Virtual Archaeology Review

# Interactive Embodied Agents for Cultural Heritage and Archaeological presentations

F. Seron, S. Baldassarri and E. Cerezo

GIGA. Advanced Computer Graphics Group
Computer Science Department, Engineering Research Institute of Aragon (I3A)
University of Zaragoza. Spain.

***Abstract***

*In this paper, Maxine, a powerful engine to develop applications with embodied animated agents is presented. The engine, based on the use of open source libraries, enables multimodal real-time interaction with the user: via text, voice, images and gestures. Maxine virtual agents can establish emotional communication with the user through their facial expressions, the modulation of the voice and expressing the answers of the agents according to the information gathered by the system: noise level in the room, observer's position, emotional state of the observer, etc. Moreover, the user's emotions are considered and captured through images. For the moment, Maxine virtual agents have been used as virtual presenters for Cultural Heritage and Archaeological shows.*

***Key words***: *MULTIMODAL INTERACTION, VIRTUAL AGENT, AMBIENT INTELLIGENCE, VIRTUAL WORLDS, CULTURAL HERITAGE, ARCHAEOLOGY*

## 1. Introduction

*"The metaphor of intelligent and human-like computer characters has been around for a long time and they are the result of the convergence of several fields such as computer graphics, computer animation, artificial intelligence, human-computer interaction and cognitive science. It also has close relationships to the robotics area since they can share the same know-how in order to model the cognitive behaviour of autonomous individuals. The impetus of the area also comes from the variety of application areas from training/education systems to human-computer interfaces and entertainment films/computer games or cultural heritage and archaeological shows. Each of these application areas requires different properties at different levels such as autonomous behaviour, natural language communication, recognition of real people, personality modelling, emotional behaviour, adaptation to environmental constraints, user needs, intentions and emotions"* (KASAP et al, 2007).

## 2. New tools for supporting new interactions forms

Currently, most research on social interfaces is related to the design of embodied conversational agents -ECAs- (CASELL et al, 2000). ECAs are agents that are visible in the interface sometimes as an animated talking face, may be displaying facial expressions and, when using speech synthesis, with

lip synchronization, and sometimes they have 3D graphical representation, with complex facial and body movements. These virtual characters are being used in a wide range of contexts (MIGNONNEAU et al, 2005), including education and learning (BOFF et al, 2005, GRAESSER et al, 2005), therapy (MARSELLA et al, 2000), persuasion (ROSIS et al, 2003, BERRY et al, 2005), marketing and entertainment (EL-NASR et al 1999, YUAN et al, 2005), among others.

These computational agents should show affective and expressive behaviors (BURLESON et al, 2004): affective expressions have been argued to be useful to help make agents "believable", expressive behaviors have additionally been associated with useful outcomes such as making agents likeable. Moreover, making agent expressions responsive to human expressions, contributes to make agents "relational," able to construct long-term social-emotional relationships with users. The general vision is that if a user's emotion could be recognized by a computer, human–computer interaction would become more natural, enjoyable, and productive (PRENDINGER et al, 2005). The computer could offer help and assistance to a confused user or try to cheer up a frustrated user, and hence react in ways that are more appropriate than simply ignoring the user's affective state as is the case with most current interfaces.

Our research concerns interfaces that employ embodied agents that support emotional and multimodal and interaction. By emulating multimodal human–human communication and displaying social cues including synthetic speech, communicative gestures, and the expression of emotion, our characters may implement the "computers as social actors" metaphor (BURLESON et al, 2004) and be useful in many different scenarios. We present their use as virtual presenters of cultural heritage and archaeological shows.

The paper is organized as follows. Maxine engine is described in Section 3. Section 4 presents a specific application of the engine for the development of presentations to be made by virtual presenters. And finally, in Section 5, the conclusions are presented, together with a description of current and future work.

## 3. The system behind: Maxine and its agents

Maxine is a script-directed engine for the management and visualization of 3D virtual worlds. In Maxine it is possible to load models, animations, textures, sounds, etc., in real-time as they are needed in the virtual environment. It has been written in C++, employing a set of open source libraries (as it is explained in [BALDASSARRI et al, 2007]). A big effort has been invested in the integration and communication of all these libraries.

In any case, we succeeded in maintaining a good real-time performance (see Table 1).

**Table 1**

*Maxine real-time performance*

*(frames per second figures correspond to full screen displays)*

| Scene | Number Vertices | Frame rate (fps) |
|---|---|---|
| Max actor | 009176 | 090 |
| Maxine actress | 016667 | 031 |
| Interactive Tutorials scene | 020570 | 103 |
| Demo group scene | 075028 | 099 |
| Rome coliseum scene | 455851 | 012 |

The overall architecture of our system is shown in Figure 1.

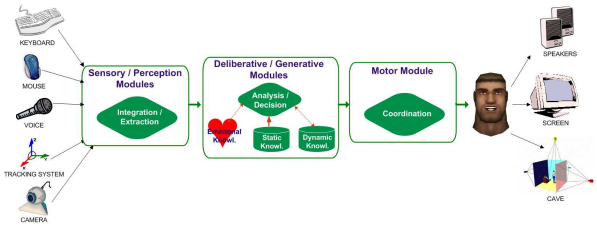


*Figure 1. Maxine's Architecture*

The engine manages scene graphs that can be built in real time, dynamically creating and manipulating its elements by means of a simple command interface. These commands can be executed via script-files when initiating the application or during execution, or can be introduced through the text console every time. The scripting language used is Lua.

A scene graph can be represented by simple objects, like images, texts, videos, geometric primitive's models, lights or 3D sound. But also by animated characters, with different types of animations including secondary animation to increase the expressivity and realism; animated actors, provided with facial animation (see Figure 2), synthetic voices with voice modulation for gaining expressivity and lip-synch.

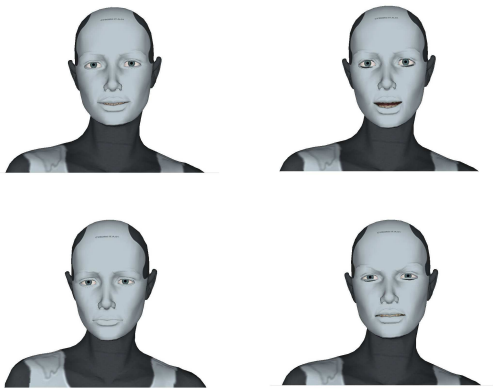

*Figure 2: Different agent's facial expressions (clockwise:happiness,surprise, sadness, anger)*

The engine can also manage several auxiliary elements like cameras, group of elements, animators (for animating group of elements)... and can include animations coming from motion capture systems.

In Maxine, virtual agents are endowed with the following differentiating features (for more details see [CEREZO et al, 2007]):

- It supports interaction with the user through different channels: text, voice (through natural language), peripherals (mouse, keyboard), which makes the use of the generated applications available to a wide range of users, in terms of communication ability, age, etc.
- It gathers additional information on the user and the environment: noise level in the room, position of the user to establish visual contact, image-based estimate of the user's emotional state, etc.
- It supports voice communication with the user in natural language and in Spanish.
- It has its own emotional state, which may vary depending on the relationship with the user and which modulates the presenter's facial expressions, the answers it gives and the modulation of its voice.

## 4. Cultural heritage and archaeological presentations

The system previously outlined enables us to undertake virtual 3D presentations with the agent acting as an Expert Coach. Information presentation is a common and necessary educational act in the Information Society we live and work within: in our group we often make guided presentations to inform, teach, motivate and attract people, showing Computer Graphics applications on cultural heritage and archaeological shows. In Figure 3 the agent is presenting the R&D activity of the group.

Agent capabilities allow to guide a people's attention with the most common and natural methods, gaze and deictic gesture: Maxine agents look at an object as point at it, can move through their environment, pointing at objects when discussing them (see Figure 3), look at the people when speaking to them... Effort has been invested in providing the virtual agent with facial and body expressiveness.


*Figure 3: An Maxine agent points at an image while presents the R&D archaeological activities of our group*

## 5. Conclusions and future work

The system previously Maxine, a powerful engine for managing virtual environments and agents was presented. The system allows the development of new applications where interaction is based on virtual agents supporting multimodal and emotional interaction. Special emphasis has been done in capturing the user's emotion through images, and in synthesizing the emotion of the virtual agent through its facial expressions, the modulation of its voice and its answers.

A specific application on cultural heritage and archaeological presentations show some of its potential. The use of Maxine system and Interactive Embodied Agents has revealed itself as a very useful instrument for guiding, informing, teaching, motivating and attracting people.

Moreover, the use of an agent in the presentation environment opens up new possibilities based on its ability to present information to a great number of people or in a more personalized form, offering a truly multimodal interface, boosting people feelings of self efficacy and being able to adapt itself to user needs.

In spite of all what it's been said, the use of the agent presents several weaknesses: agents are currently complex to create, natural language understanding technology is in its infant stages, text-to-speech suffers from robotic voices, speech recognition technology is not strong enough for widespread use, may distract users and needs students undertaking to be useful. To sum up, the experience in the GIGA group and the use of the agent tool has been very welcome, shows a high level of acceptance and it has ostensibly improved the opinion (and results) of people.

The authors are now working on a new applications for Maxine characters.

Several other research lines also remain open, most of which focus on enriching interaction between the virtual character and the user:

- to consider not only emotion but personality models for the virtual character
- to give the system learning mechanisms, so that it can modify its display rules based on what appears to be working for a particular user, and improve its responses while interacting with that user
- proper validation of Maxine system and characters by their users.

## Acknowledgment

This work has been partially financed by the Spanish "Dirección General de Investigación" (General Directorate of Research), contract number Nº TIN2007-63025, and by the Regional Government of Aragon through the WALQA agreement.

## References

BALDASSARRI, S., CEREZO, E., SERON, F. (2007): An open source engine for embodied animated agents.In *Proc. Congreso Español de Informática Gráfica: CEIG'07*, pp. 89–98.

BERRY, D.et al, (2005). Evaluating a realistic agent in an advice-giving task. In *International Journal in Human-Computer Studies*, Nº 63, pp. 304-327.

BOFF, E. et al, (2005). An affective agent-based virtual character for learning environments. *Proceedings of the Wokshop on Motivation and Affect in Educational Software, 12th International Conference on Artificial Intelligence in Education.* Amsterdam, Holland, pp 1-8.

BURLESON, W. et al, (2004). A Platform for Affective Agent Research. *Proceedings of the Workshop on Empathetic Agents, International Conference on Autonomous Agents and Multiagent Systems*, New York, USA.

CEREZO, E., BALDASSARRI, S., SERON, F. (2007): Interactive agents for multimodal emotional user interaction. In *Proc. of IADIS International Conference Interfaces and Human Computer Interaction*, pp. 35–42.

CASELL, J. et al (eds), (2000), in *Embodied Conversational Agents*. MIT Press, Cambridge, USA.

El-NASR, M. S. et al, (1999). A PET with Evolving Emotional Intelligence. *Proceedings of the 3rd Annual Conference on Autonomous Agents*. Seattle, USA, pp. 9 – 15.

GRAESSER, A. et al, (2005). AutoTutor: An Intelligent tutoring system with mixed-initiative dialogue. In *IEEE Transactions on Education*, Vol. 48, Nº 4, pp. 612-618.

KASAP, Z. and N. MAGNENAT-THALMANN (2007): "Intelligent virtual humans with autonomy and personality: State-of-the-art", in *IntelligentDecision Technologies*. IOS Press

MARSELLA S. C et al, (2000). Interactive Pedagogical Drama. *Proceedings of the 4th International Conference on Autonomous Agents*. Barcelona, Spain, pp. 301–308.

MIGNONNEAU, L. and SOMMERER, C. (2005). Designing emotional, methaforic, natural and intuitive interfaces for interactive art, edutainment and mobile communications, in *Computer & Graphics,* Vol. 29, pp. 837-851.

PRENDINGER, H. and ISHIZUKA, M., (2005). The Empathic Companion: A Character-Based Interface that Addresses Users' Affective States. In *Applied Artificial Intelligence*, Vol.19, pp.267–285.

ROSIS, F. et al, (2003). From Greta's mind to her face: modelling the dynamics of affective status in a conversational embodied agent. *In International Journal of Human-computer Studies. Special Issue on Applications of Affective Computing in HCI*, Vol 59, pp 81-118.

YUAN, X. and CHEE, S. (2005). Design and evaluation of Elva: an embodied tour guide in an interactive virtual art gallery. In *Computer Animation and Virtual Worlds*, Vol. 16, pp.109-119.