Title
Representing Events using Fuzzy Temporal Boundaries

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 27(27)

ISSN
1069-7977

Authors
Harter, Derek
Lu, Shulan

Publication Date
2005

Peer reviewed

Representing Events Using Fuzzy Temporal Boundaries

Shulan Lu (Shulan_Lu@tamu-commerce.edu)
Department of Psychology, Texas A&M University
Commerce, TX 75429 USA

Derek Harter (Derek_Harter@tamu-commerce.edu)
Department of Computer Science, Texas A&M University
Commerce, TX 75429 USA

Abstract

This study investigates whether people represent the beginnings and ends of events as fuzzy temporal frames and subsequently construct event temporal relations. The study adopted Allen's (1984; 1991) seven logical categories of temporal relations. Constructing the seven relations often requires the accurate encoding and (or) retrieval of the beginnings and ends of events. We used a recurrent neural network to simulate the performance of humans in representing event temporal relations. The network was given fuzzy event inputs, generated using Zadeh (1975) fuzzy logic functions, and trained to judge the event temporal relations. We compared the performance of the recurrent neural network to that of humans in a task where they were asked to remember and reconstruct everyday temporal relations. The simulations showed that the recurrent neural network mimicked human judgments in the correct judgments, preferences toward particular temporal relations, and directions of error. The results support that event temporal relations are best thought of as fuzzy analogue representations in humans and the simulated network.

Keywords: temporal representation; events.

Introduction

Our everyday life consists of various events we experience and enact. Consider a few examples: going to work, having dinner at a restaurant, and meeting with friends. Each of these events has a number of subevents. These subevents are related to each other in various dimensions such as by causal links. Most of all, events unfold in time and relate to each other temporally. In this paper, the term <u>event</u> refers to something that happens at a given place with a beginning and an end.

Many researchers suggest that event temporal properties provide a basic framework for structured event representations (Allen, 1984; Barsalou, 1999; Freyd, 1987; Graesser, Kassler, Kreuz, & McLain-Allen, 1998). Events have time spans (duration) over which they take place. Events also have temporal locations relative to each other (temporal relations). Most studies in psychology tend to treat events as following one another within the event hierarchy (e.g., no overlapping of events at the same level of a hierarchy). However, when two or more events are occurring, they can have overlaps in time. In addition, the overlap in time can vary. For some events, the beginnings are the same. For others, neither the beginnings nor the ends are the same. Event representations may have more dynamic event nodes linked to each other than the temporal links suggested by simple hierarchical structures (Schank, 1999).

What temporal relations are necessary for constructing event representations? In artificial intelligence, Allen (1984; 1991) proposed a representation that contains seven relational primitives. Figure 1 provides an illustration of Allen's seven event temporal relations. Each double-headed arrow in Figure 1 represents an event that occurs over some time interval, whereas each arrow-head represents either the beginning or the end of an event. The relation between each pair of events is described by one of seven predicates. These seven primitives have been used as basic temporal operators for automated planning and reasoning systems that reason with and make logical deductions about event temporal relations.

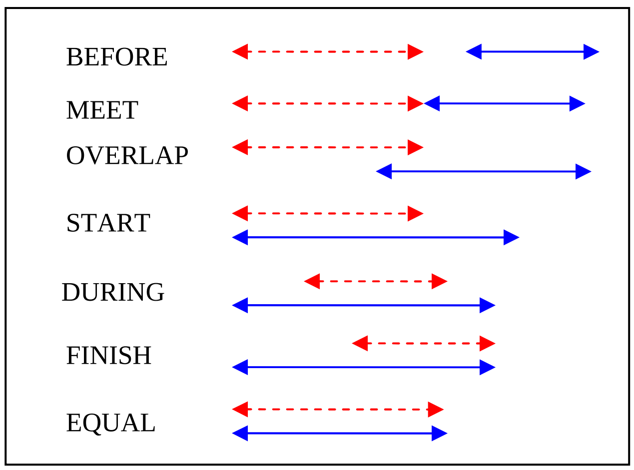

Figure 1: Temporal Relations (Allen, 1991, pp.5).

Some events have crisp, clear cut temporal frames. *Moving a coffee mug* is an example, where the beginning, the end, and the temporal trajectory are not ambiguous. Some events have fuzzy temporal frames. *Viewing sunset* is an example, where the beginning and end are ambiguous, and the temporal trajectory may be on and off. The psychological frame of an event may or may not deviate from its physical frame. Allen's representation captures some intuitive aspects of human temporal reasoning. For example, people tend to make relatively good estimations of event durations (Golding, Magliano, & Hemphill, 1992; Loftus, Schooler, Boone, & Kline, 1987), whereas people tend not to be good at estimating the points of time when events take place (Golding et al., 1992; Linton, 1975). Such evidence suggests that an interval based representation may be a more natural way of relating events and drawing inferences about the temporal relations between events.

When people perceive and remember event temporal relations, it appears that the beginnings and ends of the event psychological frame may deviate from the physical frame (Lu, 2004). Figure 2 depicts a psychological frame of event temporal properties. An Event E_i has a psychological frame of E_i (b_i, c_i, e_i), where b_i is the parameter that controls the region where the beginning point of an event E_i is likely to have occurred, c_i is the (psychological, not objective) fuzzy centroid during which Event E_i occurs, and e_i is the parameter that controls the region where the end point of event E_i is likely to have occurred. The perceived properties of an event interval may be fuzzy, that is, an approximation to the physical frame. In the figure, probability of 1 represents areas where there is complete certainty that the event occurred. Probability 0 represents times when there is certainty that the event was not occurring. Intermediate values represent fuzzy regions, where recall or perception are not completely sure of the occurrence or non-occurrence of the event.

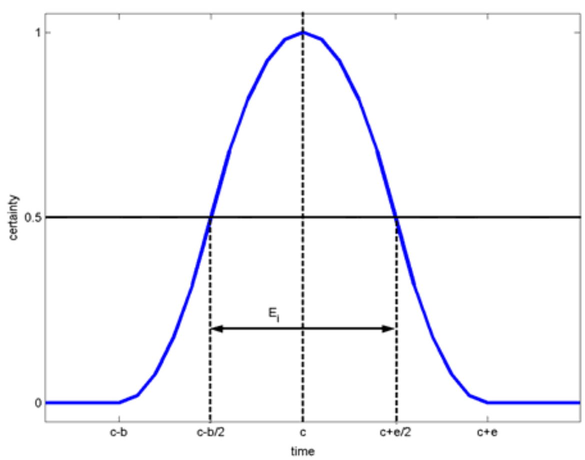

Figure 2 : Fuzzy event psychological frame.

How might humans go about constructing event temporal relations? Symbolic approaches in artificial intelligence, when representing temporal relations, typically assign temporal stamps to the events and build representational graphs. There are similar proposals in psychology on how event temporal properties are represented (Linton, 1975). However, such approaches have been challenged by the connectionist and embodiment theorists (Barsalou, 1999; Boroditsky, 2000; Elman, 1990; Lakoff & Johonson, 1999; Michon, 1993). Elman (1990) showed that temporal properties can be modeled in their implicit effects on processing rather than explicit spatial representations. Boroditsky (2000) showed that time is grounded and built up through experiential domains. In this paper, we propose that the event psychological temporal frames are fuzzy (inexact) and that event temporal relations get constructed via the dynamic processing of events and subsequent formation of temporal patterns of activity. More specifically, based on the proposal in Figure 2, the constructed temporal relation between events will be influenced by the fuzzy, analogue nature of event temporal properties, where the beginnings and ends of the events are represented to be somewhere in the fuzzy regions.

Elman (1990) showed that simple recurrent neural networks (SRNN) are ideal for modeling cognitive processes that depend on not only spatial but temporal properties. Recurrent connections save and propagate past states of a network to the current context to allow for the recognition of patterns with temporal properties. Therefore SRNN allow for the history of the inputs, and the dynamics of the system, to affect the pattern recognition. This paper investigates how humans construct temporal relations by comparing the performance of simple recurrent neural network models against the performance of human perception and memory of event temporal relations. In the SRNN simulation presented, the networks received events with fuzzy beginnings and ends. We compare the performance of the SRNN with human performance in a judgment task about event temporal relations. In the human experiment, participants made judgments of complex everyday events such as two subevents of *stirfry vegetables*.

If humans perceive event temporal properties in a fuzzy, analogue world, then the temporal representations may not preserve the detailed, accurate properties of the physical temporal frames. In the network simulation, the proportions of the correct judgments made by SRNN receiving fuzzy events should mimic the judgments made by humans remembering everyday events. In addition, the error patterns of SRNN should exhibit some of the same patterns as the human errors.

SRNN Simulation

Training

In the network simulation, we trained a SRNN network to categorize event temporal relations. The network received 25 discrete time steps as inputs. There were two inputs to the network, representing the certainty of the presence of two events, A and B. The network had 7 outputs, one for each of the 7 possible temporal relations, and a number of hidden units between the input and output layers. Networks were trained by creating a training set of 1000 examples (divided approximately equally between the 7 relations). For each given input relation, an output unit was trained to become active as soon as it recognizes its target temporal relation.

Figure 3 presents an example of an OVERLAP temporal relation used in the training. Here event A has an objective starting time of $t = 5$ and objective end time of $t = 10$, whereas event B started at $t = 7$ and ended at $t = 15$. In the network simulation uncertainty was added into the input representation. This uncertainty, we believe, better models the situation of humans where event comprehension is often noisy and uncertain. Therefore, instead of binary inputs of 0 or 1, the event inputs were real valued numbers ranging from 0 to 1 inclusive, where 1 indicates certainty that the event is occurring, 0 indicates certainty that the event is not occurring, and numbers in between indicate more intermediate certainty. When training the network output, the unit representing the OVERLAP relation would be trained, for the event in Figure 3, to begin responding with a 1 value at time step 19.

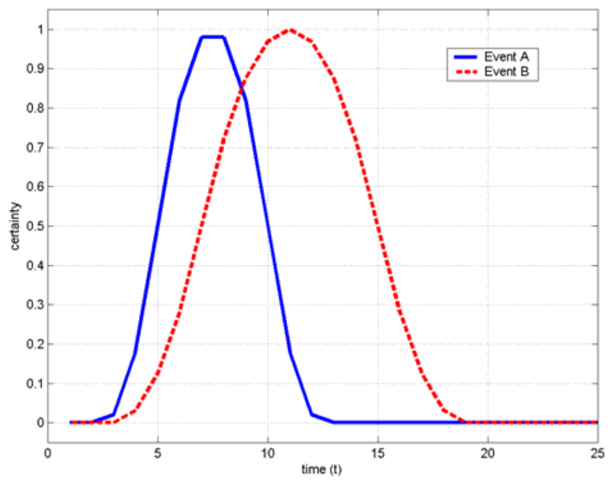

Figure 3: Example fuzzy event input to SRNN.

The s-function (Equation 1) was used to generate the real-valued certainty levels shown in Figure 3. The s-function defines a membership, or certainty curve, over some time range x. In the s-function, β serves as the inflection point, where the certainty measure is 0.5 (indicating neutrality in the perception of whether the event is occurring or not occurring), and α and γ define beginnings and ends of the event interval where their certainty ranges from 0 (absolutely certain that an event is not occurring) to 1 (absolutely certain that an event is occurring) respectively. In our simulation, if Event A had an actual begin time at t = 5, we used this time step as the inflection point (β = 5 in the s-function). We calculated α and γ based on the length of the event in such a way that the complete certainty an event is occurring only happens at the midpoint of the actual event. Since event A lasted from t = 5 to t = 10, we set γ to be at this midpoint, or 7.5. The α parameter was then set to be symmetrical to γ from the inflection point, or 2.5 in this case. The final call to the s-function for the beginning of Event A was therefore with the parameters S(x; 2.5, 5.0, 7.5), which generated the certainty curve for the beginning of the Event A shown in Figure 3. We used a similar method to generate the curve for the end of the event, but simply reversed the sense of the s-function so it generated a decreasing rather than increasing certainty curve over the end-point of the event. The equation for generating the total certainty curve is known as the Π-function and is given in Equation 2.

$$S(x;\alpha,\beta,\gamma)=\begin{cases}0 & \text{for } x\le\alpha\\ 2\left(\dfrac{x-\alpha}{\gamma-\alpha}\right)^2 & \text{for } \alpha\le x\le\beta\\ 1-2\left(\dfrac{x-\gamma}{\gamma-\alpha}\right)^2 & \text{for} \beta\le x\le\gamma\\ 1 & \text{for } x\ge\gamma\end{cases} \qquad \text{(Equation 1)}$$

$$\Pi(x;\beta,\gamma)=\begin{cases}S(x;\gamma-\beta,\gamma-\dfrac{\beta}{2},\gamma) & \text{for x}\le\gamma\\ 1 - S(x;\gamma,\gamma+\dfrac{\beta}{2},\gamma+\beta) & \text{for x}\ge\gamma\end{cases} \qquad \text{(Equation 2)}$$

Results

Number of Hidden Units We first determined an appropriate number of hidden units to use for the SRNN networks in our simulation. The networks were trained with a number hidden units ranging from 20 to 100 in increments of 5, 10 networks for each hidden unit parameter. We then used the performance of the 10 networks on a separate set of test events to determine the number of hidden units that will be selected.

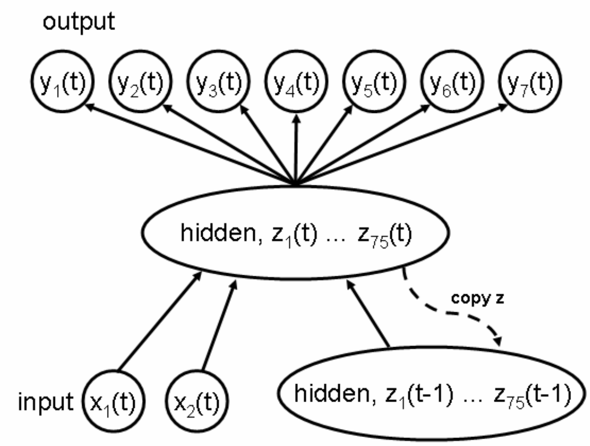

Figure 4: Network architecture used in relational judgments simulation.

The best single performance occurred with number of hidden units set to 60, where an error of .027 is reached. The best average performance occurred with 75 hidden units, which was the number of units we used to generate the performance data for the simulation in the next section.

Relational Judgments A trained Elman SRNN with 75 hidden units, as shown in Figure 4, performed temporal relational judgments. The Elman SRNN feeds back a copy of the activation of the hidden units at a previous time step in order to achieve its recurrent connectivity. In the simulation using this architecture, the probability of the network making correct judgments of event temporal relations was .54 (SD = .33) on average.

Table 1: Proportion of Network Judgments on Fuzzy Events.

	Network Judgment						
Events in the World	BEFORE	MEET	OVERLAP	START	DURING	FINISH	EQUAL
BEFORE	.738	.262	0	0	0	0	0
MEET	.132	.840	0	.028	0	0	0
OVERLAP	0	.118	.176	.426	.067	.066	.147
START	.015	.008	.015	.649	.076	.046	.191
DURING	0	0	.028	.448	.214	.083	.228
FINISH	.013	0	.084	.078	.019	.221	.584
EQUAL	0	0	0	.033	0	0	.966

Table 1 presents the confusion matrix of the network performance on the fuzzy test event set. The test set was a new set randomly generated and different from the training set used to train the network. In Table 1, rows represent the actual events presented to the network, and columns hold data on how the network actually performed. We report the proportion of the actual network judgments for a particular event temporal relation. For example, of the MEET events tested, the network correctly made MEET judgments 84% of the time. However, looking across the row, the network misjudged 13% of the MEET events as BEFORE events and made a further smaller error in judging 3% of the MEET events as START events. The network made the correct judgments on each of the seven temporal relations above chance level (assuming chance = 1 / 7 = .143), with the OVERLAP receiving the lowest proportion of correct judgments (.176). The network performed better on some event temporal relations (e.g., EQUAL) than others (OVERLAP).

Discussion The network had the best performance when categorizing BEFORE, MEET, and EQUAL events. The performance on BEFORE and MEET events are well above average at .738 and .840 respectively. The simulation does particularly well on EQUAL event relations achieving a .966 accuracy. The network had the most problems with OVERLAP, DURING and FINISH events, doing slightly better but still not well on START events.

The excellent performance on EQUAL events is intuitively not surprising. In an EQUAL relation, both events last exactly the same number of time steps and both start and end at the same time step. All these features may allow for fairly simple networks to come up with solutions to recognize EQUAL.

BEFORE and MEET appear to be fairly similar. To discriminate between the two, the network has to develop some recognition of whether a temporal gap occurred between the two events or not, and this memory may need to be held for many time steps. Of course, the larger the gap, the easier it might be to recognize and remember the gap. The confusion may increase if the temporal gap between events is smaller. Not surprisingly, the network had some confusions between BEFORE and MEET at times. Overall, the network performed well in judging BEFORE and MEET, which suggests that the network could form distinct activation patterns for BEFORE and MEET respectively.

The relations with temporal overlap (e.g., DURING, OVERLAP) and mixed synchronous/asynchronous end points (FINISH, START) seem to be the most difficult for the SRNN. These four relations appear to form a third distinct group, and can be easily confused with one another by the network.

Remembering Everyday Event Time

The dynamic mental representations of events may include the transitional states between events. For example, Freyd (1987) showed that people tend to project the next state of an event even if a picture they view does not contain that anticipated subsequent state. What tasks in everyday life have closer approximation to the fuzziness entailed in the simulated event relation judgment task? It is not hard to imagine that people routinely need to construct and reconstruct something like, “what should I do while I am doing this?” or “what should I do next?” The temporal properties of everyday events are likely to be fuzzy, for example, as a result of the intrinsic fuzzy nature of the everyday events and memory retention loss.

In the current experiment, participants were presented two events that were part of a routinely enacted activity and that were classified theoretically as having one of Allen’s seven temporal relations. Participants read the events in a context, where the events occur, such as “imagine someone boarding a plane.”, and made judgments on event temporal relations. Consider the example stimulus below.

Context: Imagine someone boarding a plane.
Events: She went through airport security screening.
Her carry-on bags were X-rayed.

The two events in the above example have the DURING relation. That is, the event of *X-raying carry-on bags* typically occurs during the process of a person *going through security check*. If humans represent the event temporal frame in fuzzy regions, the probability of correctly judging the event temporal relations would not be high.

Table 2: Proportion of human judgments on everyday events.

	Human Judgment						
Events in the World	BEFORE	MEET	OVERLAP	START	DURING	FINISH	EQUAL
BEFORE	.415	.340	.142	.048	.007	.013	.035
MEET	.165	.471	.230	.046	.018	.022	.048
OVERLAP	.092	.254	.370	.097	.017	.013	.158
START	.134	.199	.188	.156	.011	.017	.296
DURING	.129	.167	.245	.085	.033	.028	.314
FINISH	.103	.245	.287	.096	.017	.050	.204
EQUAL	.024	.044	.068	.061	.024	.020	.759

Method

Participants There were 68 college students at The University of Memphis who participated for course credit.

Materials A sample of events from everyday activities were collected. To ensure generality, the events were chosen from a wide range of everyday activities that college students routinely experience and perform. Three raters were trained to understand Allen's scheme, and made judgments on how each two events were related in time separately. The materials used in the experiment were the items agreed upon by all three judges. There were 8 test items for each of the 7 temporal relations in Figure 1. Therefore, there were 56 test items in total.

Procedure Participants were told that they would make judgments concerning the temporal relations of everyday events. Participants were shown a diagram similar to Figure 1, except that the word labels (e.g., BEFORE) were stripped. Experimenter did not launch experiment until participants understood all 7 relations.

Pairs of events along with their contexts were presented to participants one at a time on the computer screen. The two events were listed as two sentences in two rows separately. Participants were told that the presentation order of the two events was random and did not necessarily correspond to the actual order in their daily activities. Participants were instructed to read the two events and recall how they normally performed the two events in the activity they read and reconstructed. They proceeded to the next screen once they felt they comprehended the events and reconstructed the timing for enacting the events. The two events and the seven-choice diagram were presented on the screen. Participants made their judgments about the temporal relation between the two events at the end of the trial.

Results

On average, the proportion of the correct judgments in Experiment 2 was .322 (SD = .261). The proportional error rates of the seven relations were the following: BEFORE (.415), MEET (.471), OVERLAP (.370), START (.156), DURING (.033), FINISH (.050), and EQUAL (.759). The confusion matrix for the human experiments is shown in Table 2. Table 2 shows that humans clearly had preferences toward EQUAL events and that humans frequently mistook OVERLAP, START, DURING, and FINISH events as EQUAL events.

Overall the network performance was significantly correlated with the human performance (including both correct judgments and errors), $r = .67$, $p < .05$. Two sets of correlational analysis were performed on the correct human judgments and correct network judgments. The Spearman correlation showed that there were significant correlations of the order of human judgments in the experiment with that of the network judgments, $r = .79$, $p < .05$. The Pearson correlations showed that there were one way significant correlations between the proportions of correct human judgments and the proportions of the network judgments, $r = .25$, $p < .05$.

Discussion

When humans recalled and reconstructed events based on their routine activities, they appear to have remembered more details of the transitional temporal properties. For example, participants made correct judgments on four out of seven event temporal relations significantly above the chance level, and made correct judgments on START events no lower than chance level (.14). DURING and FINISH were the only event temporal relations that rarely got constructed.

Compared with the network simulation, humans' recall and reconstructions of event temporal relations had lower proportions of correct judgments on average (with a mean of .32 versus .54). Such differences were partly due to the extremely low recall of DURING (.033) and FINISH (.050) in human judgments. The proportions of correct human judgments on BEFORE, MEET and EQUAL events were also lower than the proportions of correct network judgments.

Overall, the network receiving fuzzy input representations mimicked human judgments in everyday events to a larger extent. The order in which seven temporal relations received correct judgments is not the only index.

For example, in both cases, OVERLAP, START, DURING, and FINISH events were frequently mistaken as EQUAL events. In addition, the pattern of confusions between BEFORE and MEET were similar as well. There were two exceptions worth noting. Unlike the network receiving fuzzy events, humans made more correct judgments on OVERLAP and more false alarms on OVERLAP events. Furthermore, humans did not make as many correct judgments and false alarm on START as the network did.

General Discussion

The SRNN networks and humans appear to have many of the same strengths and weaknesses when performing event temporal relation judgment tasks. Some event temporal relations are very simple to represent and recognize, whereas others are much more difficult. Humans and the simulated networks consistently demonstrated three distinct preferences in representing event temporal relations and tended to mistake difficult temporal relations as one of the three preferences. The results suggested that the event temporal properties tend to be represented in a fuzzy analogue manner and the beginning and the ends of events are not crystal clear to observers.

While BEFORE and MEET were confused with one another at times, in general both the networks and humans tended to do well in recognizing these and distinguishing between the two. EQUAL appeared to be particularly easy as a temporal pattern that can be uniquely captured when it occurred, both by the simulations and humans. Some confusions can occur between START and/or FINISH with EQUAL, but these appear to reflect the difficulty of representing the START/FINISH relation. Events that have asynchronous beginning and/or endings appear to be particularly difficult for both SRNN networks and humans to distinguish between. These relations, OVERLAP, START, FINISH and DURING, are prone to be confused with one another and are difficult to recognize with any high degree of accuracy.

The network performances also differed from humans in some ways. When humans made judgments on connected, everyday events that are part of an activity, humans failed to outperform the network. The errors were distributed more broadly. The OVERLAP relations received both higher hit rate and false alarms when humans judged everyday events, whereas the OVERALP relations received poor judgments when the network received fuzzy events. This may suggest that humans are able to construct a pattern that could represent asynchronous events. Such results may be due to some differences in the simulated task and the experiment described. The everyday events are embedded in an overarching event structure. There is possibly vagueness in linguistic expressions of everyday events. In future studies, it will be necessary to construct simulated events that have events embedded in overarching higher order constituents (e.g., schema) and systematically investigate the representational changes and potential loss of temporal properties in constructing the event temporal relations. However, the task needs to be sufficiently different from some temporal reasoning task. For example, event X occurs before event Y, and event Z and event Y occur simultaneously. When does event X occur in relation to event Z?

The results reported in this paper suggested that temporal representations of events are richer than previously assumed. Events can occur one after another, immediately follow one another, overlap with one another, and occur simultaneously. The event temporal frames are often represented in fuzzy regions.

References

Allen, J. F. (1984). Towards a general theory of action and time. *Artificial Intelligence, 23*, 123-154.

Allen, J. F. (1991). Time and time again: The many ways to represent time. *International Journal of Intelligent Systems, 6,* 341-355.

Barsalou, L. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences, 22*, 577-609.

Boroditsky, L. (2000). Metaphoric Structuring: Understanding time through spatial metaphors. *Cognition, 75*, 1-28.

Elman, J. L. (1990). Finding structure in time. *Cognitive Science, 14,* 179-211.

Freyd, J. J. (1987). Dynamic mental representations. *Psychological Review, 94,* 427-438.

Golding, J. M., Magliano, J., & Hemphill, D. (1992). When: A model for answering “when” questions about future events. In T. W. Lauer, E. Peacock, & A. C. Graesser (Eds.), *Questions and information systems* (pp. 213-282). Hillsdale, NJ: Earlbaum.

Graesser, A. C., Kassler, M. A., Kreuz, R. J., & McLain-Allen, B. (1998). Verification of statements about story worlds that deviate from normal conceptions of time: What is true about *Einstein’s Dreams*? *Cognitive Psychology, 35,* 246-301.

Lakoff, G., & Johnson, M. (1999). *Philosophy in flesh: The embodied mind and its challenge to Western thought*. New York: Basic Books.

Linton, M. (1975). Memory for real-world events. In D. A. Norman, & Rumelhart, D. E. (Eds.), *Explorations in cognition* (pp. 376-404). San Francisco, CA: Freeman.

Lu, S. (2004). Perceiving, remembering, and describing event temporal relations. *Dissertation Abstracts International,* 65 (06), 3191. (UMI No. 3137799)

Loftus, E. F., Schooler, J. W., Boone, S. M., & Kline, D. (1987). Time went by so slowly: Overestimation of event duration by males and females. *Applied Cognitive Psychology, 1,* 3-13.

Michon, J. A. (1993). Concerning the time sense: The seven pillars of time psychology. *Psychologica Belgica, 33,* 329-345.

Schank, R. C. (1999). *Dynamic memory revisited.* New York: Cambridge University Press.

Zadeh, L. A. (1975). Fuzzy Sets and Their Applications to Cognitive and Decision Processes, Academic Press.