

I Know What I Did Last Summer: Autobiographic Memory in Synthetic Characters

João Dias¹, Wan Ching Ho², Thurid Vogt³, Nathalie Beeckman⁴, Ana Paiva¹,
Elisabeth André³

¹ INESC-ID and IST, Tagus Park
Av. Prof. Cavaco Silva, 2780-990 Porto Salvo, Portugal

² University of Hertfordshire
School of Computer Science
Hatfield, Hertfordshire, AL10 9AB, UK

³ University of Augsburg
Multimedia concepts and applications
Eichleitnerstr. 30, 86159 Augsburg, Germany

⁴ University of Utrecht
Department of Information and Computing Sciences
Princetonplein 5, 3584 CC Utrecht, The Netherlands

Abstract. According to traditional animators, the art of building believable characters resides in the ability to successfully portray a character’s behaviour as the result of its internal emotions, intentions and thoughts. Following this direction, we want our agents to be able to explicitly talk about their internal thoughts and report their personal past experiences. In order to achieve it, we look at a specific type of episodic long term memory. This paper describes the integration of Autobiographic Memory into FAtiMA, an emotional agent architecture that generates emotions from a subjective appraisal of events.

1 Introduction

The concept of believability [2] in synthetic characters has been (and is still being) widely studied since traditional animators first started creating animated characters. Thomas and Johnston, creators of the well-known Disney characters, started to explore emotional expressivity and personality to achieve believability. They identified three important requisites in order to express emotions. One of them was that ”emotions affect the reasoning process and consequences should be noticeable in the actions of the characters” [17]. Another concept also adopted by researchers (e.g. Blumberg’s work [3]) to build believable characters is Daniel Dennett’s intentional stance [7]. According to the intentional stance, if we can perceive a creature’s beliefs and desires we can predict what will be its actions.

These two complementary notions suggest that if a viewer can predict a character’s actions based on his perception of the character’s emotional state, beliefs and desires, then it will become more believable to the viewer, thus allowing the suspension of disbelief. The same applies if, working backwards one can understand the actions of a given character as being influenced by its internal state.

For this reason, great part of the art in building believable characters resides in the ability to successfully portray this link between the character's internal thoughts and its external actions. The work presented in this paper tries to explore the following: making it easier for the viewer to infer the character's intentions and emotional state by having the character explicitly talking about them and report its past experiences (as we humans do to close friends). In order to achieve this, the character must have some kind of awareness over his internal thoughts and emotions, thus allowing it not only to remember past events but also what he felt at past time, and what he was planning to do.

Therefore, the main question we will be dealing is: how can we give an agent the ability to report what happened in a past experience, including his personal subjective experience? Our answer lies in a specific type of long term memory, strongly associated with the concept of self. This type of memory is called autobiographic memory. This paper describes the integration of autobiographic memory into FAtiMA[9], an emotional agent architecture based in OCC Appraisal Theory. The autobiographical memory together with the agent's ability to generate emotions from a subjective appraisal of events is used to generate summary descriptions of what happened in a past episode of the life of our characters. This summary provides the user with information about the agent's personal experience. We hope that this feature will make the user create a stronger empathic[6, 16] relation with the agent and make the agent more believable.

The rest of this paper is organized as follows: We start by an analysis of autobiographic memory, and its relation with personal experiences. The next section describes how the Autobiographic Memory is implemented and integrated into FAtiMA. Afterwards we depict the application of the resulting features into an interactive system and its effect in the interaction with the users. The last section presents some final remarks and future work.

2 Autobiographic Memory

In Psychology, autobiographic memory is a specific kind of episodic memory that contains significant and meaningful personal experiences for a human being [15]. It serves important function in providing the basis for social interaction, maintenance of a dynamic self-concept as well as the representation of the meaning of concepts [4].

2.1 Memory and Emotion

It has been widely acknowledged, in research literatures about memory, that events associated with emotional experiences partly constitute highly available memory. Psychologists and cognitive scientists also propose that when experiencing an event with emotion, humans' cognitive system is more fully engaged to processing that event, in comparison with processing of events which are not associated with emotional experience. This view can be further elaborated in that

high level emotionally intense events may also be related with high frequency of rehearsal – making a piece of memory more generally available for retrieval.

Personally significant events which are directly involved in the self memory structure, like first time experiences, can bring stronger impacts to humans' lives by creating a pre-existing knowledge structure for the rest of similar events [4]. These life events, together with events with emotions, indicate that central knowledge structures relating to the self have been employed in representing autobiographical memory.

2.2 Autobiographic Agents

Theoretically, an autobiographic agent is an embodied agent which dynamically reconstructs its individual history (autobiography) during its life-time [5]. This individual history helps autobiographic agents to develop individualised social relationships and to form communications, which are characteristic of social intelligence, and it may also lead to more appealing and human-like engaging interactions, making them more pleasant and acceptable to humans.

Different types of computational memory architectures for Artificial Life autobiographic agents have been developed and experimentally evaluated in our previous research work (e.g. [11]). These architectures include typical human memory modules which are commonly acknowledged in Psychology: short-term, long-term, and positively and negatively categorised memories. Agents embedded with these computational autobiographic memories outperform Purely Reactive agents that do not remember past experiences in surviving in both static and dynamic environments.

In the paradigm of developing synthetic agent architectures, we proposed that 1) knowledge representations in the computational autobiographic memory can be based on general episodes that agents have experienced and 2) goal structure, emotion, and attention processes, support and are influenced by the autobiographic knowledge [12]. Autobiographic knowledge may also support a long-term development and learning in synthetic agents as they gain new experience from acting in each new situation.

3 Integrating Autobiographic Memory in FAtiMA

This section describes the integration of Autobiographic Memory (AM) into FAtiMA and how the subjective appraisal of events helps in storing the agent's personal experience into memory. In order for the reader to understand this section we will provide a brief overview of FAtiMA. For a more detailed description please refer to [9, 8].

3.1 FAtiMA Overview

FAtiMA (**F**ear**N**ot! **A**ffective **M**ind **A**rchitecture) is an Agent Architecture where emotions and personality take a central role in influencing behaviour. The con-

cept of emotion used stems from OCC cognitive theory of emotions, where emotions are defined as valenced (good or bad) reactions to events. The assessment of this relationship between events and the character’s emotions is called the appraisal process. This appraisal is clearly subjective and is made regarding the agent’s goals, standards and attitudes. The same event can be positive for one character and negative for another.

The architecture incorporates two distinct levels in both appraisal and action-selection mechanisms. The first level (reactive level) provides a fast mechanism to appraise and react to a given event, and generates reactive behaviours such as crying when very distressed (which cannot be considered as planned). The deliberative level takes longer to react but allows a more sequentially complex and goal-oriented behaviour.

Whenever an external event or action is perceived, the reactive appraisal process tries to match the event with a set of predefined emotional reaction rules. These rules represent the character’s standards and attitudes and assess how generic events are appraised by defining values for OCC’s appraisal variables: desirability of the event, desirability for other (used when the event references another character other than self), praiseworthiness of the action. These variable are then used according to OCC to generate a wide range of emotions from Joy and Distress to Pity and Anger. Reactive behaviour is defined by action rules: each contains a set of preconditions that must be true to execute the action and an emotion that will trigger the action tendency.

In addition, when the event is appraised, the deliberative level checks if any goal has become active, and if so, an intention to achieve the goals’ success conditions is created generating initial Hope and Fear emotions. The agent then tries to build a plan to achieve the goal and updates such emotions accordingly. This is done using a continuous planner [1] that was extended to include emotion-focused coping[14] strategies and generates Partially Ordered Plans. Instead of executing a plan, the agent may decide to give up a plan or goal because he is scared or he is not confident it will work. This type of coping is called emotion-focused coping. If the agent succeeds or fails in achieving a goal the corresponding emotions (ex: Satisfaction, Disappointment) are also generated.

3.2 AM structure

In FAtiMA, AM is organized as a set of independent episodes. Each episode stores several actions or events that happened at a particular location and at a given time. Inspired from the research on narrative structure in life stories for humans [13], the knowledge structure representing each episode has three main components: **Abstract**, **Narrative** and **Evaluation**. Within an episode, there can be more than one important *cause-effect* action. Cause-effect actions represent actions that eventually bring a considerable impact to the agent’s emotional states. Therefore, actions that do not have an emotional impact are not considered relevant. This causal-effect association between an action and the experienced emotion allows the agent to report its feelings, but also to select

the representative actions from an episode and to generate appropriate episode evaluations.

The mechanisms that decide how the contents of each component are, are described as follows:

1. **Abstract:** Describes the type of episode, constructed by information from **Details** and **Feeling** fields from **Narrative**. If an episode contains more than one cause-effect action, the actions that create the highest emotional intensities will be selected to complete the Abstract field. Thus an agent will remember the situation (**Details** and **Feeling**) that has strongest emotional impact to itself as the rough abstract for a given episode.
2. **Narrative:** Detailed descriptions of the event, including who, when, where and how does this event happen, as well as the emotional intensities the agent had while experiencing the event.
 - To achieve flexibility while generating episode summaries, the **Time** field will record three different types of data:
 - (a) *Real time* (RT) is the time that represents the real world; it will be used for calculating the time difference when the software is used over a long term with a user.
 - (b) *Narrative time* (NT) indicates the virtual time when an episode takes place in the whole story.
 - (c) *Event sequence* (ES) simply shows the order of the event in time, e.g. the first or second event.
 - The field **Details** shows all cause-effect actions that have brought emotional impact to the agent within an episode. Consequently, each action in the field **Details** is corresponding to one set of emotional impacts in the field **Feeling**.
3. **Evaluation:** Agent's psychological interpretations as evaluations of each cause-effect action in the episodes. This essentially shows the action's consequence(s) in terms of inter-personal relationships remembered in an agent's memory.

A similar narrative structure for stories remembered by bottom-up Artificial Life agents in their autobiographic memory has been implemented and evaluated in [11]. Example in Figure 1 shows abstracted autobiographic episode structure with sample contents.

3.3 Storing Personal Experience into AM

At a given time, there is always an active episode in AM. This episode represents the present, i.e. what is currently happening. When a new event is perceived it is necessary to determine whether to store that information in the current active episode, or create a new episode to store it. An episode in AM is strongly associated with a location and time, thus whenever the agent's location changes or if nothing happens for a certain amount of time, a new episode is created and is selected as the active one.

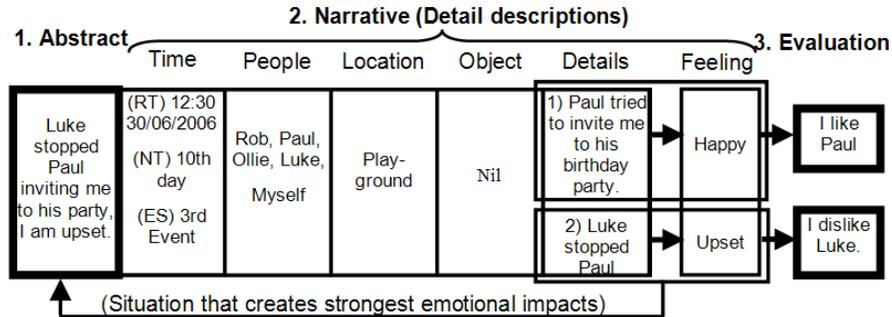


Fig. 1. Details of an episode

When the event is first perceived, it is stored into the current active episode and the corresponding emotional appraisal is associated to the event. The appraisal process might generate more than one emotion (e.g. being insulted might generate distress and anger), but only the strongest emotion is stored and associated to the event, like if the strongest one suppresses the others. The rationale for this choice is that although the agent can experience more than one emotion, it should express a clearly defined emotional state. As Thomas and Johnston put it *"the emotional state of the character should be undoubtedly perceived by the audience; emotions can be accentuated or exaggerated to be better understood;"*[17].

The subjective appraisal in FAtiMA (where the appraisal variables are specified differently for each character) gives rise to the subjective emotional experience for the character. This means that one character will remember that event with a given emotion, while another character will experience something quite different. Furthermore, the same event might be important for the first one and irrelevant to the second.

Although storing the emotional experience already gives a part of personal experience, we wanted to explore this feature further. Inspired by Dennett's intentional stance, we believe that providing explicit information about the agent's intentions will help us creating an even more believable characters. Fortunately, FAtiMA already models explicitly the activation, failure and success of intentions and activates the corresponding appraisal process. As such we can easily include that information into AM. In order to separately access the influence of this type of information in the character's believability we decided to distinguish between these two different type of events, classifying them as:

- **External Events:** all actions that happen in the virtual environment perceived by the agent.
- **Internal Events:** events associated to the intentions/goals of the agent. These are called internal because only the agent knows about them. At the moment, the internal events stored in FAtiMA are:
 - *Activation of an intention*

- *Success in achieving an intention*
- *Failure to achieve an intention*

Nevertheless, this taxonomy of event types is completely transparent to the AM. An internal event is treated in the same way as an external event. When an internal event is generated, it is stored into AM according to what is defined in Table 1. The subject of an internal event its always the agent itself, the action corresponds to the activation, success or failure of the intention, the target of the event is the intention or goal referenced. Finally, the parameters stores additional information about the intention. For instance, if the intention is to insult John, the intention’s name corresponds to ”insult” and the intention’s parameters corresponds to ”John”.

Table 1. Mapping event information into the narrative structure supported by AM

	Who?	What?	Whom?	How?
External Event	action’s subject	action’s name	action’s target	action’s parameters
Intention Activation	SELF	activate	intention’s name	intention’s parameters
Intention Success	SELF	succeed	intention’s name	intention’s parameters
Intention Failure	SELF	fail	intention’s name	intention’s parameters

3.4 Generation of Summaries

When the agent wants to generate a summary of a past episode, it first needs to get the corresponding event from the AM. The memory retrieval process is able to retrieve a past episode by specifying a set of search keys. For instance one may want to get an episode that happened at a specific location or with a specific set of characters.

Once the episode is selected, the system is then able to retrieve the necessary information in order to generate the summary. The summary of an episode consists in the following information:

- **Location:** the location where the episode happened.
- **Time:** narrative time elapsed since the episode happened.
- **EventDescription:** a list of the most relevant events that happened in the episode, ordered by event sequence.

The location and time attributes are taken directly from the information stored within the episode. However, it is necessary to determine the most relevant events to report, since we don’t want the agent to tell everything that happened, as there might exist events that are not important at all. The more relevant events are considered to be the ones that have generated a stronger emotional

impact in the agent, and thus are determined by selecting the events with the strongest emotions associated to them. At the moment, we choose randomly the number of events to report (from three to five), just in order to get a higher variability in the generation of summaries. The chosen events are then ordered by event sequence, so that the summary generated follows a coherent narrative flow.

In order to provide the user with information on the agent's personal experience about the past episode, we need to add to the event's description the emotion experienced when the event was appraised (e.g. "Luke attacked me and I felt upset"). After some initial tests, we realized that describing the emotions for all events was repetitive and seemed unnatural. Furthermore, this was also causing difficulties in giving the user the most relevant emotions. For these reasons, only the strongest emotions are reported (at the moment, only the strongest two since the summary is short).

Once the information about the summary is gathered, it is sent to a natural language System (Language Engine). This module used by FAtiMA is responsible for generating proper English or German utterances for a requested speech act. A speech act is a special type of action used to perform speech related actions. When an agent wants to perform a communicative speech to another agent, lets imagine saying hello, it gathers the type of speech act he wants to perform (greeting in this case), the sender of the speech act, the receiver of the speech act. Afterwards it sends this information to the Language Engine getting a specific utterance of the requested type (e.g. "Hello!" or "Hi"). The Speech Act with all information (including the generated utterance) is then sent to the virtual world and perceived by other agents. The episode summary is a special type of speech act because it includes additional information (as specified above) necessary to build the summary description. Nevertheless, after being generated the summary speech act is executed by the agent as any other speech act.

For the transformation of the information in the episode summaries into text, the LanguageEngine makes use of the semantic parser SPIN [10]. This is originally a parser for language analysis, but it can be — and has been [10] — also used for language generation. In this original direction, input words are parsed into a semantic type representation. Semantic parsers perform no deep syntactic analysis, so, as there is no detailed syntactic representation to generate the summaries from, this approach is very appropriate here also for generation. The task is then to map input type structures onto one result type containing the sequence of words to be returned.

A SPIN parser is built from three knowledge bases: an ontology, a lexicon and a rule set. The ontology specifies all available types, the features which are allowed for each type and the type hierarchy. The rules match type structures and define how they are replaced. In the lexicon additional information for words is defined, like stem, part of speech, semantic class and the semantic value.

In order to generate the summary text, an extract of the AM is split into events consisting of one action and subject, and optional location, time, target, parameter and emotion elements. The text of an event is then generated by

Extract of autobiographical memory	Text representation
<i>Time</i> (count:2, value:hour) <i>Location</i> (value:playground) <i>Event</i> (subject:Ollie, action: party-invitation, target: me) <i>Emotion</i> (value:joy, intensity: normal)	Ollie invited me to his party this morning at the playground which made me feel glad.
<i>Event</i> (subject: Luke, action: party-invitationsabotagehost, target: Ollie) <i>Emotion</i> (value:distress, intensity: normal)	Then Luke didn't want me to go the party. I was feeling upset,
<i>Event</i> (subject:Ollie, action: party-invitationwithdraw, target:me)	next Ollie didn't want me to come anymore.

Fig. 2. Example of the transformation of an extract of the AM into a text representation

transforming these elements into text and combining them through rules. The single utterances are concatenated using a set of connectional phrases like "and", "then", or "did you see that". By providing alternatives for the transformations of each element, a set of possible textual representations is generated for each summary from which the result is picked randomly. This ensures that successive summaries with similar events are not partly identical.

Figure 2 illustrates the transformation of a summary with an example. The mechanism is the same for English and for German. For German, additional information in the lexicon is used to observe case, number and gender agreement of words.

4 Results

To test our approach and the AM, we have embeded it into one particular example of a pedagogical system. FearNot![8] is a computer application developed to tackle and eventually help to reduce bullying problems in schools. Bullying has associated with it a wide variety of behaviours such as hitting, or kicking, in the case of direct bullying, or, in relational bullying, social exclusion or malicious rumour spreading. Thus, the overall objective of the development of FearNot!, was to build an anti-bullying demonstrator in which children age 8 to 12 experience a virtual scenario where they can witness (from a third-person perspective) bullying situations. The application runs like this: a child sees one episode where a bullying situation occurs (see Fig. 3-a). After such episode, the child is able to interact with the victim, acting as an invisible friend to a victimized character, discussing the problems that arise and proposing coping strategies. The coping

strategy suggested by the user is taken into account by the victim, and will try to apply it in the next episode.



(a) Bullying episode seen by the user

(b) Interaction - John is describing what happened in the previous episode: *"During this morning, Luke made fun of me on the street and then Luke continued teasing me. I was sad. Can you believe it, I considered to fight Luke back and then I was planning to stay safe. I was feeling scared. Can you believe it, I failed to fight Luke back."*

Fig. 3. FearNot! 2.0 snapshots

An improved version of FearNot!, FearNot! 2.0 (Fig. 3) is being developed in the context of an EU funded project (eCIRCUS). Among other features, one of the improvements over FearNot! 1.0 is the introduction of autobiographic memory into the agents. The agents store all events from previous episodes in the AM and are able to describe to the user what happened. Thus, in FearNot! 2.0 at the beginning of the interaction, the victim starts to describe a summary of the previous episode, giving the user its personal experience on what happened. This feature is important as it helps children not only to realize what happened in a past episode (that may have happened a week before for example), but also understand if the victim tried to follow their previous suggestion or not.

As an illustrative example consider Figure 3-b. It shows a description of the application of a suggested coping strategy. In an initial interaction with John (the victim), the user suggested him to fight back Luke (the bully). The next episode displayed another bullying situation where Luke provoked and teased John. However, John did nothing to prevent it or fight back as advised by the child. So, when John interacts with the user again, he will then explain that he did actually considered to fight back, but was afraid to be hurt. That made him remain passive.

The inclusion of the kind of dialogue with explanation of actions and memory experiences will allow for a more natural interaction similar to one that close friends have.

5 Final Remarks and Future Work

This paper describes our current work within FAtiMA. Our aim is to give our agents the ability to talk about their past personal experiences. This is achieved in our work through the integration of Autobiographic Memory together with a subjective appraisal model. The AM associate past events with the emotions experienced when they were initially perceived. This emotional information has the additional advantage of giving us what is relevant in an past episode and it is worth retelling. In addition, the agent's are also capable of remembering their internal intentions and associated emotional states.

We believe, that this feature makes the link between the character's internal state and its external behaviour more easily understandable by observers. This, in turn, strengthens the empathic relation between the observer and the agent, and allows for a more natural interaction.

We are planning to evaluate the influence of the AM and the summary generation in the empathic relation created between the user and the victim. We will test whether the children experience the same kind of emotions the agent is feeling. The level of empathy will be determined and compared between four different kind of agents that we want to test. Therefore we will test four different situations:

- One where agents do not have an AM component whatsoever,
- One where agents only store External Events into AM,
- One where agents only store Internal Events into AM,
- One where agents store both External and Internal Events into AM.

We expect the agents that report both external and internal events from AM to be the ones evoking most empathy by the child users. The reason lies in the summary that they are able to generate about their past significant experiences, including the emotions that were involved in that episode and their internal intentions.

6 Acknowledgments

This paper is supported by the eCIRCUS (Contract no. IST-4-027656-STP) and HUMAINE (Contract no. 507422) projects carried out with the provision of the European Community in the Framework VI Programme, and by a scholarship (SFRH BD/19481/2004) granted by the Fundação para a Ciência e a Tecnologia. The authors are solely responsible for the content of this publication. It does not represent the opinion of the European Community or the Fundação para a Ciência e Tecnologia, which are not responsible for any use that might be made of data appearing therein.

References

1. R. Aylett, J. Dias, and A. Paiva. An affectively driven planner for synthetic characters. In *Proceedings of International Conference on Automated Planning and Scheduling ICAPS06*, UK, 2006.
2. J. Bates. The role of emotion in believable agents. *Communications of the ACM*, 37(7):122–125, July 1994.
3. R. Burke, D. Isla, M. Downie, Y. Ivanov, and B. Blumberg. Creaturesmarts: The art and architecture of a virtual brain. In *Proceedings of the Game Developers Conference*, pages 147–166, 2001.
4. M. A. Conway. *Autobiographical Memory: An Introduction*. Open Univ. Press, Buckingham, 1990.
5. K. Dautenhahn. Embodiment in animals and artifacts. In *AAAI FS Embodied Cognition and Action*, pages 27–32. AAAI Press, 1996. Technical report FS-96-02.
6. M. Davis. *Empathy: a social psychological approach*. Dubuque: Brown and Benchmark Publishers, 1994.
7. D. Dennett. *The Intentional Stance*. MIT Press, 1987.
8. J. Dias. Fearnot!: Creating emotional autonomous synthetic characters for empathic interactions. Master’s thesis, Universidade Técnica de Lisboa, Instituto Superior Técnico, 2005.
9. J. Dias and A. Paiva. Feeling and reasoning: a computational model for emotional agents. In *Proceedings of 12th Portuguese Conference on Artificial Intelligence, EPIA 2005*, pages 127–140. Springer, 2005.
10. Ralf Engel. SPIN: A semantic parser for spoken dialog systems. In *Proceedings of IS-LTC*, Ljubljana, Slovenia, 2006.
11. W. C. Ho, K. Dautenhahn, and C. L. Nehaniv. A study of episodic memory-based learning and narrative structure for autobiographic agents. In *Proceedings of Adaptation in Artificial and Biological Systems, AISB 2006 conference*, volume 3, pages 26–29, 2006.
12. W. C. Ho and S. Watson. Autobiographic knowledge for believable virtual characters. In *Proceedings of Intelligent Virtual Agents 2006 (IVA 2006)*, pages 383–394. Springer LNAI, 2006.
13. C. Linde. *Life Stories: The Creation of Coherence*. Oxford University Press, 1993.
14. S. Marsella, L. Johnson, and C. LaBore. Interactive pedagogical drama. In *Proceedings of the Fourth International Conference on Autonomous Agents AAMAS*, pages 301–308. ACM Press, 2000.
15. K. Nelson. The psychological and social origins of autobiographical memory. *Psychological Science*, 4:7–14, 1993.
16. R. Plutchik. *Empathy and its Development*, chapter Evolutionary bases of empathy. Cambridge University Press, 1987.
17. F. Thomas and O. Johnston. *Disney Animation: The Illusion of Life*. Abbeville Press, New York, 1981.