

An Event Metric and an Episode Metric for a Virtual Guide

Felix Rabe and Ipke Wachsmuth

Artificial Intelligence Group, Faculty of Technology, Bielefeld University, Bielefeld, Germany

Keywords: Episodic Memory, Event Metric, Episode Metric, Virtual Agent, Guidance.

Abstract: In this paper we introduce a metric to compare events and episodes in the episodic memory system of a virtual agent. The agent, a virtual tour guide based on a belief – desire – intention cognitive architecture, uses his memories to improve the walks around a virtual city. The guide's past experiences are memorized as events and organized into episodes. Each event is indexed along six dimensions and is comparable on each dimension with a distinct distance function. This is then utilized to measure the similarity between episodes.

1 INTRODUCTION

Past experiences influence all of our actions and let us have an expectation of what might happen next. In interaction we rely on past episodes with other persons, we improve our behavior based on what we experience and store in our episodic memory. Since all of our actions are influenced by our past experiences, it is important also for a virtual agent to have an episodic memory, so his behavior in interaction with humans is improved.

Our work is centered around a virtual humanoid agent *Max*, cf. (Leßmann et al., 2006; Becker et al., 2006) that has already a lot of skills and is based on a belief – desire – intention (BDI) cognitive architecture. In (Rabe and Wachsmuth, 2012) we introduced a cognitively motivated episodic memory for our agent and a virtual guide scenario, where episodic memory is employed. The memories are episodes of events, that means past episodes (guided tours) consist of multiple events (e.g. explaining a sight). Each event can be accessed via six different indices: time, space, protagonists, intention, causality and emotion.

In this paper we introduce a metric with several distance functions to compare events among each other and measure similarity between episodes.

2 RELATED WORK

Our work on constructing a memory system for a virtual guide is influenced by theories on episodic memory, as well as on event-indexing and event segmentation. Episodic memory, as defined in (Tulving, 1972),

deals with temporally dated episodes or events, and temporal-spatial relations among these events. Every “item in episodic memory” (Tulving) is a more or less faithful record of a person's experience of an occurrence.

The Event-Indexing Model (Zwaan et al., 1995) describes how readers of short stories construct a model of the situation in the text. As readers understand what is happening in the story they update the model along five indices: Time, Space, Causality, Intentionality and Protagonists. These dimensions store answers to the questions of what happened when, where, why and how, and who was involved.

In Event Segmentation Theory, an event is defined as “a segment of time at a given location that is conceived by an observer to have a beginning and an end” (Zacks and Tversky, 2001). (Allen et al., 2008) propose that emotion is an important contextual cue for episodic memory and provide evidence that cognition is either moderated or mediated by basic affective processing. Another related field is Experience Management. The application makes use of several distance functions to provide matching solutions (Bergmann, 2002).

In earlier work (Rabe and Wachsmuth, 2012) we discussed several other applications of agents employed as guides and computational episodic memory systems. Our recent work is somewhat related to case-based reasoning (Kolodner, 1993; Aamodt and Plaza, 1994), except there is no generalization necessary for episodic memory. Concerning the event metric, (Tecuci and Porter, 2007) propose generic episodes with three dimensions: context, contents and outcome. They also provide a semantic similarity

measure (Tecuci and Porter, 2009). Nuxoll and Laird also have evaluated a variety of metrics for determining match scores on episode retrieval (Nuxoll and Laird, 2012). Their implementation is similar to our approach: They calculate a match score based on the number of elements a cue and memory have in common and the activation level of the cue that matches the memory.

3 EVENT INDICES

We define every observable occurrence as an event, in contrast to common language, where especially extraordinary occurrences are called event. Second, we follow the definition of Zacks and Tversky (Zacks and Tversky, 2001), that an event is a segment of time with a beginning and an end. Third, we are only considering events of equal level, that means we do not consider events that contain other events. In our approach events are not organized in partonomic event hierarchies, but in *episodes*. Fourth, we index events along five dimensions (Time, Space, Protagonists, Intentionality and Causality), according to Zwaan's event-indexing model. Fifth, we add emotion as sixth dimension. Figure 1 shows how episodes group events together and how events are conceptualized.

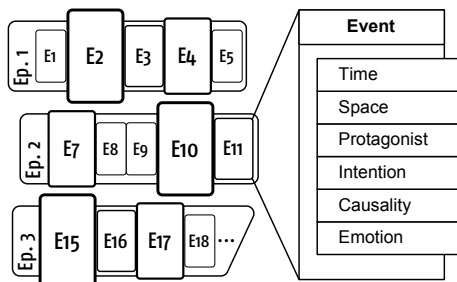


Figure 1: How episodes and events are conceptualized: Events are grouped into episodes. From the outside of episodes only events with a strong emotional impact are visible, e.g. event E2. The enlarged event E11 shows the six indices. Episode Ep. 3 is the current episode to which new events can be added (Rabe and Wachsmuth, 2012).

Independent of the six dimensions a memorized event can contain information related to the situation that it represents, we call that Payload. But in contrast to the indices the additional information can not be accessed directly.

Time is the recorded system time at the moment the event begins, and at the moment the event ends. This information is used to get the duration of the event, which is also stored. Further possible applications are to get time of the day the event occurred, and to order

events in time.

Space is represented in the virtual world coordinates, but indexed along named places to which the coordinates have been mapped. Similar to humans, who normally do not tend to memorize places in GPS coordinates, the agent memorizes the name of the place an event happened. All places are also represented using a undirected weighted location graph, that means the agents knows which places are connected and how far they are apart. Also the agent's knowledge about the places is accessible by names and not by coordinates. Nevertheless we keep the coordinates for further analysis.

Protagonists stores named representation of the individuals present during an event. This can be the agent ('I') himself and any known and named visitors. This enables the agent to remember earlier occurrences with the same visitors. In our current setup the agent can recognize only one visitor in the Virtual Environment.

Intention represents the intention the agent has during the current event. Since the agent is based on a BDI cognitive architecture the agent next actions are based on his intention. In our BDI-architecture it is the name of the plan the agent is currently pursuing. All of names the agent's plans are composed of two parts, an action and a token separated by a dash, e.g. `smalltalk-greet`, `explain-navigation`, `follow-visitor`, `show-city_hall`. All actions can be categorized into either watching, guiding, explaining, or small-talking.

Causality is what has lead to the event. This can be either a named percept of the agent (e.g. a request stated by the visitor to show a certain building), or a named action the agent performed (e.g. completely leading the way to a certain building). Similarly to the intention names causes' names consist of three parts, an actor, an action, and a token each separated by a dash, e.g. `visitor-focused-jazzclub`, `guide-proposed-city_overview`, `visitor-reached-marketplace`.

Emotion contains the current emotional state of the agent (in PAD space) and may also contain the emotional impulse the agent receives. It is our addition to the original event-indexing model, since the emotion may include clues for different things: If the guide receives good feedback, his emotional system rewards him and he is happy. If he receives negative feedback, his emotional system dampens his mood and he is sad. In further similar occurrences the guide will tend to redo the things that made him happy.

4 EVENT METRIC

Our event metric is a **metric space**, an ordered pair (M, d) where M is a set of events and d is a metric on M , i.e., a distance function

$$d: M \times M \rightarrow \mathbb{R} \quad (1)$$

We designed the distance function to not produce results greater than 1. All distance functions are equal weighted. This means we added the following constraint:

$$d: M \times M \rightarrow [0, 1] \quad (2)$$

Now we define I as a set of all indices:

$$I = \{t, s, p, i, c, e\} \quad (3)$$

Consider an event e , which is indexed along the six dimensions of time, space, protagonists, intentionality, causality and emotion:

$$e = (e_t, e_s, e_p, e_i, e_c, e_e) \quad (4)$$

The **naive distance** between two events e and f is computed as the sum of the discrete distances of the individual indices divided by the cardinality of I :

$$d_n(e, f) = \frac{1}{|I|} \sum_{i \in I} d_d(e_i, f_i) \quad (5)$$

Here we use the **discrete distance** function d_d , that means we only compare if the indices are equal or not in each dimension:

$$d_d(x, y) = \begin{cases} 0 & \text{if } x = y \\ 1 & \text{if } x \neq y \end{cases} \quad (6)$$

The minimal naive distance is 0 (all indices are alike) and the maximal is 1 (none are alike).

To enhance the event metric, we define distinct distances for each index which gives us the overall **index-distinct distance**:

$$d_i(x, y) = \frac{1}{6} (d_{td}(x_t, y_t) + d_{wp}(x_s, y_s) + d_d(x_p, y_p) + d_{in}(x_i, y_i) + d_{ca}(x_c, y_c) + d_{md}(x_e, y_e)) \quad (7)$$

We compare events on the time index using the duration of the events. To normalize the distance we divide the duration difference by the longest duration of all events. This gives us the **time-durations distance**:

$$d_{td}(x, y) = \frac{|x_t - y_t|}{max_{duration}} \quad (8)$$

We compare events on the space index using a **waypoint distance**. Due to the scenario the streets and places form an undirected weighted graph. We calculate the shortest path using Dijkstra's algorithm

and divide the length of the result by the length of the longest path of our scenario:

$$d_{wp}(x, y) = \frac{d_{Dijkstra}(x_s, y_s)}{max_{distance}} \quad (9)$$

For a different scenario it might be applicable to use the Euclidean distance instead.

We compare events on the protagonist index using the discrete distance function as **protagonist distance**. In future work we think about incorporating a relationship based distance using work from (Matar and Wachsmuth, 2012) who are building a person memory for the virtual agent.

The **intention distance** is the sum of a category based distance for the action part of the intention and the discrete distance for the token part. The category based distance is an extended discrete distance, where if an action x is not equal to y but in the same action category C as y the distance is smaller than if x and y would be in different action categories:

$$d_c(x, y) = \begin{cases} 0 & \text{if } x = y \\ 0.5 & \text{if } x \neq y \wedge x, y \in C \\ 1 & \text{if } x \neq y \wedge x \in C \wedge y \notin C \end{cases} \quad (10)$$

With this the intention distance is:

$$d_{in} = \frac{1}{2} (d_c(x_{action}, y_{action}) + d_d(x_{token}, y_{token})) \quad (11)$$

We compare events on the causality index using a category based distance for the action part of the intention and the discrete distance for the actor and token parts. Similar to the intention distance the **causality distance** is:

$$d_{ca} = \frac{1}{3} (d_d(x_{actor}, y_{actor}) + d_c(x_{action}, y_{action}) + d_d(x_{token}, y_{token})) \quad (12)$$

Finally, the agent's emotional state is mapped to a 2.5 dimensional representation of the PAD space (Becker-Asano and Wachsmuth, 2010). The maximal expansion from $(-100, -100, -100)$ to $(100, 100, 100)$ is $max_{PAD} = 200\sqrt{3} \approx 346.41$. We use this to calculate a normalized Euclidean distance, the **mood distance**:

$$d_{md} = \frac{\|x - y\|_2}{max_{PAD}} \quad (13)$$

5 EPISODE METRIC

To determine which episodes are alike we have a look at how many events of the episodes to compare are similar. Therefore we define an episode metric (M, d) (similar to the event metric) where M is a set

of episodes which consist of events. Considering two episodes E and F we define the **episode distance** as

$$d_{ep}(E, F) = 1 - \frac{\sum_{e \in E} \left(\sum_{f \in F} d_b(e, f) \right)}{|E| \cdot |F|}, \quad (14)$$

where d_b is a discrete distance which is 0 if the result of the index-distinct distance is below a certain boundary b :

$$d_b(e, f) = \begin{cases} 0 & \text{if } d_i(e, f) < b \\ 1 & \text{if } d_i(e, f) \geq b \end{cases} \quad (15)$$

This means we compare every event e of episode E to every event f in episode F and count the number of matches. This number is then divided by the product of the number of all elements in E and F and then subtracted from 1. A smaller result means that the episodes are more similar.

Note that depending on how b is selected it can be sufficient if only one or two indices of the events to compare are similar enough. If e.g. $b = 1 - \frac{2}{6}$ two events would match if two indices are alike.

6 CONCLUSIONS & OUTLOOK

We have introduced an event metric and an episode metric which our virtual guide employs to select memories matching to the current situation. He utilizes this memories in his decision process what to do next.

Next steps of our work embrace collecting more data, that means that our agent has to give many more tours to accumulate rich memories. With that we plan to evaluate the performance of the memory system and measure the quality of the chosen actions.

ACKNOWLEDGEMENTS

This project is supported by the Cognitive Interaction Technology Excellence Center (CITEC). We gratefully acknowledge the MPI for Biological Cybernetics for providing us with Virtual Tübingen.

REFERENCES

Aamodt, A. and Plaza, E. (1994). Case-based reasoning: Foundational issues, methodological variations, and system approaches. *AI Communications*, 7(1):39–59.

- Allen, P. A., Kaut, K. P., and Lord, R. R. (2008). Emotion and episodic memory. In Dere, E., Easton, A., Nadel, L., and Huston, J. P., editors, *Handbook of Episodic Memory*, pages 115–132. Elsevier, Amsterdam, Netherlands.
- Becker, C., Leßmann, N., Kopp, S., and Wachsmuth, I. (2006). Connecting feelings and thoughts - modeling the interaction of emotion and cognition in embodied agents. In Fum, D., Del Missier, F., and Stocco, A., editors, *Proceedings of ICCM06*, pages 32–37, Trieste, Italy. Edizioni Goliardiche.
- Becker-Asano, C. and Wachsmuth, I. (2010). Affective computing with primary and secondary emotions in a virtual human. *Autonomous Agents and Multi-Agent Systems*, 20:32–49.
- Bergmann, R. (2002). *Experience Management*, volume 2432 of *LNAI*. Springer, Berlin, Germany.
- Kolodner, J. L. (1993). *Case-Based Reasoning*. Morgan Kaufmann, San Mateo, CA.
- Leßmann, N., Kopp, S., and Wachsmuth, I. (2006). Situated interaction with a virtual human – perception, action, and cognition. In Rickheit, G. and Wachsmuth, I., editors, *Situated Communication*, pages 287–323. Mouton de Gruyter, Berlin, Germany.
- Mattar, N. and Wachsmuth, I. (2012). Who are you? on the acquisition of information about people for an agent that remembers. In Filipe, J. and Fred, A., editors, *Proceedings of ICAART 2012*, volume 2, pages 98–105, Vilamoura, Portugal. SciTePress.
- Nuxoll, A. M. and Laird, J. E. (2012). Enhancing intelligent agents with episodic memory. *Cognitive Systems Research*, 17–18:34–48.
- Rabe, F. and Wachsmuth, I. (2012). Cognitively motivated episodic memory for a virtual guide. In Filipe, J. and Fred, A., editors, *Proceedings of ICAART 2012*, volume 1, pages 524–527, Vilamoura, Portugal. SciTePress.
- Tecuci, D. and Porter, B. (2009). Memory based goal schema recognition. In Lane, H. C. and Guesgen, H. W., editors, *Proceedings of FLAIRS-09*, Menlo Park, CA. AAAI Press.
- Tecuci, D. G. and Porter, B. W. (2007). A Generic Memory Module for Events. In Wilson, D. C. and Sutcliffe, G. C. J., editors, *Proceedings of FLAIRS-07*, pages 152–157, Menlo Park, California. The AAAI Press.
- Tulving, E. (1972). Episodic and semantic memory. In Tulving, E. and Donaldson, W., editors, *Organization of Memory*, pages 381–403. Academic Press, New York.
- Zacks, J. M. and Tversky, B. (2001). Event structure in perception and conception. *Psychological Bulletin*, 127(1):3–21.
- Zwaan, R. A., Langston, M. C., and Graesser, A. C. (1995). The construction of situation models in narrative comprehension: An event-indexing model. *Psychological Science*, 6(5):292–297.