Enhancing Human Computer Interaction with Episodic Memory in a Virtual Guide

Felix Rabe and Ipke Wachsmuth

Artificial Intelligence Group, Bielefeld University
Universitätsstraße 25, 33615 Bielefeld, Germany
[frabe|ipke]@techfak.uni-bielefeld.de

Abstract. Have you ever found yourself in front of a computer and asking it aloud: "Why?" We have constructed a cognitively motivated episodic memory system that enables a virtual guide to respond to this question. The guide, a virtual agent based on a belief–desire–intention (BDI) architecture, is employed in a Virtual Reality (VR) scenario where he accompanies a human visitor on a tour through a city. In this paper we explain how the agents memorizes events and episodes according to an event-indexing model and how the interaction is enhanced by using these memories. We argue that due to the cognitively motivated nature of the event-indexing model every interaction situation can be described, memorized, recalled and explained by the agent.

Keywords: Episodic Memory, Event Indexing, Virtual Guide

1 Introduction

Memories are not just the autobiographical log of yourself, they are the significant foundation that enables humans to plan next actions and let us build an expectation of what might happen in the future. Especially in interaction with other human beings we rely on past episodes with other persons, we adapt to situations based on what we experience and store in episodic memory.

Our work is centered around the virtual humanoid agent *Max*, cf. [5,1]. Max has already 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 in construction tasks or "working" as a museum guide, where he is present on a large screen, conducts small talk with visitors and explains sights [7]. But up to now Max has no memory of his own actions and his experiences.

Therefore we have designed a cognitively motivated episodic memory system for Max. We have also conceived a virtual guide scenario, where Max can utilize his new skill in interaction. The episodic memory system is employed when the agent is guiding a visitor through a virtual environment. The main function is memorizing the agent's actions and interactions with the human visitor. Improved guiding recommendations are generated from similar previous episodes.

This paper is an extended version of [8]. It gives details on the guide scenario and explains how the interactions with an artificial agent are influenced by an episodic memory.

2 Related Work

The concept of episodic memory was first coined by psychologist Endel Tulving in 1972 [11]. Tulving suggested that there are two distinct types of declarative long-term memory: Episodic and semantic memory. While the latter is factual knowledge about the world, episodic memory 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, and may include the perceptible properties of that moment.

The subject of episodic memory has also attracted interest in computer science, there have been different implementations of computational models of episodic memory. Johnson [4] built an artificial fighter pilot that makes use of episodic memory to explain itself during debriefing. Ho [3] 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. Tecuci and Porter [10] created a generic memory module for events, where a generic episode has three dimensions: context, contents, and outcome, but only used it for planning. Nuxoll and Laird [6] 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. Brom et al. [2] proposed a "full episodic memory" for a non-player character of a computer role-playing game that allows the reconstruction of the character's personal story.

The foundations for our episodic memory system are the Event-Indexing Model by Zwaan and colleagues [15] and the Event Segmentation Theory by Zacks and Tversky's [14]. The Event-Indexing Model describes how readers of short stories construct a model of the situation in the text 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. We incorporate these five dimensions to describe and index the experiences of the virtual agent. Complementary to this the Event Segmentation Theory defines an event as "a segment of time at a given location that is conceived by an observer to have a beginning and an end" [14]. In further work Zacks et al. [13] found evidence that when human perception does not match the internal prediction, an event boundary is perceived. The Event Segmentation Theory also states that events are organized in partonomic hierarchies, which means they may span very long and very short periods of time, and a long event can embrace several short ones.

3 Events and Episodes

In common language *event* is a broad concept. To narrow it down 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 [14], 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 added emotion as sixth dimension, since we may use it as a cue for importance of events. Figure 1 illustrates how episodes group events together and how events are conceptualized.
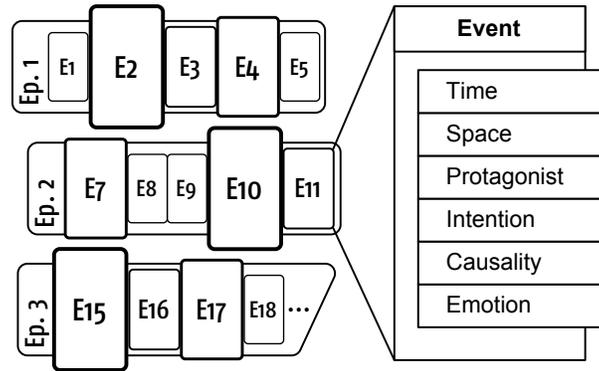

Fig. 1. How episodes and events are conceptualized (as introduced in [8]): Events are grouped into episodes. From the outside of episodes only events with a strong emotional impact are visible, e.g. event E2. The enlarged event E11 shows the six indices. Episode Ep. 3 is the current episode to which new events can be added.

3.1 Event Indices

We now give a short overview of how the individual indices are conceptualized, for a detailed explanation and the event and episode metrics we created refer to [9].

Time records the time of the event's begin and end. It also stores the duration of the event which is used for comparison.

Space holds the location of the event in virtual world coordinates. But similar to humans, who normally do not tend to memorize places in GPS coordinates, these are mapped to a named place in the world which is used for comparison.

Protagonists stores named representations of the individuals present during an event. This can be the agent ('I') himself and any known and named visitors.

Intention represents the intention the agent has during the current event. Since the agent is based on a BDI cognitive architecture the agent's next actions are based on its intention.

Causality is what has lead to the event and contains the answer to the question why. 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).

Emotion contains the current emotional state of the agent and may also contain the emotional impulse the agent receives. Events with a high emotional value are considered more significant and are more likely to be remembered.

3.2 Event Boundaries

The segmentation and memorization of events is an "unconscious" mechanism, we follow the idea stated by Zacks et al. [13] that event segmentation is a spontaneous outcome. In particular, whenever Max perceives an input, a change in the environment, or a change of his internal state, the memory system compares this to the current event and if the change is significant enough, a new event is memorized. Max's perception incorporate dialogue input and other actions of the visitor of the virtual world, e.g. navigation to a location and focusing on a specific virtual item. He also knows where he and the user are, he has knowledge of his intention, his emotional state and the actions he is performing.

4   Guide Scenario

Our scenario comprises a virtual tour guide accompanying and guiding a non-local person through a large and complex virtual environment. A visitor encounters Max in a CAVE-like virtual reality, *Virtual Tübingen*, a virtual city realistically modeled after the historic center of Tübingen in Germany (see Fig. 2).

The virtual model of Tübingen is large and complex, it covers an area of 500x150 m^2, it has 15 streets which are mostly curved and varying in width, and about 200 houses with different, photorealistic textures. Originally the model was developed at the Max Planck Institute for Biological Cybernetics as a naturalistic, controllable environment for investigating human spatial cognition [12]. We extended it by tagging places so that Max has knowledge of about 25 sights and 50 additional landmarks, which he uses as navigation points for his orientation.

In this scenario an episode is a single tour of a visitor accompanied by Max through the city. The visitor is supposed to explore the city and to look at sights of his interest. He can steer on his own using a Wii Remote controller or can ask Max to take control and guide him. After greeting the visitor, Max introduces himself and offers directly to give a guided tour or to just accompany the visitor in exploration of the town (an example dialogue is presented later on). New events are stored automatically in the memory and indexed along the six dimensions as Max and the visitor interact and move through the city together. So far, speech recognition is not provided for the visitor, he has to choose from a set of natural language sentences. Note that Max' memory is not empty at the start. It already holds some episodes, which Max could have experienced before, but in this case are pre-programmed to equip Max with some initial memories as a starting point to guide.

Fig. 2. The scenario comprises the virtual agent Max standing together with the visitor on a floating hover disk in a CAVE-like virtual reality environment. Besides wandering through the city of Virtual Tübingen and looking at sights from "normal ground perspective" the city can also be examined from a bird's eye view.

5 Episodic Memory in Interaction

For our scenario we identified four general interaction cases that make use of the episodic memory system.

An Accompanied Exploration. If the visitor chooses to explore the city on his own, Max just accompanies him (or her) and watches the visitor's actions. At any point during the exploration Max can be asked about sights encountered. Max uses his memories to suggest where to go next. These recommendations are based on what sights have been visited so far in comparison to prior episodes. Figure 3 shows an on-going episode and two related completed episodes. Although Episode 2 has a slightly different path than the on-going episode it still may be a better reference where to go next than Episode 1 based on the content of the intentionality and causality indices.

A Guided Tour. 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. If Max is not interrupted he will visit the same places and give the same information. On request by the visitor he may change the tour at a any point. He will then try to adapt to the visitor's wishes with the help of different tours he has given before and with the knowledge he has. Consider Episode 1 as the master for the on-going episode (see Fig. 3). If Max is

interrupted at stop 2 and asked for a specific sight he might remember Episode 2 where the sight would be encountered next. That episode would become the master episode and at the end Max memorizes the new episode which is a mixture of Episode 1 and Episode 2.

A Self-Explanation. If the visitor asks Max why he has recommended a certain sight or why Max has executed a specific action, Max is able to explain himself. He does this by remembering the addressed event's causality and telling it to the visitor.

A Summary. If the visitor asks for a summary Max can recall all seen sights so far from his memory and tell them in the order of visits to the visitor.

Fig. 3. Shown are two completed and one on-going episode visualized on the map of Virtual Tübingen.

5.1 Excerpts From a Tour with Max Through Virtual Tübingen

We give some examples (translated to English) of how a visitor and Max interact during a tour in our through Virtual Tübingen. First Max greets a visitor and explains the navigation:

> (MEM: Episode start)
> MAX: Now, Felix, welcome to virtual Tübingen. This model resembles the nice old town of Tübingen. Right now, we are located at the marketplace.
> (MEM: BEL Greeted)
> (GOAL: Explain navigation in VR)
> (INTENT: Find out if navigation in VR needs to be explained)

MAX: Do you know how to move about the place?
(Search MEM if CURRENT VISITOR was guided before: if YES remind CURRENT VISITOR of previous event, if NO explain navigation with Hover Disk)
VISITOR: No I don't. → (MEM: INTENT: Explain navigation)
MAX: Now, we are standing on the Hover disk together. By this we can move about the place. You can control the Hover disk by using the Wii-Mote yourself. Alternatively, I can control the Hover disk for you.
(MEM: BEL Navigation explained)
[etc.]

Then during the tour Max uses his memory to tell an interesting story:

MAX: This is the *Holzmarkt*.
(GOAL: Explain nearby sights)
(REMEMBER: Similar episode which has a strong emotional impact)
(INTENT: Tell interesting facts about Holzmarkt)
MAX: Over there you can see the old book store of *Johann Friedrich Cotta*, the most important German publisher of his time. He liked his small place right next to the market, and important authors like *Goethe*, *Schiller*, and *Hölderlin* to name only a few visited him in his shop regularly. But many foreign visitors could hardly believe that so many famous authors were coming to this place, since the market was in a really bad shape, e.g. the stairs of the *Stiftskirche* were ruins. Also the name *Holzmarkt* comes from the simple items that were sold here during that time, mostly fire wood and copper items.
VISITOR: Well, that is an interesting story. → (MEM: Positive feedback)
MAX: I can also tell you about the *Stiftskirche* while we are here. Do you want to hear about it?

Finally Max uses memory at the end of the tour to summarize which places have been visited so far and answers a question why a specific sight was visited:

VISITOR: Can you summarize what we have done so far?
(MEM: INTENT: Summarize tour on visitor request)
MAX: Ok, here we go: After I explained how you can move through Virtual Reality, we started out at the market place of Tübingen. I showed you the city hall and told you about the *Neptunbrunnen*. Afterwards we had a look at the astronomical clock and listened to the Glockenspiel of the city hall's clock. From there you went through the *Holzgasse* to the *Holzmarkt* and I showed you the *Cotta* house and the *Stiftskirche*. After it I lifted us up into the air and I gave you an overview of the whole town. Then we landed in front of the *Jazzkeller* and now we are looking at the *Krumme Brücke*.
(MEM: Summarized tour)
VISITOR: Why have we been at the *Jazzkeller*?
(REMEMBER: Event containing location *Jazzkeller*)

MAX: Because you asked me about it when we were up in the air.
VISITOR: Thank you for the great tour. Unfortunatly I have to leave now, but I'l be back. See you.
MAX: Ok. If you like we can continue from here on then. Farewell!
(MEM: BEL Bid farewell)
(MEM: Episode end)

6 Discussion

We have introduced a cognitively motivated episodic memory system which is integrated in a BDI-based agent architecture and employed in a virtual tour guide. Our scenario explains how the interaction between the virtual guide and a human visitor benefits from the use of the episodic memory. The guide can give recommendations without needing a large database of previous interactions, as he gets more knowledgeable through interaction with visitors by memorizing a new episode every time giving a tour. An added benefit of enhancing interaction lies in possibility to have the guide explain himself to the visitor, which makes the guide more comprehensible and authentic. Evaluating our approach will be subject of future work.

While it is not novel to build a system which makes use of episodic memory, the strength of our approach lies in the six event indices and their independence of a particular domain. Since these indices have been identified to describe general situations, events other than from the tour guide scenario could be memorized. Further, the fact that the described memory system is integrated within a BDI agent architecture is not essential for its application. The general requirements for a system to use our episodic memory system for enhancing human computer interaction are knowledge of the context and an agent system's capability to recognize own actions and plans.

Acknowledgements

References

1. Becker, C., Leßmann, N., Kopp, S., Wachsmuth, I.: Connecting feelings and thoughts - modeling the interaction of emotion and cognition in embodied agents. In: Fum, D., Del Missier, F., Stocco, A. (eds.) Proceedings of the Seventh International Conference on Cognitive Modeling (ICCM06). pp. 32–37. Edizioni Goliardiche, Trieste, Italy (2006)

2. Brom, C., Peškova, K., Lukavský, J.: What does your actor remember? towards characters with a full episodic memory. In: Cavazza, M., Donikian, S. (eds.) Virtual Storytelling: Using Virtual Reality Technologies for Storytelling (ICVS). LNCS, vol. 4871, pp. 89–101. Springer, Berlin/Heidelberg, Germany (2007)

3. Ho, W.C., Dautenhahn, K.: Towards a narrative mind: The creation of coherent life stories for believable virtual agents. In: Prendinger, H., Lester, J., Ishizuka, M. (eds.) Intelligent Virtual Agents. LNAI, vol. 5208, pp. 59–72. Springer, Berlin/Heidelberg, Germany (2008)

4. Johnson, W.L.: Agents that learn to explain themselves. In: Hayes-Roth, B., Korf, R.E. (eds.) Proceedings of the 12th National Conference on Artificial Intelligence (AAAI-94). vol. 2, pp. 1257–1263. AAAI Press, Menlo Park, CA, USA (1994)

5. Leßmann, N., Kopp, S., Wachsmuth, I.: Situated interaction with a virtual human – perception, action, and cognition. In: Rickheit, G., Wachsmuth, I. (eds.) Situated Communication, Trends in Linguistics, vol. 166, pp. 287–323. Mouton de Gruyter, Berlin, Germany (2006)

6. Nuxoll, A.M., Laird, J.E.: Extending cognitive architecture with episodic memory. In: Holte, R.C., Howe, A. (eds.) Proceedings of the Twenty-Second AAAI Conference on Artificial Intelligence. pp. 1560–1565. AAAI Press, Menlo Park, CA, USA (2007)

7. Pfeiffer, T., Liguda, C., Wachsmuth, I., Stein, S.: Living with a virtual agent: Seven years with an embodied conversational agent at the Heinz Nixdorf MuseumsForum. In: Barbieri, S., Scott, K., Ciolfi, L. (eds.) Proceedings of the International Conference Re-Thinking Technology in Museums 2011 – Emerging Experiences. pp. 121–131. thinkk creative & the University of Limerick, Limerick, Ireland (2011)

8. Rabe, F., Wachsmuth, I.: Cognitively motivated episodic memory for a virtual guide. In: Filipe, J., Fred, A. (eds.) ICAART 2012 – Proceedings of the 4th International Conference on Agents and Artificial Intelligence. vol. 1, pp. 524–527. SciTePress, Vilamoura, Portugal (2012)

9. Rabe, F., Wachsmuth, I.: An event metric and an episode metric for a virtual guide. In: Filipe, J., Fred, A. (eds.) ICAART 2013 – Proceedings of the 5th International Conference on Agents and Artifical Intelligence. vol. 2, pp. 543–546. SciTePress, Barcelona, Spain (2013)

10. Tecuci, D.G., Porter, B.W.: A generic memory module for events. In: Wilson, D.C., Sutcliffe, G.C.J. (eds.) Proceedings to the 20th Florida Artificial Intelligence Research Society Conference (FLAIRS-07). pp. 152–157. AAAI Press, Menlo Park, CA, USA (2007)

11. Tulving, E.: Episodic and semantic memory. In: Tulving, E., Donaldson, W. (eds.) Organization of Memory, pp. 381–403. Academic Press, New York, NY (1972)

12. van Veen, H.J.A.H.C., Distler, H.K., Braun, S.J., Bülthoff, H.H.: Navigating through a virtual city: Using virtual reality technology to study human action and perception. Future Generation Computer Systems 14(3–4), 231–242 (1998)

13. Zacks, J.M., Speer, N.K., Swallow, K.M., Braver, T.S., Reynolds, J.R.: Event perception: A mind–brain perspective. Psychological Bulletin 133(2), 273–293 (2007)

14. Zacks, J.M., Tversky, B.: Event structure in perception and conception. Psychological Bulletin 127(1), 3–21 (2001)

15. Zwaan, R.A., Langston, M.C., Graesser, A.C.: The construction of situation models in narrative comprehension: An event-indexing model. Psychological Science 6(5), 292–297 (1995)