International Journal of Agent Technologies and Systems

July-September 2009, Vol. 1, No. 3

## Table of Contents

**Editorial Preface**

i IJATS 1(3)
*Goran Trajkovski, Editor-In-Chief, IJATS*

**Research Articles**

1 **Agent Interaction via Message-Based Belief Communication**
*Adam J. Conover, Towson University, USA*
*Robert J. Hammell, Towson University, USA*

29 **Impact on Learner Experience: A Qualitative Case Study Exploring Online MBA Problem-Based Learning Courses**
*J. Heather Welzant, Laureate Education Inc., USA*

45 **Virtual Worlds and the Implication for Accountants: The Case of Second Life**
*Jorge A. Romero, Towson University, USA*

51 **Towards Learning 'Self' and Emotional Knowledge in Social and Cultural Human-Agent Interactions**
*Wan Ching Ho, University of Hertfordshire, UK*
*Kerstin Dautenhahn, University of Hertfordshire, UK*
*Meiyii Lim, Heriot-Watt University, UK*
*Sibylle Enz, Otto-Friedrich-Universitaet Bamberg, Germany*
*Carsten Zoll, Otto-Friedrich-Universitaet Bamberg, Germany*
*Scott Watson, University of Hertfordshire, UK*

**Book Review**

79 **Asking Questions: Using Meaningful Structures to Imply Ignorance**
*Reviewed by Deborah M. Seymour, Laureate Education, Inc., USA*

# Towards Learning 'Self' and Emotional Knowledge in Social and Cultural Human-Agent Interactions

*Wan Ching Ho, University of Hertfordshire, UK*

*Sibylle Enz, Otto-Friedrich-Universitaet Bamberg, Germany*

*Kerstin Dautenhahn, University of Hertfordshire, UK*

*Carsten Zoll, Otto-Friedrich-Universitaet Bamberg, Germany*

*Meiyii Lim, Heriot-Watt University, UK*

*Scott Watson, University of Hertfordshire, UK*

## ABSTRACT

*This article presents research towards the development of a virtual learning environment (VLE) inhabited by intelligent virtual agents (IVAs) and modelling a scenario of inter-cultural interactions. The ultimate aim of this VLE is to allow users to reflect upon and learn about intercultural communication and collaboration. Rather than predefining the interactions among the virtual agents and scripting the possible interactions afforded by this environment, we pursue a bottom-up approach whereby inter-cultural communication emerges from interactions with and among autonomous agents and the user(s). The intelligent virtual agents that are inhabiting this environment are expected to be able to broaden their knowledge about the world and other agents, which may be of different cultural backgrounds, through interactions. This work is part of a collaborative effort within a European research project called eCIRCUS. Specifically, this article focuses on our continuing research concerned with emotional knowledge learning in autobiographic social agents.* [Article copies are available for purchase from InfoSci-on-Demand.com]

*Keywords:* *Cultural Differences; Human-Machine Systems; Knowledge Utilization; Social Learning; Storyboards*

## INTRODUCTION

In addition to the popular utilisation of social agent simulations in areas such as education and academic research, nowadays immersive online virtual worlds allow social agents to be further enhanced through highly frequent interactions with human users. For example, Second Life (Linden Research, 2005), as a well-known and quickly evolving virtual society, has attracted millions of users to experience a new kind of social interaction in a virtual space. Powerful PCs, fast broadband connections and advanced 3D graphics offer users this alternative online reality. Interestingly, users' social activity in such a virtual society involves high levels of cultural and emotional learning, as many real cases demonstrate that are reported by Ananthaswamny (2007). How well can we expect intelligent virtual agents (IVAs) to be able to cope with interactions similar to those that human users experience in a comparable social context?

To answer this question, one of the primary goals from our research project eCIRCUS (2006) "Education through Characters with emotional-Intelligence and Role-playing Capabilities that Understand Social interaction" is to promote intercultural empathy. Cross-cultural conflicts have been the source of violent acts in many countries worldwide, including conflicts involving people who come to live in a different country. This background forms a strong motivation for our goal to develop a virtual learning environment (VLE) that supports intercultural learning and fosters intercultural empathy skills for its users. Through developing an educational role-play game named ORIENT with character-based emergent narrative (Aylett et al, 2005), we aim to establish a fun way to educate boys and girls at the age of thirteen to fourteen years in the UK and Germany. The learning outcomes will be designed specifically for children native to the host countries and will be used by the entire school class.

Not surprisingly, in virtual worlds like Second Life or other popular online games, the large international user population and the freedom given to individuals to openly "live" in the environment introduce a certain level of difficulty in intercultural communication. Naturally most users can handle problems emerging from cultural differences – without much effort they can understand both verbal and non-verbal expressions from other characters, which may be either non-player characters (NPCs) or avatars controlled by other users who may be from different backgrounds[1]. To achieve the same level of social and cultural understanding, IVAs are expected to have an *interdependent* (rather than *independent*) "self" – being more attentive to themselves and sensitive to others[2] (Markus & Kitayama, 1991). Consequently, they need to have the ability to be aware of cultural differences and to learn from others. Therefore, we argue that IVAs will benefit from a type of memory which records events that are meaningful to the agent personally and also allows them to extend their knowledge about others' cultural expressions.

Like human beings, these agents should also respond to a stimulus situation with mediation by cognitive processing at several different stages. As reported by social psychologists (Wyer & Srull, 1989; Kim & Ko, 2007), effects of two factors play an important role in most phases of social and cultural information processing: (a) the affect or emotion that one experiences at the time information is processed and a judgment or decision is made, and (b) the agent's "self"[3]. If we assume that one's own emotional reaction to a stimulus is essentially an aspect of oneself, and emotions can be both cause and effect in relation to one's perceptions, then these two factors are clearly related.

In the psychological literature, two different definitions of emotion are commonly used. According to one tradition emotion is viewed as a set of internal cognitive processes of self-maintenance and self-regulation, e.g. Young defined emotion as an "acute disturbance of the individual as a whole" (Young, 1936, p. 263). Another tradition views emotion as an adaptive and organising response because it can motivate forces and direct activity (Salovey & Mayer, 1990). Both definitions seem to naturally assume that these cognitive processes or responses are universal (e.g. Plutchik, 1994), however, the emotional experience is very much culture-based (Solomon, 1984). The reason is that "emotional meaning is a social rather than an individual achievement – an emergent product of social life" (Lutz, 1988; p. 5). Furthermore, like our thoughts and the language that we use to express them, meaning and regulation of emotion are also shaped by "the interest of cultural cohesion", as pointed out by Bruner[4] (1990, p. 58).

To create the dynamic representation of different meanings of agents' emotions, it is beneficial to use schemas for agents' long-term memory that record significant events (events that are meaningful from the agent's point of view). Since the 1920s, psychologists like Bartlett (1932) have been illustrating that humans appear to deal with complex structural knowledge by using memory schemata – not by storing everything in a 'semantic bin' which works like a warehouse. In computational terms, the memory schema provides us with empty slots to be filled in by recalling meaningful episodic information from our long-term memory. This whole remembering process involves memory reconstructions of one's past experiences and thus produces coherent narratives[5] based on meaningful events which tend to be unforgettable in one's life. Psychologists in recent decades identified this type of episodic long-term memory with personally significant events for individuals as *autobiographic memory* (Conway, 1990).

In light of our interests in agents' cultural and emotional learning using computational autobiographic memory, this article reports our ongoing research into the emergence of agents' knowledge about the "self" and emotion in social simulations. We aim to develop a minimal cognitive agent architecture with essential components for IVAs to perform cultural understanding, to extend emotional knowledge and thus eventually to be able to empathise with other agents or human users. In the rest of the article, we first review psychological literature illustrating cognitive representation of emotion concepts in human long-term memory. Next we discuss research in IVAs using emotional models. Then we introduce autobiographic memory for IVAs and our previous research using autobiographic agents with emotional expressions for narrative storytelling. Following that, we briefly introduce ORIENT (Overcoming Refugee Integration using Empathic Novel Technology) – a software application that is being developed that aims at helping teenagers in emotional and social learning. The next section illustrates several important modules in our agent architecture to support the implementation of agents' personality and emotion, and the integration of autobiographic memory for extendable emotional knowledge for individual agents. Finally, future work is discussed in the concluding section.

## BACKGROUND

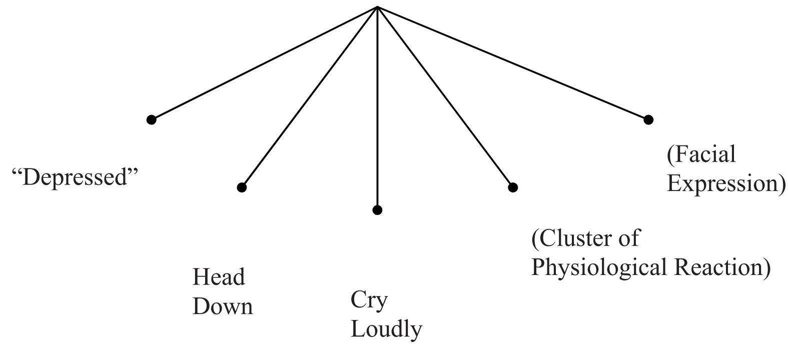
In this section we introduce some essential ideas that support our study in cultural and emotional learning for IVAs. These ideas are very much interdisciplinary, and they cover mainly the areas including 1) *the presentation of emotion concepts in human long-term memory,* 2) *empathy,* 3) *memory and emotion*, and 4) *computational autobiographic agents*. We will illustrate the interrelations that exist between them.

### Representation of Emotion Concepts

Culture not only impacts the language that we use, non-verbal behaviour that we display, our gender roles, etc, but also influences our emotional expressions. Here we specifically focus on agents' emotions because much psychological research has shown that emotion plays a critical role in our memory-encoding, decision-making, interaction with others, and overall, intelligence (Salovey & Mayer, 1990). To create socially "friendly" IVAs, emotion is certainly a critical component to be integrated into the agent architecture. From this perspective, representing concepts of emotions is fundamentally important in agents' memory.

In psychology, the way that humans encode concepts of different emotions has been commonly identified as a type of semantic representation. Emotion concepts, together with other concepts of semantic knowledge, are stored permanently in our long-term memory. Wyer & Srull (1989) suggested that, like concepts of personality, each emotion can be also conceptualised as a central node attaching different configurations as features. They further explained that these features may include verbal labels, representation of overt behaviours, etc. As an example, Figure 1 illustrates their idea about the cognitive representation of the emotion concept of "sad".

*Figure 1. Cognitive representation of the emotion concept "sad"*




Understandably, features attached to the central node are derived from social learning. They are, however, not only used for expressing the internal emotional state of oneself; more interestingly, in social interactions they are often used directly to interpret others' emotions. Consider the following scenario: A mother who observes her young son slamming the house door or performing another behaviour that she personally labels as "angry" may ask him "Why are you angry?". Therefore, emotion concepts stored in the long-term memory are crucial to daily life because the retrieval of these concepts can often help process observable information in interaction with others and in the process of making a decision of how to react or respond to the situation.

## Believable Agents and Modelling Emotion

Research on believable agents in Artificial Intelligence (AI), inspired by work on believable characters in the Arts, has been pioneered by the Oz project under the guidance of Joe Bates. A believable character is "... one that provides the illusion of life, thus permitting the audience suspension." (Bates 1994, p. 122). Work on believable agents in AI has grown significantly, a detailed overview in believable emotional agents can be found in (Loyall, 1997).

Integrating an emotion model into the design of an agent architecture is important for believable interactive virtual agents in applications such as computer games or training software. Marsella and Gratch (2003) utilised appraisal processes from the OCC model (Ortony et al, 1988) for believable characters that perform in various applications. They state that appraisal variables enable agents to characterize the significance of events from the individual's perspective as the interpretation of each event is altered by an agent's own beliefs, desires and intentions, and past events. To continue this research direction with computational autobiographic memory, our previous work developed a mechanism that creates *significant experience* for believable virtual characters through OCC appraisal processes (Ho et al. 2007a).

## Empathy

In addition to inferring others' internal states, emotion concepts are fundamentally important for the *empathic process*. Empathy refers to "any process where the attended perception of the object's state generates a state in the subject that is more applicable to the object's state or situation than to the subject's own prior state or situation" (Preston & de Waal, 2002, p. 4). Contemporary empathy research agrees on two main aspects of the empathic process, one is cognitive, and the other is affective in nature (Davis, 1996; 2006). Both aspects are clearly relevant to our research. Cognitive empathy focuses on the "intellectual or imaginative apprehension of another's condition or state of mind without actually experiencing that person's feelings" (Hogan, 1967, p. 308), resulting in an understanding of the target's inner state. Affective empathy is described as "a vicarious affective response to another's feeling" (Hoffman, 1977). Therefore, learning emotional expressions from another culture seems to rely on the understanding of internal states of the members of the culture and creating affective evaluative reactions towards them.

Furthermore, there can be two types of development for emotional reactions during the empathic process (Davis, 2006): 1) Through a repeated simultaneous experience of an emotional expression in a target person and an emotion in the observer, 2) through associative processes, which work similarly, but rely more on memory representations of past experiences of the observer; these associations may emerge from visual perception of expressions or situational cues.

## Autobiographic Memory and Emotion

Being aware of how high level empathic processes help in increasing the cultural understanding of others, we now study further a rather basic but essential concept of attributing emotion to a specific kind of episodic memory – *autobiographic memory*. In psychology, autobiographic memory has been introduced as memory that contains significant and meaningful personal experiences for a human being (Nelson, 1993). It serves important functions in providing the basis for social interaction, maintenance of a dynamic self-concept as well as the representation of the meaning of concepts (Conway, 1990). Moreover, when memories were plotted in terms of age-at-encoding highly similar life-span memory retrieval curves were observed: the periods of childhood amnesia[6] and the reminiscence bump[7] were the same across cultures – suggesting that there are culturally invariant features of autobiographical memory that yield structurally similar memories across cultures (Conway et al, 2005).

In the literature on human memory, it has been widely acknowledged that events associated with emotional experiences partly constitute highly available memory. Psychologists and cognitive scientists also propose that when experiencing an event with emotional content, a human's cognitive system is more fully engaged in processing that event, in comparison to the processing of events which are not clearly associated with any strong emotional experience. This view can be further elaborated with respect to the frequency of rehearsal (Conway, 1990) – in that highly emotionally intense events are more readily available for retrieval.

Personally significant events which are directly involved in the self memory structure, like first time experiences, can have stronger impacts on humans' lives by creating a pre-existing knowledge structure for other similar events (Conway, 1990). These life events, together with events with emotions, indicate that central knowledge structures relating to the self have been employed in representing autobiographic memory. Nevertheless, studies e.g. from Markus & Kitayama (1991) suggest that emotional processes may differ from the nature of the self-system. They further pointed out that both 1) the predominant eliciting conditions, and 2) the intensity and frequency of emotions expressed and experienced by individuals, vary dramatically according to one's construal of the self.

## Computational Autobiographic Agents

Conceptually, an autobiographic agent is an embodied agent which dynamically reconstructs its individual history (autobiography) during its life-time (Dautenhahn, 1996). This individual history helps autobiographic agents to develop individualised social relationships and to communicate with others, which are characteristics of social intelligence. It has been suggested that autobiographic memory for agents may also lead to more appealing and human-like engaging interactions, making the agents 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. in (Ho et al, 2006). These architectures include memory modules which are commonly acknowledged in psychology: short-term, long-term, and positively and negatively categorised memories. In a series of simulation experiments we showed that 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 previously 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 (Ho & Watson, 2006). Autobiographic knowledge may also support long-term development and learning in synthetic agents as they gain new experience from acting in each new situation.

Our previous research also includes the investigation of how the engagement of users can be increased in an interaction through the inclusion of believable agents with their own emotions and autobiographic memory (Ho et al, 2007a; 2007b). Specifically, in Ho et al. (2007a), we incorporated the psychological view that emotions can arise in response to both internal and external events. More importantly, both types of events have "*a positively or negatively valenced meaning*" (Salovey & Mayer 1990, p. 186) for an individual agent.

More recently our study in Ho et al. (2008) shows that, embedded with communicative autobiographic memory, agents' behaviours can be understood as intentional, narratively structured, and temporally grounded. Furthermore, the communication of experience can be seen to rely on emergent mixed narrative reconstructions combining the experiences of several agents.

Earlier in this article we discussed the semantic representation of emotion concepts and how these concepts can be learnt. Now we define *learning* in this article as more than just acquiring knowledge. Agents use the emotional knowledge to improve their performance – in the sense of achieving believability and performing cultural interactions with other agents and human users. We therefore focus on low level symbolic learning for agents aiming to increase their cultural and emotional knowledge through gaining experiences from interactions with other agents/users and using autobiographic memory.

## CULTURAL AND EMOTIONAL LEARNING FOR SOCIAL AGENTS

This section presents the challenge of creating IVAs that are able to extend their emotional knowledge based on cultural interactions with other agents or users. We start with identifying several limitations of existing research in the area. We then introduce the software application ORIENT developed within the eCIRCUS project, the agent architecture and examples of story scenarios in which agents' intercultural learning ability is required for human-agent interactions. Next, we present our agent architecture design and specify the *minimum requirements* for agents embedded with computational autobiographic memory in ORIENT to be able to learn emotional knowledge through social and empathic interactions.

## Issues and Challenges

As discussed in the background section, humans represent concepts of emotion in long-term semantic memory. Importantly, these concepts are *not* innate but each individual forms them either from one's own cultural inheritance (i.e. knowledge acquired via folk tales, books or other media etc.) or from past meaningful personal experiences. For this reason, we may not be able to infer others' emotions from purely observable information during the interaction. Instead, backed up by the reviewed psychological literature, we argue that both cultural background knowledge as well as autobiographical experiences of a person or an agent play an important role in recognising and learning others' emotional expressions.

One possible approach is to create a synthetic culture for a group of agents, so that each agent can have the same set of semantic knowledge for emotional expressions that represent their culture – this is our first step in the current implementation of ORIENT, which will be introduced in the next subsection. Unfortunately, when these agents are interacting with users in different circumstances, this set of *static* semantic knowledge is too limited.

Since developing a personality model which is valid across cultures is also essential to our research, we suggest re-interpreting cultural dimensions in terms of personality traits. According to Triandis & Suh (2002), the basic personality traits in the Big Five model (Costa & McCrae, 1992) are the most suitable to incorporate cultural dimensions. They argue that 1) the Big Five model is applicable to a wide variety of cultures, 2) traits are supposedly biologically based and they show the same pattern of developmental change in adulthood, and, most importantly, 3) acculturation effects can be found in the predicted direction. To give an example of point 3: if people live as part of a culture different from their own, then their behaviour and attitudes are likely to change by acclimatising to the lived in culture.

Overall, the Big Five model represents five replicable, broad dimensions of personality: Extraversion, Agreeableness, Conscientiousness, Neuroticism and Openness to experience. These five dimensions are, however, descriptive concepts that focus on illustrating how human behaviour can be traced back to psychological variables and their interplay. Therefore, they still need to be explicated in low-level structures and processes for our implementation.

With the consideration of the developmental stages that a person may pass through, the transition includes social and psychological changes. Hence, one shall undergo a time of important changes, which can be a cause of conflict, and a positive development of personality or a clearer sense of psychological identity (Piaget, 1952). Computationally capturing these changes from all agent interactions requires a careful design specification of episodic memory that allows agents to remember those significant events. Therefore, in addition to the underlying personality model that the agent might be using, computational autobiographic memory is necessary here to establish the learning of emotional knowledge and to emphasise the cultural differences of a particular individual.

Furthermore, although we discussed above that empathy is a crucial psychological process to understand others' internal states, modelling the empathic or similar processes, such as Theory of Mind (Premack & Woodruff, 1978; Baron-Cohen, 1991; Leslie, 1994), for agents is a highly challenging task. In the research field of believable IVAs, researchers have attempted to create empathic agents by, for example, using a data-driven affective architecture with human teaching examples (McQuiggan & Lester, 2006), or analysing human users' physiological responses through skin conductance and electromyography to guide the animated interface agents in performing empathic behaviour (Prendinger et al, 2006). However, in both examples the agent's emotional knowledge is predefined by the designer and is not extendable.

Integrating emotion into agents' behaviour modulation mechanisms has recently been widely studied, e.g. as part of a large European project called HUMAINE (HUMAINE, 2004). However, not many studies of agents' emotional learning based on culture can be found. One can imagine the difficulties: learning new emotion concepts means to match the particular expression the user shows with his/her internal emotional states – however, typically virtual agents can neither perceive the user's facial expression nor detect their physiological changes easily in real-time.

We argue that the first important step towards creating IVAs with abilities of learning emotional knowledge about other cultures, and ultimately empathising with others, is to enable the agents' long-term learning from *meaningful* events. For example, through remembering its significant experiences in the past agent_A learned that event_X always has a negative impact on its internal states and thus it generates a dislike emotion. Agent_A, after sometime, creates a concept "*event_X is harmful and leading to a dislike emotion*" and an associated *avoidance* action. With this conceptual knowledge derived from its own experiences, initially agent_A assumes that all other agents whom it meets have the same concept. Therefore when agent_A sees agent_B expressing a dislike emotion, it automatically infers that event_X (or similar events that have a same effect) has happened to agent_B. Afterward, intuitively agent_A expresses its concern to agent_B and expects agent_B's "feedback". The feedback can be either a confirmation of the concept or an unexpected result, both of them will lead to an update of agent_A's existing concept for event_X and thus increase agent_A's cultural emotional knowledge. Note that the complete description of empathic processes is shown in the next subsection *Specifications for computational autobiographic memory*.

In addition to embedding computational autobiographic memory in the agent architecture, it is also essential to consider the "minimum requirements" as specifications for agents interacting with users or other agents. We aim to achieve that, eventually, both the agent's long-term "self" and emotional knowledge can emerge from such social and cultural interactions.

## The Character-Based Approach

In this section we elaborate the character-based approach, with the features that computational autobiographic memory can provide, as a potential solution to address the issues we raised above. In order to allow our agents to adapt themselves to distinct cultures and to be able to establish empathic relations with others, the computational autobiographic memory model focuses on agents' knowledge representations and how information retrieved from autobiographic memory can support agents' goal processing with the PSI theory's "needs" as a foundation (see below for a detailed explanation of the PSI theory). In this model, meaningful episodic knowledge derived from an agent's past experience forms events, episodes, themes and life periods[8] in a bottom-up fashion. Each of them has an abstraction generated for representing general meanings to the agent itself during different periods of its lifetime in a temporal sequence. The technical design for this model can be found in (Ho et al, 2007b).

The main advantage of our autobiographic memory model is that it attributes changes of internal states (e.g. emotions) in action-situation patterns to show the significance (to an agent itself) of past episodic experience. Thus this approach improves agents' learning and adaptation. Moreover, various levels of abstraction can feature the production of narrative storytelling for describing agents' past experiences as well as forming changeable personalities. Using autobiographic memory knowledge to bias planning current goals is particularly suitable in creating dramatic acting for synthetic agents.

Since the full description of the story scenario in ORIENT (eCIRCUS, 2006) goes beyond the scope of this article, we provide 1) the design of a fundamental personality model with its integration into the main architecture, 2) an example of the story scenario, and more importantly, 3) the detailed specifications of requirements for computational autobiographic memory based on the design for the ORIENT scenario and user-agent interaction.

## Personality Model and Agent Architecture Implementation

In ORIENT we aim to use the descriptive trait concepts from Big Five for modelling personality. However, in order to allow agents to generate realistic and expressive behaviours in real-time from the personality model, it is not possible to simply use those five static parameters as initiated. It means that we will need to create a set of low-level internal states for agents (as main characters interacting with users in the game) to capture the dynamic changes of these states and to map the Big Five personality traits into them. As pointed out by Schaub (1999), this approach will involve three levels of modelling work:

1. Descriptive level (Big Five) that provides information regarding each main character's general tendency in five (trait) dimensions.
2. Behavioural level (e.g. being friendly, aggressive) indicates characters' expression based on their personality during the game-play.
3. Low-level implementations of the system that generate behaviour through internal states (e.g. needs, intentions and goals)

The first step is to define the value of each character's personality trait using the Big Five dimensions[9]. Story writers for the ORIENT scenario and psychologists collaborate and revise these trait values based on the original design of each character in the story. Then these (1st level) trait values will be "translated" into low-level internal states (3rd level) for generating characters' motivations, intentions and goals. Finally, we expect that agents' behaviour (2nd level) will emerge from the dynamics of internal states, computational autobiographic memory and planning processes happening in the agent architecture.

To create the low-level system, the first step is to integrate the motivation module from PSI theory proposed by psychologist Dietrich Dörner (Dörner, 2003; Dörner & Hille, 1995). The original PSI theory is based on the idea that humans are motivated emotional-cognitive beings; therefore it integrates cognitive processes, emotions and motivation. The motivation module in our agent architecture includes existential "needs" (*Existence Preservation*, *Species Preservation* and *Affiliation*), and intellectual "needs" (*Competence* and *Certainty*). By mapping the corresponding Big Five trait values predefined by the story writers and psychologists to the set-point (threshold) of each need, an individual character's dynamic personality is created. For example, if character A was defined as a person who is secretive, wise and rarely speaking to new people, then part of his Big Five traits are: High Conscientiousness, low Extraversion and low Agreeableness. Therefore the relevant set-points for PSI needs translated from traits are: High Competence, High Certainty and Low Affiliation. Table 1 shows a brief guideline for the translation from Big Five traits to PSI needs. Note that the mapping from traits to needs is not one-to-one in most cases.

All needs of each agent must be maintained in order to "survive" in the game environment. Therefore individual agents will try to reduce the deviation of each of their own need from

*Table 1.*

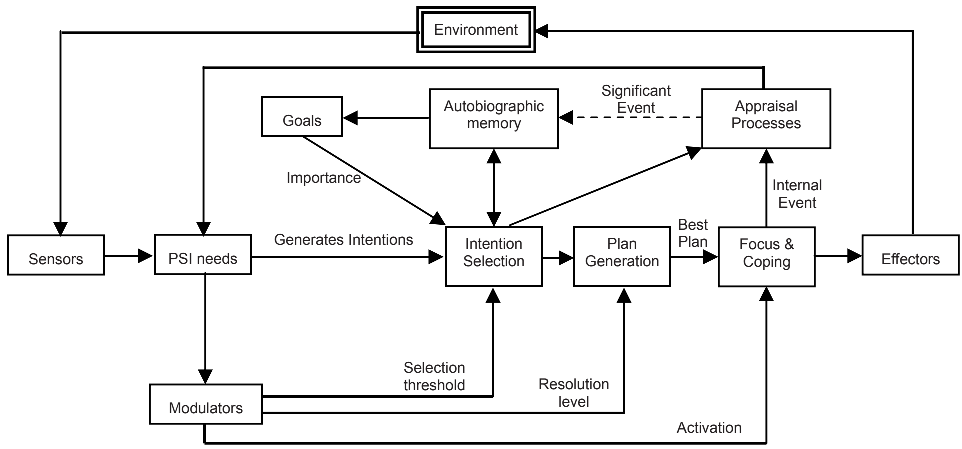
| Big Five trait | PSI needs |
|---|---|
| Neuroticism | Many needs are highly activated at the same time |
| Extraversion | High need for affiliation |
| Agreeableness | Needs are generally not very activated or need long to exceed threshold |
| Conscientiousness | High need for competence and certainty |
| Openness | Low need for certainty, high need for competence |

the set-point as much as possible at all times. Intentions of an agent are built according to the strength (as determined by the deviation) of a need, the success probability to satisfy the need from the semantic knowledge, past significant experiences in autobiographic memory, and urgency (the timeframe to satisfy a need). Whenever a need deviates from the set point, it activates the corresponding intentions. In the case when several intentions are active in a given time, the strongest one will be selected and executed. To help depict the main ideas of the motivation-based architecture as well as the interactions between the computational autobiographic memory and other components, the overall design of the agent architecture is shown in Figure 2.

After the integration of needs for intention generation, we also utilise three types of behaviour modulators from PSI theory:

- **Activation:** Similar to the psychological concept of "arousal", activation represents the preparedness for perception and reaction on the agent – speed of information processing. It increases with the general pressure from the motivation system and the strength of the currently active intention. For example, if the environment poses threats to the agent and its activation level is high, we will expect a short reaction time from the agent – it needs relatively quick cognitive adaptations for the satisfaction of current needs. As a result, only superficial perception and planning may be observed. Thus the agent will be highly cautious and try to adapt its responses to the environmental conditions.
- **Selection threshold:** To prevent the currently active intention to be easily replaced by another equally strong intention, a selection threshold increases the strength of the current intention and prevents oscillation of behaviour by giving priority to the currently active intention. Therefore, low selection threshold means that the agent is easily distracted from its current intention, and vice versa.
- **Resolution level:** Resolution level determines the carefulness and attentiveness of an agent's behaviour –

*Figure 2. ORIENT agent architecture that links components such as needs, goals, emotion and autobiographic memory*




the accuracy of cognitive processes, e.g. perception, planning, action regulation. It changes inversely to the activation value. Generally when the resolution level of the agent is high, it performs more extensive memory retrieval and generates either more alternative plans or a very detailed plan for achieving a selected intention.

Altogether these three modulators in the agent architecture establish the dynamic model of emotion as modulation of "cognition" (here we refer to "cognition" as an agent's perception, action-selection, memory access and planning). This unique aspect was derived from PSI theory, in which emotions are not explicitly defined but emerge from modulation of information processing and adaptation to the environment, see Hille (2007) for details. In this case, an agent's complex behaviours become apparent due to the value of these modulators modified by needs.

Taking other agents and events from the environment into account, not only will they be internally represented symbolically in the agents' semantic memory (world knowledge), but the proposed agent architecture will also include their individual influence on the low-level needs. The strength of the influence is dependent upon an agent's personality specification, and is also affected by its own culture, and its existing experiences. Details of events which have brought significant impact to any of the agent's internal needs, together with other agents involved (if there were any), will be stored in the autobiographic memory for guiding the agent's future emotional and behavioural responses to the same type of events or agents. The relevant process is named *appraisal*, in line with other approaches to modelling emotions in IVA architectures, e.g. (Marsella & Gratch, 2003). Similar to our previous work (Ho, 2007a), for the interpretation of different types of emotional impact, e.g. in order to generate an agent's personal narrative, we utilise the OCC taxonomy (Ortony et al, 1988) as appraisal variables in the architecture.

Therefore, we also project the low-level needs to these appraisal variables. For example, *Like/Dislike* refers to how an event impacts the character's need for *affiliation*; *Desirability (of an event to oneself)* is determined by whether it satisfies or threatens to some extend the character's overall needs; *Desirability (for others)* – depending on how an event affects the other character's needs – is indeed important for an agent's cultural and emotional knowledge learning, and defined by a self-as-first-person evaluation process (see the next subsection *Empathising with others* for details). Furthermore, *Praiseworthiness*[10] is influenced by the agent's need for *certainty* and need for *competence*, also taking into account the social and cultural background of individuals.

At the beginning of an interaction with a user or other agents , each agent's internal needs will be initialised with random values together with a set of predefined "personality thresholds". Based on the level of deviation that each of its current needs has, it generates intention(s) and activates goal(s) that are relevant to the perceived circumstances. The goal with the highest priority to be executed can be determined by the fact that it can satisfy the current most deviated need. Moreover, the agent's competence level plus the past experiences in its autobiographic memory will provide information for the calculation of the success probability for a particular goal. Therefore, both goal importance and goal success probability will be used to manage intention and thus action selections.

We believe that, at a certain level, the composition of an agent's internal needs, goals and autobiographic memory in the architecture can attribute "meanings" to its emotional expressions. During the processes of appraisal, the agent evaluates both 1) internally, the effect of the coping action that it has carried out on its needs and 2) externally, the environmental changes based on that action. This mechanism allows agents with autobiographic memory to link their emotional expressions (for either internal or external events) to the changes of their internal needs. More importantly, their autobiographic memory records these "event-needs" relations as part of their life experiences, and thus they are able to infer the meaning of other agents' emotional expressions.

## An Example from ORIENT Story Scenario

ORIENT is focused on friendship and integration strategies and is aimed at 13-14 year olds. It involves small groups of users interacting with a 3D virtual environment populated by intelligent agents from different fantasy cultures, using innovative communication and control devices and a mission that requires the teenage users to work together with the ORIENT cultures to save the planet. The users' task is to interact with each of the cultures through a number of engagement scenarios with the ultimate aim of saving the planet.

One of the scenarios will be a formal meal where agent_C has invited agent_D, which is from another culture on the planet, to join. The impact to agent_C's internal needs starts when it observes agent_D perform (what is to agent_C) a completely novel ritual when eating – the perceived situation reduces the level of both *certainty* and *competence* because agent_D's actions conflict with its own knowledge and expectation. This will then lead to an increase for these needs, and therefore, agent_C will adapt its future behaviour to take into consideration what they have just experienced. Since agent_C finds no relevant actions in the past to cope with the current situation, through the processes of *Intention Selection* and *Plan Generation* it chooses an emotional coping strategy with the display of *surprised* emotion to agent_D. The coping behaviour and changes to internal needs, together with the context as well as the reaction from agent_D are encoded in agent_C's autobiographic memory as a significant experience.

Since agent_D is from another culture, it may have a different reaction when perceiving agent_C's ritual. This is due to the inclusion of the Big Five which provides a foundational personality for each agent to show their own differences form others. The difference of reaction further leads to self-reflection and the creation of empathy – being aware of how others behave (see next subsection for details).

Generally when encountering an unexpected situation, how much agents adjust their behaviour will depend on their personality and past experiences individually. In the same way each significant event, that brings changes to their needs and triggers the elicitation of their emotion, allows them to learn from experience. Therefore, the learning of cultural emotional knowledge can emerge from intercultural agent-agent and user-agent interactions.

## Specifications for Computational Autobiographic Memory

Given that the agent architecture supports both motivation-based personalities and appraisal models, the next key step is to produce the requirements that guide the integration of computational autobiographic memory. On the one hand these requirements are important to the conceptual specification of the overall architecture for ORIENT agents. On the other hand they are necessary in providing a method to validate the implementation of computational autobiographic memory.

**Agents with "a life":** Autobiographic Memory stores and reconstructs significant experiences that an agent derives from a *long-term* *interaction* with its environment and surroundings (objects, other agents, human users). By using the rich amount of information in autobiographic memory, the agent can largely extend its *temporal horizon* (Nehaniv & Dautenhahn, 1998; Nehaniv, 1999) which serves as an important foundation for an agent's planning process. The extended temporal horizon means that an agent's autobiographic memory and the remembering process provide 'extrasensory' meaningful information for the agent to modulate or guide its immediate or future behaviour – planning for future actions and story-telling about past or imagined events.

For example, it can relate the current situation to some particular moments in its memory, and then an action can be chosen based on the outcome desired by the agent. Therefore, ORIENT agents with autobiographic memory can be seen as "having a life" – a relatively long period of time – through developing such a kind of *interaction history*.

**Repeated user-agent interactions:** To facilitate the development of autobiographic memory, we aim to have users *interact repeatedly with the same* agent for a considerable amount of time in ORIENT. On the one hand, with an initial semantic world knowledge which is consistent with the agent's cultural background, the users' input can often violate an agent's expectations and thus create significant events in its autobiographic memory. On the other hand, while learning a specific culture in ORIENT, it is important for users to create a long-term relationship with an agent from that culture. This relationship has two important aspects: 1) the behaviour performed by the agent precisely reflects its culture – avoiding knowledge inconsistency which may be created by other agents or users that the agents is interacting with; 2) users can easily relate a currently perceived experience from a particular situation to past interactions with a specific agent. The latter aspect indicates that individualised interactions, as opposed to anonymous ones, can be more effective in learning other cultures for users in ORIENT.

Most importantly, a long-term relationship implies that users are familiar with the agent – understanding its personality through typical reactions to external events. This allows users not only to infer the agent's cultural background easily, but also to naturally engage empathically with the agent in various special occasions.

**Forming relationships with agents:** Other game characters in ORIENT which can provide pieces of information as cues for users to solve problems can be modelled as reactive agents (no autobiographic memory architecture) with the sole purpose of guiding users in ORIENT. In contrast, in long-term relationships developed through repeated interactions between the agent with autobiographic memory and users, both sides can gradually recognise each other's role e.g. as friend, ally, or enemy, etc. From this perspective, user-agent interactions in ORIENT emerge socially – the relationship drives the agent to behave differently when users are present, and also to behave differently in interactions with alternative users. This aspect also facilitates the link between an agents' emotion and attitude. Furthermore, agents' narratives can also be personalised for users – reconstructing events from autobiographic memory based on the history of the users' interaction.

**Supporting Emotional Coping:** In Ho et al (2007a) we showed that FearNot! agents (child-like characters inhabiting a virtual environment designed to teach school children how to better cope with bullying) possessing a simplified version of autobiographic memory can express their emotion by remembering past significant events (Ho et al, 2007; Dias et al, 2007). In FearNot! a user can interact with a victim character by offering potential coping strategies, after watching episodes of bullying take place in a virtual school environment. Our aim in the ORIENT agent architecture is to utilize the process of expressing emotion as one of the coping strategies when the agent is experiencing an unexpected situation.

Based on the information available to the agent from its autobiographic memory, a new set of goals will need to be formed in order to cope with the current unexpected situation and to reduce this discrepancy between expected and actual events. If no relevant information can be found to re-formulate the goals, the agent will be forced to fall back on emotion-based coping strategies. In our previous work (Dias et al, 2007), the agent architecture FAtiMA was implemented with a series of appraisal mechanisms and emotion-based coping strategies rooted in the OCC model (Ortony et al, 1998).

In addition to influencing the normal goal formulation process, agents with emotion can also generate coping strategies in the case that their autobiographic memory is lacking. Finally, these new coping strategies are themselves encoded into the autobiographic memory and can be used in the future to create and continuously update the working "self" to support the goal accomplishment.

**Empathising with others:** In ORIENT, we aim to create a concrete agent model with autobiographic memory and a feasible approach to enable agents to, at a certain level, "understand[11]" and thus empathise with others. The idea is to allow agents, by reconstructing past experiences, to reach the perceived physical-psychological states (e.g. emotions) of the target agents as closely as possible – perception of the target agent's feelings is a projection of the self. From the long-term autobiographic memory base, agents can remember past experiences which associate with specific physical-psychological states individually, as described in Ho et al (2007b). When a target agent's physical-psychological states are perceived by the empathiser (an agent that possesses autobiographic memory) then the empathiser will attempt to "imagine" a series of events in order to understand the target agent's situation. This first step models cognitive empathy. Afterward, through rehearsing the relevant experiences (internally re-perceiving these events from memory), the empathizing agent reaches similar physical-psychological states. This second step models affective empathy. At the end of this process the empathiser will express its physical-psychological states to the target agent; the target agent then may receive this empathy and feel experientially "understood", as suggested by Dautenhahn (1997, p. 20) "the concept of 'experiential understanding' can be described by dynamic mechanisms of resonance and synchronization". She further proposed that the metaphor of viewing artefacts as dynamic systems, studying interactions between an artefact and its environment, and correlating them with dynamics inside the agent could be a useful approach to experiential grounding of 'social understanding' in agents. The computational modelling of empathy discussed in this article is inspired by this view.

Based on the empathy cycle (Dautenhahn, 1997, Barrett-Lennard, 1993) the following necessary requirements and steps for modelling empathy in agents with autobiographic memory are suggested:

1. The "willingness" of the empathising agent after perceiving the target agent's physical-psychological states: a precondition for empathy to occur is to listen personally with truly interested attention and non-judging receptivity (Barrett-Lennard, 1997).
2. The available experience of the empathising agent: this enables the agent to "imagine" and thus "understand" the target agent's situation – remembering experiences, based on its own experiences or told by other agents, to allow an agent to reach the similar physical-psychological states that the target agent is currently in.
3. Rehearsal of the most appropriate experience: after all relevant experiences are retrieved from the autobiographic memory base and reconstructed with general event representations (Ho et al, 2007b), the empathiser selects the most similar experiences to rehearse internally[12]. The selected experience can bring the closest physical-psychological states (perceived from the target agent) to a) itself – if the experience is from its own or to b) other agents – if the experience was derived from storytelling or observation. Note that the reconstruction and selection of the most appropriate event is based on the empathiser's own perspective, therefore cultural differences may create 'difficulties' for this agent to have a good understanding of the target agent's situation. This exploration of such 'difficulties' is very relevant for ORIENT in order to support intercultural empathy.

4. Verbal or other behavioural expression toward the target agent: based on the selected experience that was rehearsed internally, the empathiser performs a certain behaviour to show its understanding of the situation to the target agent.
5. Finally, the target agent will receive empathy and thus feel "understood". Cultural differences may again create problems for the target agent to understand the expressed empathising action from the empathizer.

With this set of minimum requirements given to both agent architecture design and human-agent interaction specification, the *autobiography* of each agent can be richly developed and thus the agent can react to situations in a more believable (human-like) way. Based on the psychologically inspired approach, we can assume that 1) the agent's conceptualisation of emotional knowledge becomes more solid because meaning is now assigned to the individual agent's memory contents, e.g. as emotional impact; and 2) the inference of another's emotions based on unique subsets of internal physiological reactions can be established through the empathic processes.

As a result of having a dynamic personality model, appraisal processes and specifications of computational autobiographic memory for the implementation of the ORIENT agent architecture, we expect that agents in ORIENT can achieve social learning of the "self" and emotion knowledge from other cultures in their autobiographic memory. Eventually the selective memory encoding processes for emotion concepts from an agent will include: observing the target agents' behaviour, empathising, receiving feedback and exchanging experiences through communication with the agent.

## CONCLUSION

In this article we discussed how agents can extend their emotional knowledge in social and cultural interactions with other agents or users. The computationally feasible and psychologically supported approach we proposed here for the implementation of IVAs integrates computational autobiographic memory with a cross-cultural personality model and the specification of requirements that facilitate the interactions. Essentially, computational autobiographic memory consists of temporal and episodic information encoded with sensory and emotional contents. As discussed in the previous section, it is critical to enable agents to have this kind of "autobiography" to derive the meaning of emotions and thus to learn new emotion concepts from another culture.

The main direction for this work is to complete the technical implementation of both the personality model and the computational autobiographic memory and the integration of them into the main agent architecture for game characters in ORIENT. As described in the previous section, the main agent architecture will be developed based on the PSI theory, and the emotion and appraisal models from the OCC taxonomy. Autobiographic memory will then play an important role in generating emotions and perception within the character's "mind", e.g. the character may perform individual emotional reactions while encountering other characters or users. In order to achieve this, appraisal variables must be dynamic and determined from the agent's experiences through the information stored in autobiographic memory. We also aim to carry out an evaluation in order to study whether the proposed memory model can increase the believability and interactivity of agents in the game played by different users.

Another future direction is to explore the social interaction between users and agents through gestures. Since using gestures can be a way to express one's emotion, and meanings of a gesture may vary in different cultures, this part of the research can extend the current approach of cultural and emotional knowledge learning for both users and agents. It is supposed that at any time the user can perform any gesture in front of game characters, but some gestures might have very different connotations depending on the context of the social and cultural interaction. Currently our project is examining the possibility of using the Nintendo Wii Remote[13] for recognising users' gestures as new input interface for ORIENT.

Finally, we are also interested to investigate *group* simulations in cultural learning and interactions. As reported by studies in psychology, the peer-group is increasingly important to adolescents and 90% of them identify themselves with a peer group (Palmonari et al, 1992). Also, according to the group socialization theory (Harris, 1996), teenagers' identities are shaped more by their peers than their parents because of peer pressure present in their environment. Therefore, modelling power relationships and the structure among members within a group is interesting – the expression of an agent's emotion can be influenced by both cultural background and the position of the member in the group.

## ACKNOWLEDGMENT

This work was partially supported by European Community (EC) and is currently funded by the eCIRCUS project IST-4-027656-STP. The authors are solely responsible for the content of this publication. It does not represent the opinion of the EC, and the EC is not responsible for any use that might be made of data appearing therein.

## REFERENCES

Ananthaswamny, A. (2007). Technology virtual worlds: a life less ordinary offers far more than just escapism, *New Scientist, 195*(2618), 26-27.

Aylett, R., Louchart, S., Dias, J., Paiva, A., Vala, M., Woods, S., & Hall, L. (2006). Unscripted Narrative for Affectively Driven

Characters, *IEEE Computer Graphics and Applications, 26*(3), 42-52. IEEE Computer Society.

Baron-Cohen, S. (1991). Precursors to a theory of mind: Understanding attention in others. In A. Whiten (Ed.), *Natural theories of mind: Evolution, development and simulation of everyday mindreading*. Oxford: Basil Blackwell, 233-251.

Bates, J. (1994). The role of emotion in believable agents, *Communication of the ACM, 37*(7), 122–125.

Bartlett, F.C. (1932). *Remembering: A Study in Experimental and Social Psychology*. Cambridge University Press, Cambridge, Great Britain.

Barrett-Lennard, G.T. (1993). The phases and focus of empathy. *Br. J. Med. Psychol.*, *66*, 3-14.

Barrett-Lennard, G.T. (1997). The recovery of empathy – toward others and self. In A.C. Bohard & L.S. Greenberg (Eds.), *Empathy reconsidered – new directions in psychotherapy*. American Psychological Association, Washington, DC, (pp. 103-121).

Bruner, J. (1990). *Acts of Meaning*. Cambridge, MA/London: Harvard University Press

Conway, M.A. (1990). *Autobiographical Memory: An Introduction*. Buckingham: Open Univ. Press.

Conway, M.A., Wang, Q., Hanyu, K., & Haque, S. (2005). A cross-cultural investigation of autobiographical memory: On the universality and cultural variation of the reminiscence bump. *Journal of Cross-Cultural Psychology, 36(6),* 739-749.

Costa, P.T., & McCrae, R.R. (1992). *Revised NEO Personality Inventory (NEO PI-R) and NEO Five Factor Inventory. Professional Manual*. Odessa, Florida: Psychological Assessment Resources.

Davis, M.H. (1996). *Empathy: A social phsychological approach.* Westview Press: A Division of Harper Collins Publishers.

Davis, M.H. (2006). Empathy. In J.E. Stets & J.H. Turner (Eds.). *Handbook of the Sociology of Emotions*. New York: Springer, (pp. 443-466).

Dias, J., Ho, W.C., Vogt, T., Beeckman N., Paiva, A., & Andre, E. (2007). I Know What I Did Last Summer: Autobiographic Memory in Synthetic Characters. *Proceedings of Affective Computing and Intelligent Interaction 2007*, ACM Press.

Dautenhahn, K. (1996). Embodiment in animals and artifacts. In *AAAI FS Embodied Cognition and Action*, pp 27-32. AAAI Press. Technical report FS-96-02.

Dautenhahn, K. (1997). I could be you: The phenomenological dimension of social understanding. *Cybernetics and Systems, 28*, 417-453.

Dörner, D. (2003). The mathematics of emotions. In F. Detje, D. Dörner & H. Schaub (Eds.), *Proceedings of the Fifth International Conference on Cognitive Modeling*, (pp. 75-79), Bamberg, Germany, Apr, 10-12 2003.

Dörner, D., & Hille, K. (1995). Artificial souls: Motivated emotional robots. *In Proceedings of the International Conference on Systems, Man and Cybernetics*, (pp. 3828-3832).

eCIRCUS (2006). http://www.e-circus.org/. Last accessed: 30-Nov-2007.

Harris, J. R. (1995). Where is the child's environment? A group socilization theory of development. *Psychological Review, 102,* 458-489.

Hille, K. (2007). *A theory on emotion.* http://web.unibamberg.de/ppp/insttheopsy/dokumente/Hille_A_theory_of_emotion.pdf, 2007. Last accessed: 18-Dec-2007.

Ho, W.C., Dias, J., Figueiredo, R., & Paiva, A. (2007a). Agents that remember can tell stories: integrating autobiographic memory into emotional agents. *Proceedings of Autonomous Agents and Multiagent Systems (AAMAS)*, ACM Press.

Ho, W.C., Dautenhahn, K., & Nehaniv, C.L. (2006). 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, 3*, 26-29.

Ho, W.C., Dautenhahn, K., & Nehaniv, C.L. (2008). Computational Memory Architectures for Autobiographic Agents Interacting in a Complex Virtual Environment: A Working Model. *Connection Science*, *20*(1), 21 - 65.

Ho, W.C., & Watson, S. (2006). Autobiographic knowledge for believable virtual characters. *Intelligent Virtual Agents 2006 (IVA 2006)*, Springer LNAI, (pp. 383-394).

Ho, W.C., Watson, S., & Dautenhahn, K. (2007b). AMIA: A Knowledge Representation Model for Computational Autobiographic Agents. *Proceedings of IEEE International Conference on Development and Learning (ICDL) 2007*.

Hoffman, M.L. (1977). Sex differences in empathy and related behaviors. *Psychological Bulletin, 84,* 712-722.

Hogan, R. (1969). Development of an empathy scale. *Journal of Consulting and Clinical Psychology 33,* 306-316.

HUMAINE (2004). http://emotion-research.net/. Last accessed, 25-Nov-07.

Kim, H.S., & Ko, D. (2007). Culture and self-expression. In C. Sedikides & S. Spencer (Eds.), *Frontiers of social psychology: The self*. New York: Psychology Press.

Leslie, A.M. (1994). Pretending and believing: Issues in the theory of ToMM. *Cognition, 50*, 211-238.

Leung, K., & Bond, M.H. (1984). The impact of cultural collectivism on reward allocation. *Journal of Personality and Social Psychology, 47*, 793-804.

Linden Research (2005). http://www.secondlife.com/. Last accessed, 9-Nov-07.

Loyall, A.B. (1997). *Believable Agents*. PhD thesis, Carnegie Mellon University, Pittsburgh, Pennsylvania. CMU-CS-97-123.

Lutz, C. (1988). *Unnatural emotions*. Chicago: University of Chicago Press.

Markus, H.R., & Kitayama, S (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review, 98*(2), 224-253.

Marsella, S., & Gratch, J. (200)3, Modelling coping behaviour in virtual humans: Dont worry, be happy, *Autonomous Agents and Multiagent Systems (AAMAS)*, ACM Press, (pp. 313–320).

McQuiggan, S.W., & Lester, J.C. (2006). Learning empathy: a data-driven framework for modeling empathetic companion agents. *Proceedings of the fifth international joint conference on autonomous* *agents and multiagent systems (AAMAS 06)*, Hakodate, Japan, (pp. 961-968).

Nehaniv, C.L. (1999). Narrative for artifacts: Transcending context and self. *Narrative Intelligence*, ser. AAAI Fall Symposium 1999. AAAI Press, pp. 101–104, Technical Report FS-99-01.

Nehaniv C.L., & Dautenhahn, K. (1998). Semigroup expansions for autobiographic agents. In *Proceedings of the First Symposium on Algebra, Languages and Computation.* Osaka University, 77–84.

Nelson, K. (1993). The psychological and social origins of autobiographical memory. *Psychological Science, 4*, 7-14.

Ortony, A., Clore, G., & Collins, A. (1988). *The Cognitive Structure of Emotions*. Cambridge University Press.

Palmonari, A., Pombeni, L., & Kirchler, E. (1992). Evolution of the self concept in adolescence and social categorization processes. *European Review of Social Psychology, 3*, 287-308.

Pike, K.L. (1954). Emic and etic standpoints for the description of behavior. In K.L. Pike (Ed.), *Language in relation to a unified theory of the structure of human behavior. Part I,* 8-28. Glendale California: Summer Institute of Linguistics.

Plutchik, R. (1994). *The psychology and biology of emotion*. New York: HarperCollins.

Premack, D.G. & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences, 1,* 515-526.

Prendinger, H., Becker, C., & Ishizuka, M. (2006). A study in users' physiological response to an empathic interface agent. *International Journal of Humanoid Robotics, 3*(3), 371-391.

Preston, S.D., & de Waal, F.B.M. (2002). Empathy: Its ultimate and proximate bases. *Behavioral and Brain Sciences, 25,* 1–72.

Rubin, D.C., Rahhal, T.A., & Poon, L.W. (1998). Things learned in early adulthood are remembered best. *Memory & Cognition*, *26*(1), 3-19.

Salovey, P., & Mayer, J.D. (1990). Emotional Intelligence. *Imagination, Cognition, and Personality, 9,* 185-211.

Schaub, H. (1999). Die Person als Synergieeffekt. Persönlichkeit als Ergebnis der Interaktion basaler Informationsverarbeitungsprozesse. Presentation at *the 8. Herbstakademie „Selbstorganisation in Psychologie und Sozialwissenschaften"* 27.-29.9.1999 in Jena, Germany.

Solomon, R.C. (1984). Getting angry: The Jamesian theory of emotion in anthropology. In R.A. Shweder & R.A. LeVine (Eds.), *Culture theory: Essays on mind, self, and emotion*. Cambridge: Cambridge University Press, (pp. 238-254).

Tomasello, M. (1999). *The Cultural Origins of Human Cognition*. Harvard University Press.

Triandis, H.C., & Suh, E.M. (2002). Cultural influences on personality. *Annual Review of Psychology, 53*, 133-160.

Wetzler, S.E., & Sweeney, J.A. (1986). Childhood amnesia: An empirical demonstration. In D.C. Rubin (Ed.), *Autobiographical memory* (pp. 202-221). Cambridge, UK: Cambridge University Press.

Wyer, Jr. R., & Srull, T. (1989). *Memory and Cognition in its Social Context*. Hillsdale, New Jersey: Lawrence Erlhaum

Young, P.T. (1936). *Motivation of behavior*. New York: John Wiley & Sons.

## ENDNOTES

1 Developmental psychologist Tomasello argues that many unique characteristics of humans are elaborations of one trait that arises in human infants at about nine months of age: the ability to understand other people as intentional agents (Tomasello, 1999).

2 Markus and Kitayama (1991) argue that the independent and interdependent views of the self in psychology can have a systematic influence on different aspects of cognition, emotion and motivation. They suggest that, in many cultures of the world, the Western notion of the self is seen as an *independent* entity "containing significant disposition attributes and as detached from context". Next they point out, however, that in many other constructs the self is viewed as *interdependent* with the surrounding context, and it is the "other" or "the "self-in-relation-to-other" that is focal in individual experience.

3 Note, we do not claim that the artificial agents possess a concept of "self" comparable to that of e.g. human beings. We are using this concept in a computational sense with reference to an autobiographic memory architecture.

4 Bruner (1990) quotes the results of Barlett's serial reproduction experiments in *Remembering* (1932), namely that the most distinctive characteristics of human memory schemata are 1) being under control of an affective cultural "attitude", and 2) any "conflicting tendencies" likely to disrupt individual poise or to menace social life are likely to destabilize memory organization.

5 In this article we define narrative as "a story being told by, perceived by, or remembered (reconstructed) by an agent", thus a narrative requires a story and an agent interpreting this story. The agent's motivations, goals and other internal states, as well as the context of when and where the story is being told, perceived or remembered, will influence how the story is being (re-)created. 'Narrative story-telling' refers to the specific process of how a story is being told by an agent.

6 Childhood amnesia suggests that children from birth to approximately 5 years of age do not seem to form extensive personal episodic memories. For details, see (Wetzler& Sweeney, 1986).

7 The reminiscence bump is the effect in the temporal distribution of autobiographical memory which suggests that people tend to recall

more personal events from adolescence and early adulthood (10-25 years) than personal events from other lifetime periods (Rubin et al, 1986).

8 Life-period' or 'life-time' of an agent refers to the computational duration of its memory, in the sense of how long the agent interacts with and remembers experiences in its environment.

9 Note that we are not using the Big Five model to operationalise cultural differences in agents. We use it to create roles for a virtual role-play approach which is only an aid or auxiliary means to implement personality differences in agents through need states. Personality dimensions such as the Big Five can be compared across cultures – empirical data shows that their meaning is comparable across different cultures in each dimension (Pike, 1954; Leung & Bond, 1984).

10 Praiseworthiness is the action that an agent performed a praise- or blame-worthy behaviour, from the observer's point of view. Therefore it relates to an agent's abilities to 1) predict the consequences of actions (*need for certainty*) and 2) master problems and tasks, e.g. satisfy one's needs.

11 Note that "understanding" as used in this article in a computational sense refers to the agent's ability to relate experiences to its internal states and past experiences, it does not relate to the phenomenological nature of understanding in biological systems.

12 Rather than assigning the physical-psychological states associated with a past experience directly to an agent, rehearsing this experience is necessary to allow the agent to have another comprehension of this event based on the existing semantic knowledge and activated goals. Furthermore, when an agent rehearses a past experience internally, both the importance and endurance of this experience in the agent's memory are increased.

13 Please refer to the official Nintendo Wii controllers website: http://wii.nintendo.com/controller.jsp.

*Wan Ching Ho received his first BSc and PhD degrees from University of Hertfordshire in 2002 and 2005 respectively. The title of his PhD thesis is* Computational Memory Architectures for Autobiographic and Narrative Virtual Agents. *He is now working as a full-time postdoc research fellow in the Adaptive Systems Research Group in the same university. His research mainly focuses on developing control architectures for narrative and autobiographic virtual agents for an EU Framework 6 funded project eCircus (Education through Characters with emotional-Intelligence and Role-playing Capabilities that Understand Social interaction). Ching Ho's research interests are interdisciplinary, including: applying theories from cognitive science and psychology to computational agent control architectures and developing narrative and autobiographic agents for educational software and computer games. The common aim of these is to increase the agents' believability and the interactivity of the software application.*

*Kerstin Dautenhahn received her PhD degree from the Department of Biological Cybernetics at University of Bielefeld, Germany. She is professor of artificial intelligence in the School of Computer Science at University of Hertfordshire where she coordinates the Adaptive Systems Research Group. She has published more than 200 research articles on socially intelligent agents, artificial life, social robotics, robot learning, human-robot interaction and assistive technology. Dautenhahn has edited several books and frequently organises international research workshops and conferences including hosting the AISB05 convention and IEEE RO-MAN 2006. She was general chair of HRI08, the 3rd ACM/IEEE International Conference on Human-Robot Interaction. She is involved in several FP6 and FP7 European projects (Cogniron, Robotcub, Iromec, eCircus, LIREC, I-Talk) and is editor in chief of the journal* Interaction Studies: Social Behaviour and Communication in Biological and Artificial Systems.

*Mei Yii Lim was born in Malaysia. She did her diploma, followed by advanced diploma in computer science at Tunku Abdul Rahman College, Malaysia from 1998 to 2002. At the same time, she pursued a BSc in computer science (major in computer science and mathematics) from University of Campbell, North Carolina, USA and graduated in June 2002 with Summa Cum Laude. She was the winner of Wallace Ewart Prize for Computer Science and the Dwight Lamar Norwood prize for Mathematics. Later she received a MSc with distinction in virtual environments from University of Salford, UK in July 2004. In summer 2007, she completed a PhD in computing entitled* Emotions, Behaviour and Belief Regulation in An Intelligent Guide with Attitude *from Heriot-Watt University, Scotland. Currently, she is a research associate for the EU FP6 project eCircus. Her research is focused on emotional models, emotional memories, emotional virtual agents, narrative and mobile technologies.*

*Sibylle Enz is research fellow at the Institute of Theoretical Psychology at the Bamberg University where she has been involved in the EU-funded projects VICTEC and eCIRCUS. Her scientific interests centre on empathy, problem-solving in social contexts, self-presentation and evaluation. Currently, she is also working on her PhD on the role of empathy in conflicts at the workplace.*

*Carsten Zoll works as a research fellow at the department for general psychology at the University of Bamberg. He studied psychology and computer science. Before joining the department for general psychology, he worked at the professorship for economic and social psychology at the Technical University of Chemnitz. In his dissertation he dealt with decision making among experts of the financial markets. His scientific work focuses on simulation of humanlike behaviour, modelling the human mind, empathy, decision making and behavioral finance. Apart from his academic career, he works as a child psychologist in both a medical office doing diagnostics and psychotherapy and for the family courthouse as an expert witness.*

*Scott Watson graduated from the University of Lincoln (UK) in 2000 with a degree in psychology, and from the University of Hertfordshire (UK) with a MSc in research methodology for psychology. He taught personality, statistics, and research methods in the School of Psychology at the University of Hertfordshire for three years. In 2005 he became a research assistant on the EU Framework 6 project e-CIRCUS (Education through Characters with emotional-Intelligence and Role-playing Capabilities that Understand Social interaction). He is responsible for the administration of a large-scale longitudinal evaluation of the FearNot! anti-bullying software, and analysis and publication of data from these evaluations. He is also investigating the efficacy of a computational model of autobiographic memory, developed alongside Wan Ching Ho. His professional interests are personality theory, human-computer/robot interaction, and the development of ecologically valid models of artificial intelligence. Moreover, he enjoys applying rigorous psychological principles to the HCI/HRI field.*