machine evolves from one instant of time to the next.

For example, consider a MSP machine consisting of two local states: event A and event B. The state A might refer to the event “Mary heard the ice-cream truck”, while the state B might refer to the event “Mary wanted to buy some ice-cream.” Figure 1 introduces a graphical notation which formally states that the probability (belief) that event B occurs at the current instant of time is functionally dependent upon the value of the event A at the previous instant of time. In particular, Figure 1 depicts a two-dimensional global state of the machine (i.e., a point in situation state space) where only the A state is present (marked) at some time t. With this marking of the machine, the transition entering state B will “fire”, and the B state will become marked with a high probability at the next instant of time. Notice that the firing of the transition does not affect the value of the A state since A is connected to B via a control (thin arrow). Another way to visualize the flow of information in Figure 1 is to imagine a directed excitatory link from state A at time t to state B at time t+1.

Insert Figure 1 about here

Figure 2 shows another two-state machine where transitions are *leaving*, as well as *entering* local states. To fix ideas, let C refer to the event that "Mary is eating ice-cream."In this example, the probabilistic transition (dark arrow) connecting states B and C allows the marker in state B to "pass through" to state C with a high probability. Or more formally, the probability that C will be marked at the next instant of time is functionally dependent upon the value of state B at the current instant of time, *and* the probability that B will be unmarked at the next instant of time is functionally dependent upon the value of state C at the current instant of time. Another way to visualize the flow of information in Figure 2 is to imagine a directed excitatory link from state B at time t to state C at time t+1, and an inhibitory link from state C at time t to state B at time t+1.

Insert Figure 2 about here

Appendix A describes more complicated examples where the multi-state probabilistic machine notation is used to represent the causal knowledge structure underlying four specific stories. These stories were parsed into "event clause" propositions following Trabasso, Secco, and Van Den Broek (1984). All probabilistic transitions are designed to "fire" with probability 0.9. Like the notation of Trabasso and his colleagues, the links in

this representation are derived from a combination of intuitive considerations supported by counterfactual arguments. This notation has the advantage, however, of permitting multiple local states to be simultaneously active, and permitting multiple local transitions to simultaneously fire. We also exploit a Markov chain framework to formally link the firing of the "local probabilistic" transitions, with the global subjective probability function which assigns a probability to each trajectory in the system (Appendix B). Thus, once the probabilistic causal chain representation has been constructed, it is possible to implicitly assign a degree of belief to all possible trajectories in situation state space through the use of interpretable local probability distributions.

The Parallel Distributed Processing Model

The theory of story comprehension and recall we have been describing is based, as we have previously noted, on a two-step process. First, a highly probable (believable) trajectory is computed by the reader which is consistent with the constraints of the story. This story trajectory is the reader's mental model of the story. The reader learns the story trajectory so that given the initial portion of the trajectory (story) as a retrieval cue, the remainder of the story is retrieved.

We began by constructing MSP representations for each of the four stories analyzed by Trabasso et al. (1984) (see Appendix 1). These four stories were titled *Epaminondas, Judy's Birthday, The Tiger's Whisker*, and *The Fox and the Bear*. These stories are for the most part temporally ordered, have only simple syntactic structures, possess a strong causal structure, and vary in their degree of causal coherence.

Network architecture. We have devised a recurrent parallel distributed processing model which searches for highly probable trajectories where the probability of a trajectory is formally defined with respect to a MSP representation consisting of d local states or propositions.. Figure 3 shows a convenient way of thinking about this system. Instead of trying to visualize the complex temporal dynamics of a d-unit system, a system of d*M units is considered where the unit in the i*th* row and j*th* column of the array corresponds to the value (absent or present) of the i*th* feature at time j. Thus, the each column of the array represents a point in the d-dimensional situation state space, and the M columns of the array correspond to an ordered sequence of M points in situation state space. Notice that the connectivity pattern between units at adjacent time intervals is identical because this two-dimensional array is intended to model a two-layer recurrent network. The network shares important similarities with recurrent back-propagation (Jordan, in press; Rumelhart, Hinton, and Williams, 1986), the elastic net algorithm (Durbin and Willshaw, 1987), and auto-associative networks such as the brain-state-in-a-box model (Anderson et al., 1977; Golden, 1986a) and Hopfield-type networks (Hopfield, 1982; Hopfield, 1984).

The weights of the system (i.e., the parameters b_{ij} and a_i) are set such that:

1. If all excitatory inputs to a unit are inactive, the local probability that the unit would

be active would be 0.1 or less.

2. If any excitatory input to a unit is active and all inhibitory inputs are inactive, the local probability that the unit would be active would be 0.9 or greater.

3. If all inhibitory inputs to a unit are active and some subset of excitatory inputs are active, the local probability that the unit would be active would be 0.1 or less.

Comprehension process. The comprehension process is modelled by beginning with a story which has already been parsed into a situation state space trajectory. If situation feature i at time j is equal to one (indicating feature i at time j is present), then the activation value of the unit in the i*th* row and j*th* column of the array in Figure 3 is clamped to the value of one during the comprehension process. The term *clamped* refers to the case where the activation of a unit is not permitted to change. The unclamped unit representing situation feature i at time j then computes a weighted sum of the activations of the units at time j-1 and time j+1, and uses this weighted sum to increment or decrement the unclamped unit's activation value. The unclamped unit's activation value also decays toward zero (i.e., the value is decremented) by an amount inversely proportional to ρ so that the system searches the high-dimensional trajectory space in a region near the original story constraint trajectory. Finally, the activations of all units in the system are constrained to lie in the range of zero to one. When the activations of the unclamped units are updated in this manner, then the network may be formally viewed as seeking a trajectory in situation state space which is highly probable with respect to the system's belief (probability) function (Appendix B). The system updates the activations of the units for some pre-determined maximum number of iterations, or until the system stabilizes. The resulting pattern of activation over the d * M units is the constructed story trajectory, and represents the system's "understanding" of the story.

Insert Figure 3 about here

Learning and recall processes. It is assumed that people learn the story trajectory (mental model) which they constructed during the comprehension process. Rather than attempt to model the details of this learning process, we are currently content to model the end results of that process. In particular, it is assumed that when people are asked to recall a story from memory, an episodic memory trace of the story trajectory is available to guide the recall process. In the model, the strength of this episodic memory trace is the retention interval parameter ρ. Small values of ρ correspond to short retention intervals where the strength of the episodic memory trace of the story trajectory is strong. Large values of ρ correspond to long retention intervals where the strength of the episodic memory trace of the story trajectory is weak.

The recall process is then similar to the comprehension process, except that only the first few features of the story in the first three time steps of the trajectory are clamped

as a retrieval cue. The network must then reconstruct the remainder of the trajectory. Although this is a very difficult parameter estimation problem, the introduction of the constraints imposed by the retention interval parameter help since they restrict the search to trajectories close in an Euclidean distance sense to the original story trajectory. In particular, each unit at time t uses a weighted sum of the activations of the units at times t-1 and times t+1 to update its activation value, but each unit at time t also has its activation value modified by the episodic memory trace of the story as well. In particular, if the i*th* situation feature in situation t in the story trajectory is active, then the i*th* unit at time t has its activation value incremented by an amount inversely proportional to ρ. If the i*th* situation feature in situation t in the story trajectory is not active, then the i*th* unit at time t has its activation value decremented by an amount inversely proportional to ρ.

The reconstructed recall trajectory is then used to generate the model's free recall responses. In particular, the "most active" unit in each situation is assumed to be the model's summary of that situation. Moreover, since the situation vectors are ordered in the recall trajectory, the responses of the model will be ordered as well. If the maximally active unit's value was less than 0.5, the model does not recall a situation feature summarizing that situation. If several units in a situation are maximally active, then exactly one of the units is randomly chosen to summarize the situation.

Note that both of the comprehension and recall search processes in the model involve backward and forward chaining as the recurrent network cycles forwards and backwards in time to construct a trajectory (see Appendix B for additional details). Trabasso et al. (1984) and Black and Bern (1981) have reported evidence supporting the hypothesis the probability of recalling a story consequence given the story cause is roughly the same as the probability of recalling a story cause given the story consequence. Trabasso et al. (1984) interpreted these findings as supporting a retrieval model which exploits bi-directional search strategies.

Simulation Experiments

Human memory for simple causally coherent stories is characterized by at least four fundamental phenomena (Trabasso et al., 1984). First, propositions with more causal connections are more likely to be on the main causal chain of the story and are more likely to be recalled. Second, as the retention interval between reading and recalling the story is increased, the percentage of propositions recalled decreases but those propositions will be more likely to lie on the main causal chain of the story. And third, more propositions will be recalled from stories which are more causally coherent (i.e., stories which possess more propositions on the main causal chain of the story). In this first set of simulations, we were interested in whether the proposed model would exhibit some of these qualitative phenomena. Two different performance measures were used to evaluate the model's performance.

The first performance measure was average causal connectivity. The average *causal connectivity* of a group of propositions was computed by counting the number of forward and backward connections in the MSP representation each statement in the group possessed, and then averaging over the group of propositions.

The self-information of a temporally ordered group of propositions was also computed as a second performance measure. The second performance measure was to construct a trajectory of story constraints for a temporally ordered group of propositions. Note that a trajectory which is constructed in the manner is an ordered sequence of situations such that each situation in the sequence has exactly one "active" feature. The relative self-information of a trajectory is defined as the negative logarithm of the model's degree of belief (probability) of the trajectory plus any additive constant (Appendix B). Thus, since "less likely" trajectories are less causally cohesive, "less likely" trajectories are more informative (i.e., surprising) in an information theoretic sense and possess larger relative self-information values.

Each of the four stories were individually "comprehended" by the model, "learned" by the model, and "recalled" by the model as previously described. The ρ parameter during the recall process was varied, and took on the values:$\rho = 0.1$, $\rho = 0.2$, and $\rho = 0.3$ corresponding to short, medium, and long retention intervals respectively. A flow chart which summarizes the basic simulation program is provided in Figure 4.

Put Figure 4 about here

Construction and Evaluation of the MSP Causal Chain Representations

Because the finite-state machine causal chain representations of Trabasso et al. (1984) were so successful at predicting which propositions would be recalled from memory by human subjects, we were curious whether the MSP representation preserved roughly the same causal connectivity characteristics of the Trabasso et al. representation. To test this hypothesis, correlation coefficients for each of the four stories were computed. Each correlation coefficient compared the number of connections a story clause possessed with respect to the Trabasso et al. representation to the number of connections a story clause possessed with respect to the MSP representation. Thus, the number of event clauses in a story corresponded to the number of observations required to estimate a particular correlation coefficient. Table 3 shows the results of this preliminary analysis of "representations."

Insert Table 3 about here

Table 3 suggests that the MSP machine representations preserve some of the important functional properties of the causal chain representation used by Trabasso et al. (1984). In

particular, nodes with more causal connections in the Trabasso representation also have more causal connections in the MSP representations described here.

Fewer Propositions are Recalled as the Retention Interval is Increased

Figure 5 shows how variation of the retention interval parameter of the model ρ affects the number of propositions recalled by the model from memory. As retention interval is increased, the percentage of propositions in the story which are recalled decreases. A similar pattern of results were obtained by Trabasso et al. (1984) and Trabasso and Van Den Broek (1985) for data obtained from human subjects.

The reason for this effect in the model is simply that the retention interval parameter of the model directly affects the strength of the episodic memory trace of the story. Thus, event clauses without causal connections will be forgotten for large values of the retention interval parameter (i.e., long retention intervals). Similarly, Trabasso et al. (1984) and Trabasso and Van Den Broek (1985) found that memory loss over time for statements which were not on the main causal chain of the story was greater than memory loss over time for statements which were on the main causal chain.

Put Figure 5 about here

An analysis of variance of the data using a story by retention interval design where the story factor was treated as random, however, did not support the hypothesis that as retention interval is increased in the model, the percentage of recalled statements decreases although a borderline effect was observed ($F(2,6) = 4.6$, $p < 0.10$).

Self-information is a Good Predictor of the Percentage of Recalled Propositions

One potential application of the theory described here is a measure of the "readability" or "causal coherence" of a text. Trabasso, Secco, and Van Den Broek (1984) in their analysis of childrens' recall data, note that the percentage of propositions recalled by people is a linear function of the percentage of propositions on the main causal chain of a story which is subjectively identified.

The causal coherence of a story is naturally computed within this theoretical framework, by expressing the story as a trajectory through situation state space (note that each situation will only have one active proposition in this case) and computing the likelihood (probability) of that trajectory (see Appendix B).

Table 4 shows how the percentage of propositions recalled from each of the four stories by the model is correlated with the above two measures of causal cohesiveness. For comparison purposes, the recall data collected by Trabasso et al. (1984) was also used in the analysis. Note that the recall performance of the model correlated highly with the recall performance of human subjects. Also note that as the information content of the story decreases (i.e., as the probability of the story trajectory increases), the percentage of recalled statements increases although this latter correlation was significant on the

borderline. Moreover, the causal connectivity measure does not seem to be a good predictor of recall performance since greater causal connectivity suggests improved recall performance yet causal connectivity is negatively correlated with recall performance.

Insert Table 4 about here

Table 4 suggests that the self-information of a story might be a good predictor of both the human and model recall data, while the causal connectivity measure accounts for a lower percentage of the variance. Figure 6 shows the relationship between recall performance and trajectory self-information is qualitatively quite similar. This is a promising observation since the self-information content of a temporally ordered group of propositions may be objectively computed using only the MSP machine representation and the story constraint trajectory representation of the story.

Insert Figure 6 about here

Notice that the "Epaminondas" story which is the most causally coherent as determined by the self-information measure is also recalled most effectively by the model and by people. Also notice that the "Fox and Bear" story which is least causally coherent is recalled least effectively by the model and people.

Groups of Propositions with Less Self-information are Recalled as the Retention Interval is Increased

The effects of varying the retention interval parameter of the model upon the set of propositions recalled by the model were also considered. The recalled propositions were used to construct a trajectory in the state space, and then the self-information of that trajectory was computed. The self-information of the original story trajectory was then subtracted from the self-information of the recalled trajectory to arrive at the *relative self-information*. Both of these quantities were computed with the retention interval factor set equal to zero. The relative self-information is small if the recalled propositions have much less information content than the original story. The relative self-information can equivalently be viewed as a ratio of the likelihood of the original story trajectory divided by the likelihood of the recalled story trajectory. Simulations of the model showed that propositions with less relative self—information (greater causal cohesiveness) are more likely to be recalled as retention interval increases (Figure 7). An analysis of variance of the data using a story by retention interval design where the story factor was treated as random supported the hypothesis that this effect was significant ($F(2,6) = 15.6$, $p < 0.01$). Moreover, the average relative self-information was significantly less than zero ($t(11) = -4.8$, $p < 0.01$), indicating that the causal coherence of the recalled statements was greater than the causal coherence of the original story.

This behavior is consistent with the fact that the model seeks trajectories which are more likely or probable with respect to a probability measure which is dependent upon both the causal knowledge structures required for understanding the story, and the temporal episodic memory trace of the story. The relative self-information measure, on the other hand, is only dependent upon the causal structure of the story. These two observations account for the general behavior of the model which demonstrates a tendency to recall trajectories which are more consistent with its causal knowledge as the retention interval factor is decreased. Notice that the Epaminondas story which is very causally coherent does not show this effect since the belief (probability) of any subset of recalled propositions tended to be less than the belief (probability) of the original story.

Insert Figure 7 about here

Propositions with Greater Causal Connectivity are More Likely to be Recalled

Figure 8 shows how variation of the retention interval parameter of the model affects which propositions are recalled by the model from memory. The dependent measure was *relative causal connectivity* which was defined as the average number of causal connections per proposition in a story subtracted from the average number of causal connections per proposition in the set of propositions recalled by the subjects. Like human subjects (e.g., Trabasso et al., 1984; Trabasso and Van Den Broek, 1986), propositions with greater causal connectivity (greater causal cohesiveness) are more likely to be recalled as retention interval increases because propositions with fewer causal connections will be less likely to be recalled.

Insert Figure 8 about here

An analysis of variance of the data using a story by retention interval design where the story factor was treated as random, however, did not support the hypothesis that statements with greater causal connectivity are more likely to be recalled as retention interval increases although the effect was borderline significant ($F(2,6) = 4.2$, $p < 0.10$). On the other hand, the average relative causal connectivity was significantly greater than zero ($t(11) = 2.4$, $p < 0.05$), indicating statements with greater causal connectivity are more likely to be recalled (e.g., Trabasso et al., 1984; Trabasso and Van Den Broek, 1986) was replicated by the model.

The reason for the model's behavior is simply due to the fact that the number of causal connections a proposition has with other propositions in a story may be considered to be an indirect measure of the magnitude of the self-information of a trajectory. The model claims that the self-information measure is a more sensitive indicator of performance than the causal connectivity measure.

Misordered Propositions Drift More Towards their "Canonical" Positions in the Story as Retention Interval is Increased

Another important aspect of human memory for stories with a strong causal structure is that if propositions in such stories are "displaced," they will "migrate" back towards their canonical positions in the text (Bischofshausen, 1985; Bower, Black, and Turner, 1979; Mandler, 1978). Moreover, the magnitude of these effects tend to increase as retention interval is increased (Bischofshausen, 1985). We would expect to see these effects in the model as well since as the retention interval factor is decreased, the local probability of predicting a feature at the next instant of time becomes less dependent upon the episodic memory trace of the story formed during the comprehension process.

To model these phenomena, the simulations described above were repeated using exactly the same four causal knowledge structures (i.e., sets of connections among the units), but using stories which were slight distortions of the original four stories. For example, a "distorted" story was generated from the original *Fox and Bear* story by switching statement #8 with statement #21. Three distorted stories for each of the original four stories were generated in this manner, and the recall of the model for each of the twelve stories was recorded. Because the retention interval factor was varied in this experiment as well, thirty six independent simulation runs of the model were done. The number of steps along the recalled trajectory a misordered statement "drifted" towards its canonical position was used as a dependent measure.The results of these experiments are shown in Figure 9.

Insert Figure 9 about here

An analysis of variance of the data using a story by retention interval design where the story factor was treated as random supported the hypothesis that as retention interval increased, the magnitude of the drift increased as well ($F(2,6) = 15.6$, $p < 0.01$). Also the drift measure was significantly greater than zero ($t(11) = 6.1$, $p < 0.01$) indicating that misordered statements did indeed “drift” towards their canonical positions in the text rather than in the opposite direction.

Summary and Conclusions

This paper has provided a framework for representing knowledge in some forms of text. The resulting knowledge representation are used to assign a "degree of belief" or probability to any ordering of any subset of propositions derived from the original text. This probability is a good predictor of the "causal coherence" of a story as measured by the percentage of propositions recalled from a story.

A connectionist model which can implicitly use this probabilistic knowledge representation to chain forward and backward in time to make necessary causal links was also described. Thus, the model provides some insights into the types of *psychological mechanisms* people may use to “link up” the causal structure of a story. The model is

also noted to be formally searching for the most probable value of the “unknown” features of the trajectory given the “known” features (Appendix B), thus we have attempted to begin to formalize the nature of the computational problem of story understanding within a probabilistic framework.

The performance of the model was then qualitatively compared to the performance of human subjects. The model had difficulty recalling the same stories which human subjects had difficulty recalling. In addition, like human subjects, propositions with more causal connections were more likely to be recalled, and misordered propositions tended to be recalled closer to their canonical positions in the text. These properties of the model are due to the fact the model constructs trajectories which are (i) consistent with the constraints of the episodic memory trace of the story, and (ii) are causally coherent with respect to the self—information measure.

In addition, like human subjects, the magnitude of the causal connectivity and misordered proposition effects increase as retention interval is increased. This property of the model is simply due to the fact that the strength of the episodic memory trace contribution is weighted by the retention interval parameter ρ. When ρ is small, the episodic memory trace contribution dominates and the model is biased to construct trajectories which are more consistent with the story’s episodic memory trace, and less consistent with the model’s causal knowledge structure. When ρ is large, the causal knowledge contribution dominates and the model is biased to construct trajectories which are less consistent with the story's episodic memory trace, and more consistent with the model's causal knowledge structure.

Finally, preliminary simulation data support the hypothesis that the relative self-information performance measure may be more effective than the causal connectivity measures typically used in the experimental literature. Two obvious advantages of the relative self-information performance measure are: (i) the measure is a natural one with respect to the probabilistic framework proposed here, and (ii) the self-information performance measure takes the *sequential order* of the recalled propositions directly into account.

One major problem with the current instantiation of our more general theory, however, is its computational adequacy. In particular, how does one determine or estimate a set of probabilistic rules (i.e., probabilistic temporal logic expressions) for the MSP machine that represent an "adequate" understanding of a story? Another issue of interest is the functional utility of this theoretical framework for developing theories of story comprehension and recall that can account for the empirical data. We are currently pursuing both experimental and theoretical avenues in an attempt to address these and related questions.

References

Ackley, D. H., Hinton, G. E., & Sejnowski, T. J. (1985). A learning algorithm for boltzmann machines. *Cognitive Science*, *9*, 147–169.

Allen, R. B. (1990). Connectionist language users. *Connection Science*, *2*, 279–311.

Anderson, J. A., Silverstein, J. W., Ritz, S. A., & Jones, R. S. (1977). Distinctive features, categorical perception, and probability learning: Some applications of a neural model. *Psychological Review*, *84*, 413–451.

Barwise, J. & Perry, J. (1983). *Situations and Attitudes*. Cambridge: MIT Press.

Bischofshausen, S. (1985). Developmental differences in schema dependency for temporally ordered story events. *Journal of Psycholinguistic Research*, *14*, 543–556.

Black, J. B. & Bern, H. (1981). Causal coherence and memory for events in narratives. *Journal of Verbal Learning and Verbal Behavior*, *20*, 267–275.

Bower, G. H., Black, J. B., & Turner, T. J. (1979). Scripts in memory for text. *Cognitive Psychology*, *11*, 177–220.

Bower, G. H. & Morrow, D. G. (1990). Mental models in narrative comprehension. *Science*, *247*, 44–48.

Bransford, J. D., Barclay, J. R., & Franks, J. J. (1972). Sentence memory: A constructive versus interpretive approach. *Cognitive Psychology*, *3*, 193–209.

Bransford, J. D. & Johnson, M. K. (1972). Contextual prerequisites for understanding: Some investigations of comprehension and recall. *Journal of verbal learning and verbal behavior*, *11*, 717–726.

Cox, R. T. (1979). Of inference and inquiry: An essay in inductive logic. In Levine & Tribus (Eds.), *The maximum entropy formalism* (pp. 119–167). Cambridge: MIT Press.

Durbin, R. & Willshaw, D. (1987). An analogue approach to the travelling salesman problem using an elastic net method. *Nature*, *326*, 689–691.

Dyer, M. (1986). *In-depth understanding*. Cambridge: MIT Press.

Fenstad, Halvorsen, Langholm, & Benthem. (1987). *Situations, language, and logic*. Holland: D. Reidel.

Gabrielian, A. & Iyer, R. (to appear). Integrating automata and temporal logic: A framework for specification of real-time systems and software. In C. Rattray & R. Clark (Eds.), *The Unified Computation Laboratory, Institute of Mathematics and its Applications*. Stirling, Scotland: Oxford University Press.

Golden, R. M. (1986a). The brain-state-in-a-box neural model is a gradient descent algorithm. *Journal of Mathematical Psychology*, *30*, 73–80.

Golden, R. M. (1986b). A developmental neural model of visual word perception. *Cognitive Science*, *10*, 241–276.

Golden, R. M. (1986c). Representing causal schemata in connectionist systems. In *Proceedings of the Eighth Annual Conference of the Cognitive Science Society* (pp. 13–22). Hillsdale, NJ: Erlbaum.

Graesser, A. C. & Clark, L. F. (1985). *Structures and procedures of implicit knowledge*. Norwood, NJ: Ablex.

Hopfield, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences, USA*, *79*, 2554–2558.

Hopfield, J. J. (1984). Neurons with graded response have collective computational properties like those of two-state neurons. *Proceedings of the National Academy of Sciences, USA*, *81*, 3088–3092.

Johnson-Laird, P. N. (1983). *Mental models*. Cambridge, MA: Harvard.

Jordan, M. I. (in press). Supervised learning and systems with excess degrees of freedom. *Journal of Mathematical Psychology*.

Kintsch, W. & VanDijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review*, *85*, 363–394.

Luenberger, D. G. (1979). *Introduction to dynamic systems: Theory, models, and applications*. New York: Wiley.

Mackie, J. L. (1980). *The cement of the universe: A study of causation*. Oxford: Clarendon Press.

Mandler, J. M. (1978). A code in the node: The use of a story schema in retrieval. *Discourse Processes*, *1*, 14–35.

McClelland, J. L., St.John, M., & Taraban, R. (1989). Sentence comprehension: A parallel distributed processing approach. *Language and cognitive processes*, *4*, 287–335.

Miikkulainen, R. (1990). Script recognition with hierarchical feature maps. *Connection Science*, *2*, 83–101.

Miikkulainen, R. & Dyer, M. (1989). A modular neural network architecture for sequential paraphrasingof script-based stories. *Technical Report UCLA-AI-89-02*.

Nelson, K., Fivush, R., Hudson, J., & Lucariello, J. (1983). Scripts and the development of memory. In M. Chi (Ed.), *Trends in memory development research contributions to human development* (Vol. 9, pp. 52–70). New York: Karger.

Norvig, P. (1989). Marker passing as a weak method for text inferencing. *Cognitive Science*, *13*, 569–620.

Raaijmakers, J. (1987). A formal model for associative memory. In E. E. Roskam & R. Suck (Eds.), *Progress in Mathematical Psychology-1* (pp. 85–116). North-Holland: Elsevier Science Publishers.

Rumelhart, D. E. (1977a). Toward an interactive model of reading. In S. Dornic (Ed.), *Attention and performance VI* (pp. 265–303). Hillsdale, NJ: Lawrence Erlbaum.

Rumelhart, D. E. (1977b). Understanding and summarizing brief stories. In *Basic processes in reading: Perception and comprehension* (pp. 265–303). Hillsdale, NJ: Lawrence Erlbaum.

Rumelhart, D. E. (1984). Understanding understanding. In J. Flood (Ed.), *Understanding reading comprehension* (pp. 1–20). Newark: Delaware: International Reading Association.

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning internal representations by error propagation. In D. E. Rumelhart & J. L. McClelland (Eds.), *Parallel distributed processing. Volume 1: Foundations* (pp. 318–362). Cambridge, MA: MIT Press.

Schank, R. C. & Abelson, R. P. (1977). *Scripts, plans, goals, and understanding*. Oxford: North-Holland.

Schank, R. C. & Riesbeck, C. K. (1981). *Inside computer understanding*. Hillsdale, NJ: Lawrence Erlbaum.

Spiro, R. J. (1977). Remembering information from text: The state of schema approach. In R. C. Anderson, R. J. Spiro, & W. E. Montague (Eds.), *Schooling and the acquisition of knowledge*. Hillsdale, NJ: Lawrence Erlbaum.

Spiro, R. J. (1980). Accommodative reconstruction in prose recall. *Journal of Verbal Learning and Verbal Behavior*, *19*, 84–95.

St.John, M. & McClelland, J. L. (1990). Learning and applying contextual constraints in sentence comprehension. *Artificial Intelligence*, *46*, 217–257.

Trabasso, T., Secco, T., & VanDenBroek, P. (1984). Causal cohesion and story coherence. In H. Mandl, N. L. Stein, & T. Trabasso (Eds.), *Learning and comprehension of text* (pp. 83–111). Hillsdale, NJ: Lawrence Erlbaum.

Trabasso, T. & Sperry, L. L. (1985). Causal relatedness and the importance of story events. *Journal of Memory and Language*, *24*, 595–611.

Trabasso, T. & VanDenBroek, P. (1985). Causal thinking and the representation of narrative events. *Journal of Memory and Language*, *24*, 612–630.

Trabasso, T., VanDenBroek, P., & Suh, S. Y. (1989). Logical necessity and transitivity of causal relations in stories. *Discourse Processes*, *12*, 1–25.

VanDenBroek, P. (1988). The effects of causal relations and hierarchical position on the importance of story statements. *Journal of Memory and Language*, *27*, 1–22.

VanDenBroek, P. & Trabasso, T. (1986). Causal networks versus goal hierarchies in summarizing text. *Discourse Processes*, *9*, 1–15.

Warren, W. H., Nicholas, D. W., & Trabasso, T. (1979). Event chain and inferences in understanding narratives. In *New directions in discourse processing*. Hillsdale,

NJ: Lawrence Erlbaum.

Appendix A

MSP Causal Chain Representations of the Four Stories

Analysis of the Epaminondas story

*Epaminondas.*Once there was a little boy who lived in a hot country. One day his mother told him to take some cake to his grandmother. She warned him to hold it carefully so it wouldn't break into crumbs. The little boy put the cake in a leaf under his arm and carried it to his grandmother's. When he got there the cake had crumbled into tiny pieces. His grandmother told him he was a silly boy and that he should have carried the cake on top of his head so it wouldn't break. Then she gave him a pat of butter to take back to his mother's house. The little boy wanted to be very careful with the butter so he put it on top of his head and carried it home. The sun was shining hard and when he got home the butter had all melted. His mother told him that he was a silly boy and that he should have put the butter in a leaf so that it would have gotten home safe and sound.

Insert Table 5 about here

Using the above coding scheme, the abstract representation of the Epaminondas story would be the following ordered sequence of features:

$$1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22$$

which was then used to construct the Judy's Birthday story trajectory. This trajectory consists of 22 22–dimensional situation state vectors each of which contains only one active feature.

Insert Figure 10 about here

Analysis of the Judy's Birthday story

Judy's Birthday. Judy is going to have a birthday party. She is ten years old. She wants a hammer and a saw for presents. Then she could make a coat rack and fix her doll house. She asked her father to get them for her. Her father did not want to get them for her. He did not think that girls should play with a hammer and a saw. But he wanted to get her something. So he bought her a beautiful new dress. Judy liked the dress but she still wanted the hammer and saw. Later she told her grandmother about her wish. Her grandmother knew that Judy really wanted a hammer and a saw. She decided to get them for her because when Judy grows up and becomes a woman she will have to fix things when they break. Then her grandmother went out that very day and bought the tools for Judy. She gave them to Judy that night. Judy was very happy. Now she could build things with her hammer and saw.

Insert Table 6 about here

Using the above coding scheme, the abstract representation of the Judy's Birthday story would be the following ordered sequence of features:

$$1, 2, 3, 4, 5, 6, 6, 7, 8, 9, 10, 11, 3, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25$$

which was then used to construct the Judy's Birthday story trajectory. This trajectory consists of 25 23–dimensional situation state vectors each of which contains only one active feature. For example, the third and thirteenth state vectors in this trajectory would have feature number three active (set equal to one) and the remaining features inactive (set equal to zero).

Insert Figure 11about here

Analysis of the Tiger's Whisker story

*The Tiger's Whisker.*Once there was a woman who needed a tiger's whisker. She was afraid of tigers but she needed a whisker to make a medicine for her husband who had gotten very sick. She thought and thought about how to get a tiger's whisker. She decided to use a trick. She knew that tigers loved food and music. She thought that if she brought food to a lonely tiger and played soft music the tiger would be nice to her and

she could get the tiger's whisker. She went to a tiger's cave where a lonely tiger lived. She put a bowl of food in front of the opening to the cave. Then she sang soft music. The tiger came out and ate the food. He then walked over to the lady and thanked her for the delicious food and lovely music. The lady then cut off one of his whiskers and ran down the hill very quickly. The tiger felt lonely and sad again.

Insert Table 7 about here

Using the above coding scheme, the abstract representation of the Tiger's Whisker story would be the following ordered sequence of features:

$$1, 3, 2, 3, 4, 5, 6, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23$$

which was then used to construct the Tiger's Whisker story trajectory. This trajectory consists of 25 23–dimensional situation state vectors each of which contains only one active feature.

Insert Figure 12 about here

Analysis of the Fox and the Bear story

The Fox and the Bear. There was a fox and a bear who were friends. One day they decided to catch a chicken for supper. They decided to go together because neither one wanted to be left alone and they both liked fried chicken. They waited until night time. Then they ran very quickly to a nearby farm where they knew chickens lived. The bear, who felt very lazy climbed up on the roof to watch. The fox then opened the door of the henhouse very carefully. He grabbed a chicken and killed it. As he was carrying it out of the henhouse the weight of the bear on the roof caused the roof to crack. The fox heard the noise and was frightened but it was too late to run out. The roof and the bear fell in killing five of the chickens. The fox and bear were trapped in the broken henhouse. Soon the farmer came out to see what was the matter.

Insert Table 8 about here

Using the above coding scheme, the abstract representation of the Fox and Bear story would be the following ordered sequence of features:

$$1, 3, 2, 3, 4, 5, 6, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 20, 21, 22, 23, 24, 25$$

which was then used to construct the Fox and Bear story trajectory. This trajectory consists of 26 25–dimensional situation state vectors each of which contains only one active feature.

Insert Figure 13 about here

Appendix B

Mathematical Analyses

Equivalence of local and global probability distributions

Let X(t) be the global state of the multi-state probabilistic machine at time t or equivalently a d-dimensional vector (point) in situation state space.

The *local probability* $p_i(t)$ is the probability that the i*th* element of X(t) will have the value of one, given knowledge of X(t-1). The probability of a trajectory through situation state space $p(X(0), ..., X(M))$ is called the *global probability*. Given the local probabilities of the model are completely specified, it is possible to uniquely compute the probability of any trajectory in situation state space.

In particular, the parametric form of the local probability distributions are given by:

$$p_i(t) = S(\psi_i(X(t-1))) \tag{1}$$

where

$$S(\xi) = [1 + exp[-\xi]]^{-1} \tag{2}$$

and

$$\psi_i(X(t-1)) = \sum_{j=1}^{d} b_{ij} x_j(t-1) + a_i + \lambda \left(x_i^0(t) - 0.5\right) \tag{3}$$

with X(-1) defined as a vector of zeros.

The probability (belief) of a trajectory of length M in a d-dimensional situation state space (i.e., an ordered sequence of M d-dimensional global states) is represented in terms of the local probabilities as a first-order Markov chain by the following equations:

$$p(X(1),...,X(M)) = \prod_{t=0}^{M} p(X(t)|X(t-1)) \tag{4}$$

where

$$p(X(t)|X(t-1)) = \prod_{i=1}^{d} [x_i(t)p_i(t) + (1-x_i(t))(1-p_i(t))] \tag{5}$$

Now since:

$$x_i(t)p_i(t) + (1-x_i(t))(1-p_i(t)) = (1/Z_{it})exp[x_i(t)\psi_i(X(t-1))] \tag{6}$$

we have substituting into (5), we have:

$$p(X(t)|X(t-1)) = (1/Z_t)exp\left[\sum_{i=1}^{d}[x_i(t)\psi_i(X(t-1))]\right] \tag{7}$$

which can then be substituted into (4) to obtain:

$$p(X(0)...X(M)) = p(X(0))\prod_{t=1}^{M}(1/Z_t)exp\left[\sum_{i=1}^{d}[x_i(t)\psi_i(X(t-1))]\right] \tag{8}$$

which simplifies to:

$$p(X(0)...X(M)) = (1/Z)exp\left[\sum_{t=1}^{M}\left[\sum_{i=1}^{d}[x_i(t)\psi_i(X(t-1))]\right]\right] \tag{9}$$

Formal definition of relative self-information

Equation (9) may be used to easily compute the information content of a trajectory X(0)...X(M) up to an additive constant. In particular, the self-information of trajectory {X(0) ... X(M)} is simply equal to:

$$-LOG_2[p(X(0)...X(M))] = -\sum_{t=1}^{M}\left[\sum_{i=1}^{d}[x_i(t)\psi_i(X(t-1))]\right] + K \tag{10}$$

The constant K is a mathematically intractable function of the b_{ij} and a_i parameters, but is not a function of the trajectory {X(0) ... X(M)}. This observation is useful when we compare trajectories by subtracting the information content of one trajectory from another.

In this paper we have found it useful to define the *relative self—information* of a trajectory {X(1) ... X(M)} in situation state space as:

$$I(X(0),...,X(M)) = -LOG_2[p(X(0)...X(M))] - K \tag{11}$$

where it is assumed that the parameter $\lambda = 0$ and the constant K is as defined in (10). Also note that we have found it convenient to define the *retention interval parameter* ρ of the model as $\rho = 1/\lambda$.

Suppose that some subset of the features of a trajectory are "known", while other features are unknown and have to be estimated. That is, let κ be a subset of the features of a trajectory and their assigned values so that an element of κ has the form $x_i(t) = 1$ or $x_i(t) = 0$. Let ω be the remaining features of the trajectory. Thus, $p(X(0), ..., X(M)) = p(\kappa \cap \omega)$. Now we would like to compute MAP (maximum a posteriori) estimates of ω given κ, so we want to assign values to the features in ω such that the quantity:

$$p(\omega|\kappa) = p(\kappa \cap \omega)/p(\kappa)$$

is maximized. Since κ is not a function of the unknown feature values, we can therefore equivalently maximize the likelihood of the trajectory $p(X(0), ..., X(M)) = p(\kappa \cap \omega)$. This argument demonstrates that an algorithm which seeks local minima of the relative self-information is seeking "most believable" values of the unknown features in the trajectory given the values of the trajectory features that are known with respect to a belief structure (i.e., probability distribution) of the form of (9).

Convergence properties of the algorithm

Golden (submitted) describes sufficient conditions for the special recurrent back-propagation algorithm described in this paper to converge to a strict local minimum of the relative self-information measure by showing the back-propagation algorithm is a special case of a generalized version of the Brain-State-in-a-Box model algorithm (Anderson et al.,

1979; Golden, 1986a). The proof shows that a step constant of fixed size exists such that at each step of the algorithm, the relative self-information measure is strictly decreasing if an equilibrium point of the algorithm has not been reached. Moreover, the proof shows how to calculate the step constant explicitly, and provides a simple test for deciding if an activation pattern is an asymptotically stable equilibrium point of the algorithm.

For completeness, the recurrent back-propagation algorithm is now formally defined. Let $X^k(0)...X^k(M)$ be the current estimate of the trajectory at iteration k of the algorithm. For convenience, define the following "linearized" sigmoidal function, $\sigma(\mu)$, as $\sigma(\mu) = \mu$ for $0 \leq \mu \leq 1$, $\sigma(\mu) = 1$ for $\mu > 1$, and $\sigma(\mu) = 0$ for $\mu < 0$. Also note that:

$$dI/dx_i^k(t) = \sum_{k=1}^{d} x_k(t+1)b_{ki} + \sum_{j=1}^{d} b_{ij}x_j(t-1) + a_i + \lambda\big(x_i^0(t) - 0.5\big) \qquad (12)$$

so the algorithm is formally described by updating all "unclamped" units at time k according to the formula:

$$x_i^{k+1}(t) = \sigma\Big[x_i^k(t) + \gamma dI/dx_i^k(t)\Big] \qquad (13)$$

where the notation $dI/dx_i^k(t)$ refers to the derivative of I with respect to $x_i(t)$ evaluated at $X^k(0)...X^k(M)$. The parameter γ is a strictly positive scalar constant.

Author Notes

This research was supported in part by National Institute of Health post-doctoral fellowship HD06943 to the first author while he was at Stanford University, and the Program in Cognition and Neuroscience at the University of Texas at Dallas.

Correspondence and requests for reprints should be sent to Richard Golden, University of Texas at Dallas, School of Human Development, Program in Cognition and Neuroscience, GR41, Box 830688, Richardson, Texas 75083–0688. Email address: golden@utdallas.edu.

Table 1

A four-dimensional situation vector or equivalently a "point" in situation state space.

Situation Identification	Situation at time = t
f_1 = "Mary heard the ice-cream truck."	1 (present)
f_2 = "Mary wanted to buy some ice-cream."	1 (present)
f_3 = "Mary is eating ice-cream."	0 (absent)
f_4 = "Mary is sleeping."	0 (absent)

Table 2

A sequence of three four-dimensional situation vectors or equivalently

a "trajectory" in situation state space.

Situation Identification	Situation at time=1	Situation at time=2	Situation at time=3
f_1	1 (present)	1 (present)	0 (absent)
f_2	0 (absent)	1 (present)	0 (absent)
f_3	0 (absent)	0 (absent)	1 (present)
f_4	0 (absent)	0 (absent)	0 (absent)

Table 3.

Correlations between Trabasso et al. (1984) and MSP representations

for each of the four stories.

Epaminondas	Judy's Birthday	Tiger's Whisker	Fox and Bear
0.45*	0.67**	0.73**	0.52**

* $p < 0.05$ ** $p < 0.01$

Table 4

Correlations among human and model recall data and the causal connectivity and self—information performance measures.

	% Recall (model)	% Recall (Trabasso et al., 1984)	Causal Connect	# of Bits (surprise)
% Recall (model)				
% Recall (Trabasso et al., 1984)	0.96**			
Causal Connect	-.69	-.51		
Self Information (surprise)	-.83	-.93*	.40	

** $p < 0.05$, N=4 $R_{critical}(2) = 0.95$

* $p < 0.10$, N=4 $R_{critical}(2) = 0.90$

Table 5

Propositional representation of Epaminondas story. B = boy, M = mother, C=cake, G=grandmother.

Feature Number	Event Clause (proposition)
1.	B(Little)
2.	Lives(B, In HC)
3.	Tell(M, Bring(B, C, To G))
4.	Tell(M, B, Hold(B, C, Carefully))
5.	Tell (M,B,Prevent(Break(C(Crumbs))))
6.	Had_Put(B, Under(In(C, Leaf), Arm))
7.	Carry(B, C, To G)
8.	Arrive(B, At G, From M)
9.	C(Crumbs)
10.	Told(G, B, Is(B, Silly))
11.	Told(G, B, Carry(B, C, On Head))
12.	Told(G, B, Prevent(C(Crumbs)))
13.	Tell (G, B, Carry(B, Butter, To M))
14.	Desire (B, Hold(B, Butter, Carefully))
15.	Had_Put(B, Butter, On Head)
16.	Carry(B, Butter, To Home)
17.	Sun (Shining Hard)
18.	Arrive(B, At Home, From G)
19.	Butter(Melted)
20.	Told(M, B, Is(B, Silly))
21.	Told(M, B, Put(B, Butter, In Leaf))
22.	Told(M, B, Keep(Butter(Safe))

Table 6

Propositional coding scheme for Judy's Birthday story. J=Judy, H and S = Hammer and Saw, G=Grandma.

Feature Number	Event Clause (proposition)
1.	Has(J, BP, Soon)
2.	Is(J, 10 years)
3.	Desire(J, Have(J, H and S))
4.	Desire (J, Make(J, Coat Rack))
5.	Desire(J, Fix(J, DH))
6.	Ask(J, F, To Get (J, H and S))
7.	Desire (F, Not_Have (J, H and S)
8.	Desire(F, Not_Play(J, with H and S))
9.	Desire (F, Get (J, Present))
10.	Buy(F, Dress, for Judy)
11.	Liked(J, Dress)
12.	Tell (J, G, Desire (J, Have (J, H and S)))
13.	Knew (G, Desire (J, Have (J, H and S)))
14.	Desire (G, Have (J, H and S))
15.	Thought (G, Is (J, Adult))
16.	Thought (G, Is (J, Woman))
17.	Thought (G, Things (Broken))
18.	Thought (G, Have (J, Fix (J, Things (Broken))))
19.	Went (G, Out)
20.	Bought (G, J, Tools)
21.	Give (G, J, Tools)
22.	Happy (J)
23.	Think (J, Build (J, Things, with H and S))

Table 7

Propositional coding scheme for Tiger's Whisker story. W= woman, TW = tiger's whisker, M=music, F=food

Feature Number	Event Clause (proposition)
1.	Exists (W)
2.	Afraid (W, Of Tigers)
3.	Desire (W, Have (W, TW))
4.	Desire (W, Make (W, Medicine, for H))
5.	Husband (Sick)
6.	Think (W, About How (Get (W, TW)))
7.	Desire (W, Use (W, Trick))
8.	Knew (W, Love (Tigers, F and M))
9.	Think (W, Bring (W, F, To Tiger (Lonely)))
10.	Think (W, Play (W, M, To Tiger (Lonely)))
11.	Think (W, Tiger (Nice))
12.	Think (W, Get (W, TW))
13.	Went (W, Tiger's Cave)
14.	Live (Tiger, In Tiger's Cave)
15.	Had_Put (W, F, In Front of Cave)
16.	Had_Sang (W, M, In Front of Cave)
17.	GoTo (Tiger, Front of Cave)
18.	Eat (Tiger, Food)
19.	Walk (Tiger, To W)
20.	Thank (Tiger, W, For F and M)
21.	Take (W, TW, From Tiger)
22.	Run (W, Down Hill)
23.	Tiger (Lonely)

Table 8

Propositional coding scheme for Fox and Bear story. F=fox, B=bear, R=roof.

Feature Number	Event Clause (proposition)
1.	Exists (F and B)
2.	Friends (F, B)
3.	Desire (F and B, Catch (F and B, Chicken, For Supper))
4.	Desire (F and B, GoTo (F and B, Chickens, Together))
5.	Dislike (F and B, Being Separated)
6.	Like (F and B, Fried Chicken)
7.	Wait (F and B, Until Night Time)
8.	Ran (F and B, To Farm X, Very Quickly)
9.	Knew (F and B, Farm X, Has Chickens)
10.	B(Lazy)
11.	Climbed(B, On Roof)
12.	Desire (B, Watch (B, Get (F, Chicken)))
13.	Open (F, Door (Henhouse), Very Carefully)
14.	Grab (F, Chicken)
15.	Kill (F, Chicken)
16.	Carry (F, Chicken, Out of Henhouse)
17.	Cause (W(B), Roof (Cracks))
18.	Heard (F, Noise)
19.	F(Frightened)
20.	Unable (F, Escape Henhouse)
21.	Fall (R and B, Into Henhouse)
22.	Kill (R and B, Five Chickens)
23.	Trapped (F and B, In Henhouse)
24.	Arrive (Farmer, At Henhouse)
25.	Desire (Farmer, Investigate (Noise))

Figure Captions

Figure 1. A two-state MSP machine's "marking" at the current instant of time, and the next instant of time. The heavy dark arrow is called a "transition", while the thin dark arrow is called a "control". The presence of a dark circle in a local state indicates that local state is "marked." More formally, the local probability $p(B(t+1) = 1|A(t))$ in this case is a monotonically increasing function of the binary random variable A(t).

Figure 2. A two-state MSP machine's "marking" at the current instant of time, and the next instant of time. Notice how "markers" flow though transitions, but do not flow through controls. More formally, the local probabilities $p(C(t+1) = 1|B(t))$ and $p(B(t+1) = 0|B(t))$ are monotonically increasing functions of the binary random variable B(t).

Figure 3. The recurrent network can best be visualized as a feed-forward network (Rumelhart et al., 1986) where the pattern of connectivity from one time instant to the next is identical. In this example, a trajectory consisting of five four-dimensional situation state vectors is represented as a feed-forward network.

Figure 4. Flow chart of the operation of the simulation model for comprehending and recalling a single story from memory.

Figure 5. Percentage of statements recalled from a story as a function of retention interval. Larger values of ρ correspond to longer retention intervals.

Figure 6. The percentage of statements recalled from each of the four stories. Data from Trabasso et al. (1984) is plotted for comparison purposes.

Figure 7. Average relative self-information as a function of retention interval. Smaller values of the information measure correspond to larger measures of "causal coherence." Larger values of ρ correspond to longer retention intervals.

Figure 8. Relative causal connectivity as a function of retention interval. Larger values of the relative causal connectivity measure correspond to larger measures of "causal coherence." Larger values of ρ correspond to longer retention intervals.

Figure 9. Number of steps a misordered statemente "drifted" towards its canonical position as a function of retention interval. Larger values of ρ correspond to longer retention intervals.

Figure 10. Epaminondas MSP causal chain representation. The numbered states refer to propositions in Table 5.

Figure 11. Judy's Birthday MSP causal chain representation. The numbered states refer to propositions in Table 6.

Figure 12. Tiger's Whisker MSP causal chain representation. The numbered states refer to propositions in Table 7.

Figure 13. Fox and Bear MSP causal chain representation. The numbered states refer to propositions in Table 8.

Figure 1.

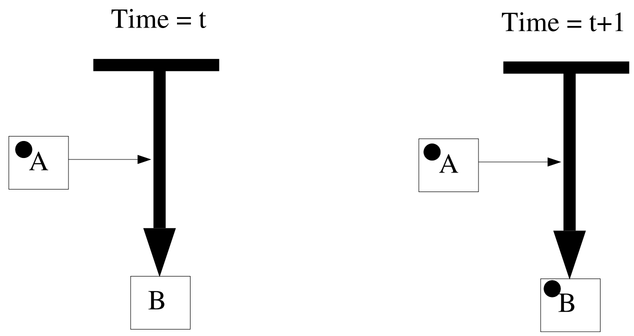

Figure 2.

Figure 3.

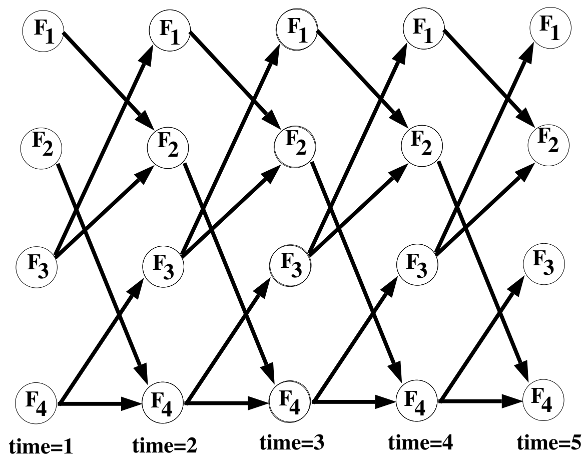

Figure 4.

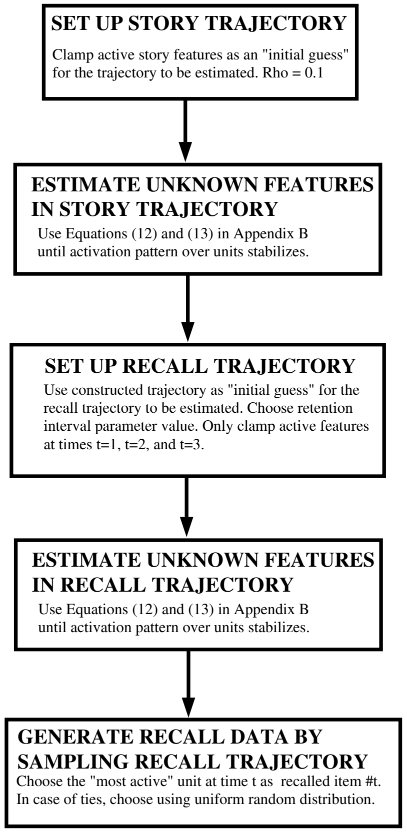

Figure 10.

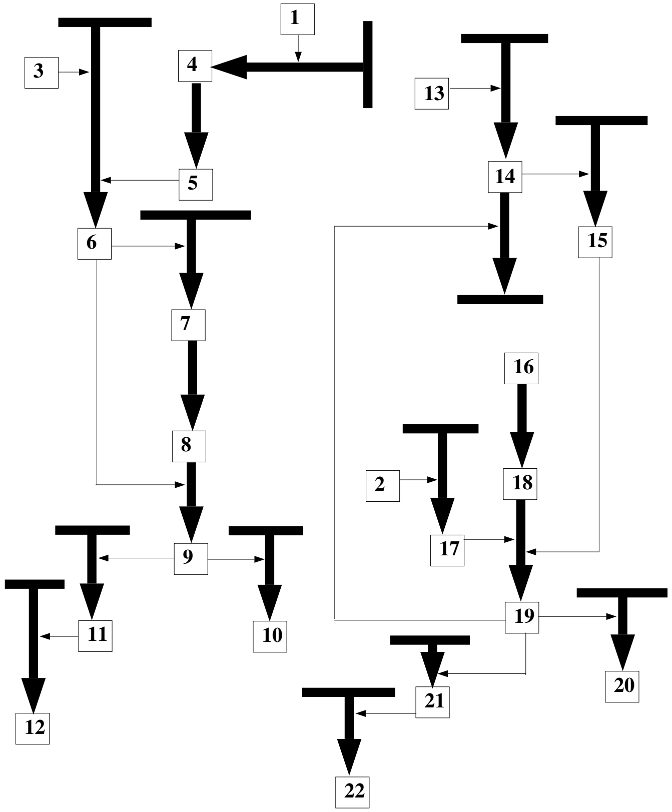

Figure 11.

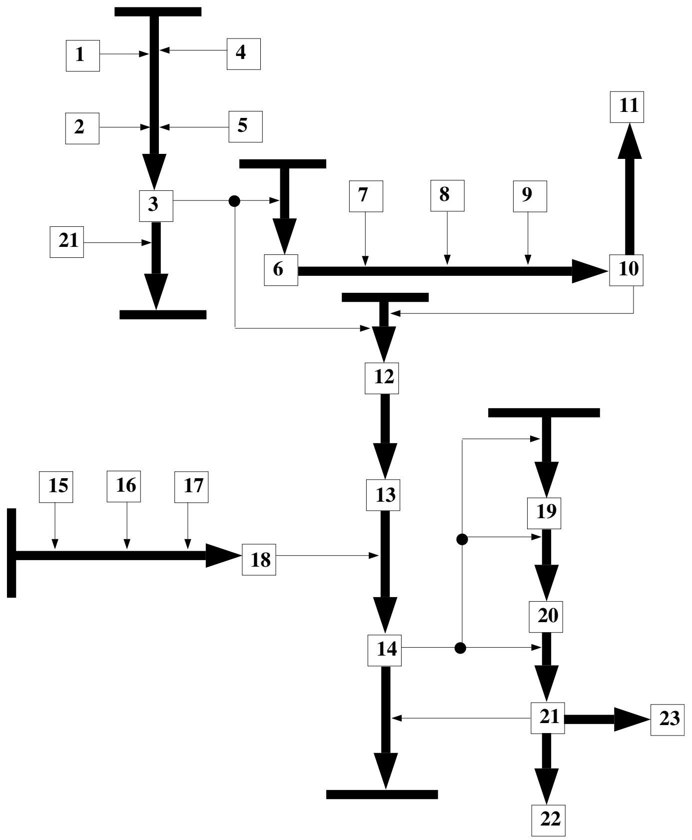

Figure 12.

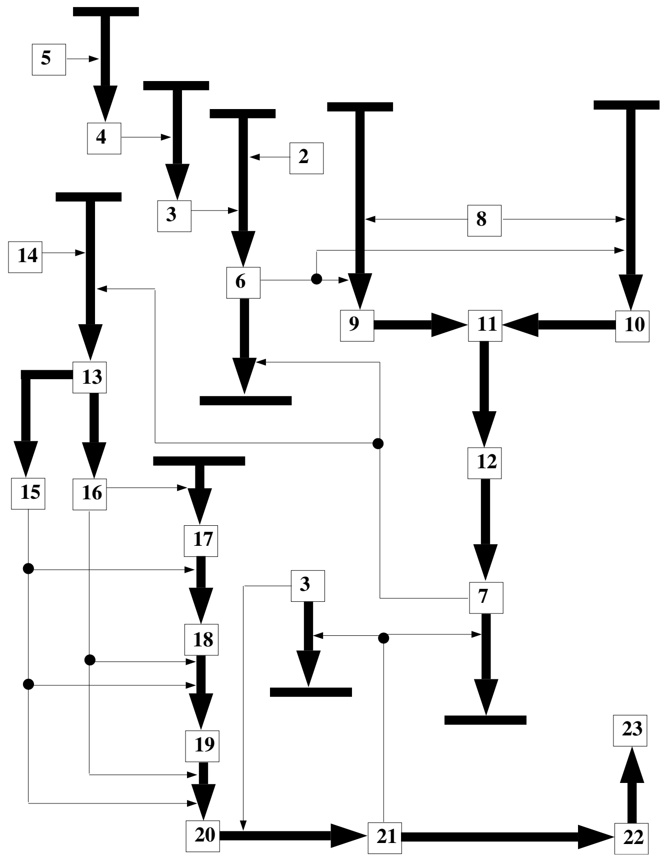

Figure 13.