

# COGNITIVELY MOTIVATED EPISODIC MEMORY FOR A VIRTUAL GUIDE

Felix Rabe and Ipke Wachsmuth

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

**Keywords:** Episodic memory, Virtual reality, Guidance, Event index, Embodied conversational agent.

**Abstract:** This paper describes how the guiding capabilities of a virtual agent with a belief – desire – intention cognitive architecture can be enhanced by adding an episodic memory. We describe how the theories of episodic memory and event segmentation can be applied to the architecture of our virtual agent *Max*, and how to create an index according to the event indexing model. Having memories of past experiences will enable our agent to have improved plans how to react in future interaction.

## 1 INTRODUCTION

Past experiences influence all of our actions and let us have an expectation of what might happen next, so that we are able to plan well. Especially 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. To improve the behavior of virtual agents in interaction with humans the agents need episodic memory, too.

In the past we have introduced the virtual humanoid agent *Max* (Leßmann et al., 2006) who already has lots of skills and is based on a belief – desire – intention (BDI) cognitive architecture. He has proven a helpful interaction partner, e.g. in assisting a human in construction tasks. In a different scenario *Max* is “working” as a museums guide, where he stands in a museum, conducts small talk with visitors and explains sights. But up to now *Max* has no memory of his own actions and his experiences. He can only memorize some facts about persons visiting him. With the cognitively motivated episodic memory this is to be changed.

We have conceived a virtual guide scenario, where *Max* can utilize his new skill in interaction. We picked a virtual reality (VR) scenario, since VR is where humans and machines meet and can interact at eye level. The episodic memory system will be especially useful when the agent is guiding a visitor through a virtual environment. The main focus will be on the agent memorizing his own actions, the interaction with the human visitor, and the improvement of the guidance by assisting the human.

## 2 RELATED WORK

### 2.1 Psychological Background

Endel Tulving suggested that there are two distinct types of declarative long-term memory: Episodic and semantic memory. Semantic memory is factual knowledge about the world. Episodic memory deals with temporally dated episodes or events, and temporal-spatial relations among these events. Every event, or “item in episodic memory”, as it is called by Tulving, is a more or less faithful record of a person’s experience of an occurrence, and can include the perceptible properties of that moment (Tulving, 1972).

Another theory we build on is the Event-Indexing Model (Zwaan et al., 1995). It describes how readers of short stories construct a model of the situation in the text. As the readers understand what is happening in the story they update the model along five indices: Time, Space, Causality, Intentionality and Protagonist. These dimensions store the answers to the questions of what happened when, where, why and how, and who was involved. Causality is what lead to the current event and Intentionality are the intentions the protagonists had during the event. Zwaan and colleagues suggested that events can be compared along these indices: The more indices differ, the more distinct are the events.

In Zacks and Tversky’s Event Segmentation Theory (Zacks and Tversky, 2001), 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.” Second, events can be organized in partonomic

hierarchies, which means they may span very long and very short periods of time. A long event can embrace some short events. In further work Zacks et al. found evidence that when human perception does not match the internal prediction, an event boundary is perceived (Zacks et al., 2007).

Allen, Kaut, and Lord 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 (Allen et al., 2008).

## 2.2 Computational Episodic Memory

In computer science, Ho built an autobiographic memory system for an agent to locate resources encountered before and used it in further work to enhance virtual characters, so that they were able to talk about personal past experiences (Ho and Dautenhahn, 2008).

Tecuci and Porter created a generic memory module for events, where a generic episode has three dimensions: context, contents, and outcome. Context is the general setting in which an episode happened, contents is the ordered set of events, that make up the episode, and outcome is an evaluation of the episodes effect (Tecuci and Porter, 2007).

Nuxoll and Laird extended a cognitive architecture (Soar) with episodic memory. They showed that an autonomous agent, a virtual tank, performs better if it can use episodic memory for reasoning (Nuxoll and Laird, 2007).

Brom et al. proposed a that virtual characters need a “full episodic memory” and implemented it for a non-player character of a computer role-playing game. The system allows the reconstruction of the personal story of the character (Brom et al., 2007).

Sieber and Krenn proposed an episodic memory for a companion to enhance its dialogue capabilities. The episodes are stored as RDF triples (Sieber and Krenn, 2010).

## 2.3 Applications

In the field of virtual humanoid agents, there are a few scenarios where the agent is a guide.

Theune et al. presented a virtual tour guide with focus on conversational interactions. Their guide served as direction giving guide in a virtual museum. They focused on dialogue management and how the directions can be given with language and gesture (Theune et al., 2007).

Jan et al. presented a virtual guide in Second Life. Their agent can be asked about items on an island and

can navigate a user to different spots by walking in front of him (Jan et al., 2009).

Lim and Aylett built an affective mobile guide, which has an emotion memory. The stories the guide tells about the environment are influenced by emotions he experienced beforehand (Lim et al., 2009).

## 3 MEMORY MODEL

We will explain the cognitive architecture of Max and how we extend it with memory (see fig. 1).

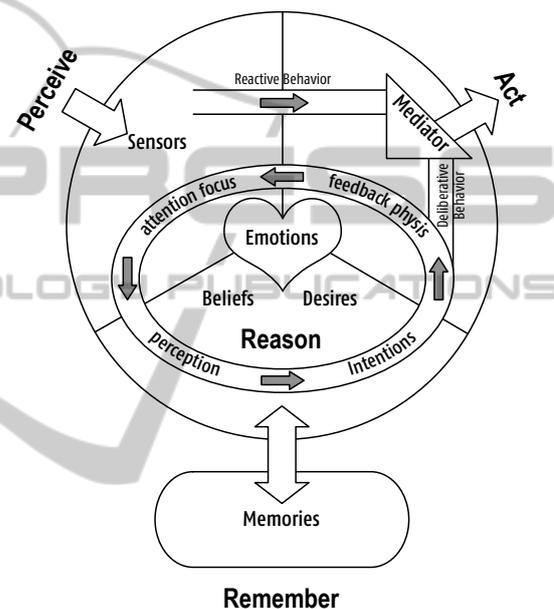


Figure 1: This figure shows the cognitive architecture of Max, extended with memory.

In the virtual world Max perceives dialogue input and other actions of the user, e.g. navigation to a location and focusing on a specific virtual item. Max also knows where he and the user are.

Due to what Max perceives he reasons and deliberately chooses his actions. The perception process also delegates the input to the memory which has an “unconscious” mechanism that segments the input into events, following the idea stated by Zacks et al. that event segmentation is a spontaneous outcome. During reasoning Max can choose to actively remember episodes and events from his memory. Also there is the possibility that his memory comes up with an episode or event that is similar to the current situation, Max then has a spontaneous inspiration.

### 3.1 Events and Episodes

In common language *event* is a board concept. To nar-

row it down for our purposes, first of all an event is an observable occurrence, and for us any observable occurrence is an event, not only extraordinary occurrences. Second, we follow the definition of Zacks and Tversky, that an event is a segment of time, that an event has a beginning and an end. Third, at the moment we are only considering events of equal level, that means we don not consider events that contain other events. So, our events are not (yet) organized in partonomic event hierarchies, but in episodes. Figure 2 shows how we organize events and episodes.

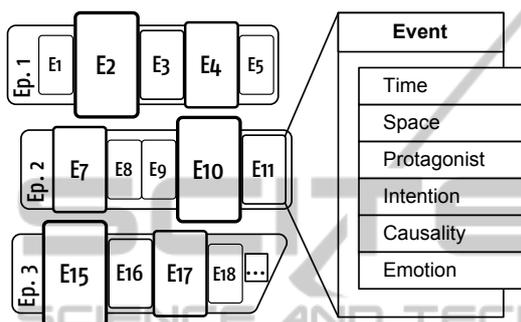


Figure 2: This figure shows how episodes and events are conceptualized. From the outside of episodes only events with a strong emotional impact are visible, e.g. event E2. The enlarged event E11 shows the six indices. Episode Ep. 3 is the current episode to which new events can be added.

Max's events are indexed according to Zwaan's event-indexing model along five indices. But since Allen et al. state that emotion is also an important cue for episodic memory we extended the model with a sixth dimension.

**Time.** Time is the recorded system time at the moment the event begins, at the moment the event ends. This information can be used to get either the length of the event and to get time of the day the event occurred.

**Space.** Space is represented in the virtual world coordinates. The agent can map these to named places.

**Protagonist.** Can be the agent ('I') himself and any known and named visitor.

**Intention.** This is the intention the agent has in the current event. It is a reference to the plan the agent is currently pursuing.

**Causality.** This is what has lead to the event. This can be either a named action the agent performed, or a named percept of the agent.

**Emotion.** This is the current emotional state of the agent as well as the emotional impulse the agent receives.

**Additional Payload.** Furthermore we add additional "payload" to the event. This may include perceptual stimuli of the agent, e.g. a snapshot of what he heard, saw, and thought.

Events are compared along their indices, the more indices are similar the more two events are alike. This is useful in finding related events. If the agent tries to remember something he first will only be able to access events with a strong emotional impact, since these are the events which function as key events to episodes. Should the agent not be satisfied with the recovered memories, he tries to remember events that may be hidden in episodes, since the emotional impact is not that strong. But he should be able to access similar episodes relatively fast, since the indices are similar and therefore near.

## 4 GUIDE SCENARIO

Our scenario is accompanying and guiding a non-local person through a large and complex virtual environment: A visitor encounters Max in a CAVE-like virtual reality of Virtual Tübingen, a virtual city realistically modeled after the historic center of Tübingen. It has been originally developed at the Max Planck Institute for Biological Cybernetics as a naturalistic, controllable environment for investigating human spatial cognition (van Veen et al., 1998). The model covers an area of 500x150 m<sup>2</sup>, it has 15 streets which are mostly curved and varying in width, and about 200 houses with different, photorealistic textures. Max has knowledge about 25 sights and 50 additional waypoints, he uses for his orientation.

After greeting the visitor Max introduces himself and offers to give a guided tour or to just accompany the visitor in exploration of the town.

If the visitor chooses to explore the city on his own Max just accompanies him and watches the visitors actions. Should Max remember an episode which fits to some of the visitors actions, Max will ask if he may provide guidance to additional sights he remembers.

If the visitor chooses to be guided Max has the goal to give a tour and takes over the steering control. He remembers a "master" episode where he has given a tour before, he will visit the same places and give the same information. If he is not interrupted at all in the end the current episode will be similar to the master episode. But since Max can be interrupted at a certain point he may change the tour due to the wishes of the visitor. He will then try to adapt with the help of different tours he has given before and with the knowledge he has. For example, if the visitor utters that he is very interested in seeing old churches, then

Max tries to remember an episode where he showed churches. If he does not remember anything he will plan a new tour on his own, since he knows that there are churches. In the end Max then has memorized a new episode with lots of churches in it.

On the end of a tour Max can be asked to give a summary of the tour. He will then remember the whole episode and tell the visitor what places were visited and which sights have been explained. Furthermore the visitor can ask Max why he has shown a specific sight. Due to the Causation and Intentionality indices of the events Max can generate adequate answers.

## 5 OUTLOOK

We have introduced a cognitively motivated episodic memory system to be integrated in a BDI-based agent architecture. Future work is in evaluating the quality and the benefit of the episodic memories with respect to Max's event segmentation capabilities.

One possible way would be to video tape Max guiding a person through Virtual Tübingen. Afterwards the recording is played back to a certain number of probands who segment the video clip into events. These events than can be compared to the events Max has segmented on his own and from that we can see how human like Max event perception is.

Another possibility would be having a human guiding a visitor through Virtual Tübingen with Max only watching the tour without interfering. Afterwards the visitor is asked to retell the tour in form of events which is compared to Max's summary.

## 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

- 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.
- Brom, C., Peškova, K., and Lukavský, J. (2007). What does your actor remember? towards characters with a full episodic memory. In Cavazza, M. and Donikian, S., editors, *Virtual Storytelling: Using Virtual Reality Technologies for Storytelling (ICVS)*, LNCS 4871, pages 89–101, Berlin. Springer.
- Ho, W. C. and Dautenhahn, K. (2008). Towards a narrative mind: The creation of coherent life stories for believable virtual agents. In Prendinger, H., Lester, J., and Ishizuka, M., editors, *Intelligent Virtual Agents*, LNAI 5208, pages 59–72, Berlin. Springer.
- Jan, D., Roque, A., Leuski, A., Morie, J., and Traum, D. R. (2009). A virtual tour guide for virtual worlds. In Ruttkay, Z., Kipp, M., Nijholt, A., and Vilhjálmsón, H. H., editors, *Intelligent Virtual Agents*, LNAI 5773, pages 372–378, Berlin. Springer.
- 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.
- Lim, M., Aylett, R., Ho, W. C., Enz, S., and Vargas, P. (2009). A socially-aware memory for companion agents. In Ruttkay, Z., Kipp, M., Nijholt, A., and Vilhjálmsón, H. H., editors, *Intelligent Virtual Agents*, LNCS 5773, pages 20–26, Berlin. Springer.
- Nuxoll, A. M. and Laird, J. E. (2007). Extending cognitive architecture with episodic memory. In Holte, R. C. and Howe, A., editors, *Proceedings of the 22<sup>nd</sup> AAAI Conference on AI*, pages 1560–1565, Menlo Park, CA. AAAI Press.
- Sieber, G. and Krenn, B. (2010). Episodic memory for companion dialogue. In *Proceedings of the 2010 Workshop on Companionable Dialogue Systems*, pages 1–6, Uppsala, Sweden. ACL.
- 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, CA. AAAI Press.
- Theune, M., Hofs, D., and van Kessel, M. (2007). The virtual guide: A direction giving embodied conversational agent. In *INTER\_SPEECH-2007*, pages 2197–2200, Antwerp, Belgium.
- Tulving, E. (1972). Episodic and semantic memory. In Tulving, E. and Donaldson, W., editors, *Organization of Memory*, pages 381–403. Academic Press, New York, NY.
- van Veen, H.-J. A. H. C., Distler, H. K., Braun, S. J., and Bülthoff, H. H. (1998). Navigating through a virtual city: Using virtual reality technology to study human action and perception. *Future Generation Computer Systems*, 14(3-4):231–242.
- Zacks, J. M., Speer, N. K., Swallow, K. M., Braver, T. S., and Reynolds, J. R. (2007). Event perception: A mind–brain perspective. *Psychological Bulletin*, 133(2):273–293.
- 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.