humaine

D6h: Final report on WP6

Workpackage 6 Deliverable

Date: 30th January 2008

IST project contract no. **Project title**	507422 **HUMAINE** **Human-Machine Interaction Network on Emotions**
Contractual date of delivery	*January 30, 2008*
Actual date of delivery **Deliverable number**	*February 14, 2008* D6h
Deliverable title	Final report on WP6
Type	Report
Number of pages	12
WP contributing to the deliverable	WP 6
Task leader	Paris 8
Author(s)	C Pelachaud
EC Project Officer	Philippe Gelin

Address of lead author: Catherine Pelachaud
IUT de Montreuil
Université de Paris VIII
140 rue de la Nouvelle France
93100 Montreuil
France

Table of Contents

1 The place of this report within HUMAINE

The final workpackage reports in HUMAINE have been designed as brief statements that complement other sources. On one hand, substantive content – particularly technical content – has been reported in deliverables throughout the project. On the other hand, several potentially overlapping sources are part of the final reporting process. In particular, the activity report for the final period sets out the goals achieved during the final period; and the final report for the project as a whole includes a relatively non-technical account of what the workpackage has achieved.

The core of this report is shared with the section on WP6 in the final report for the project as a whole. That consists of a short, non-technical summary of what the workpackage has achieved. However, there are details that do not belong in a report intended for the general reader (such as information about the participating institutions, meetings organised, etc). They are covered here.

Workpackage members invested a substantial amount of effort in the deliverables, and they form a continuing resource available to the community on the HUMAINE portal. They are listed for completeness. In some senses it would make sense to do the same for publications, but they are too difficult to allocate to individual workpackages. A full list is given in the final report for the project as a whole.

2 Workpackage remit, membership and structure

Within WP6, we are particularly interested in the role of emotion in interaction, in particular for interaction between a human and an Embodied Conversational Agent (ECA). An ECA can be defined as an autonomous entity, that is often represented by an anthropomorphic figure. An ECA is endowed with verbal and nonverbal means allowing her to engage in dialog with humans. To be able to be a communicative interactant, an ECA ought to be endowed with abilities in three domains: perception, interaction and generation, as an ECA must also be able to perceive and interpret emotion as well as show emotion in order to interact in a desirable emotional manner. Thus, the core WP6 exemplar is to propose the definition of an Affective Interactive Embodied Conversational Agent system with capabilities beyond those of present day ECAs.

The field covered by Workpackage 6

An ECA needs to perceive the user's emotional state and the context; she has to react to the user's action and emotional states as well as to the events happening within the context of the interaction; she also needs to show emotions, and to adapt her behaviour to the user behaviour. She needs to catch the user's attention, to maintain it while she is conversing with the user, and so on. WP6 tackles these issues over three core domains:

- • Perception domain: In order to interact in a proper manner, an ECA must first be able to pay attention to, perceive and interpret her conversational partners (e.g. the user) and the context she is placed in; only then can she hope to adapt her (verbal and nonverbal) behaviour depending on the user, her role, and the social and cultural context she is placed in. At a fundamental level, the ECA should have some notion of the details of the conversational partner, their emotion and gesture and the environment.
- • Interaction domain: Interaction is the central theme of WP6. Thus, we are particularly interested in the role of emotion (including emotion proper, mood, interpersonal stance and attitude) in interaction, i.e., how to create an interactive ECA capable of not only showing and perceiving emotion, but also adapting behaviour to convey some emotion or to induce some emotional state in the human user. Interaction is therefore the bridge between perception and generation.
- • Generation domain: Even though the presentation side of ECAs is the most developed in ECA research, further improvement of this area is required, especially as regards a) multimodal integration and the display of emotional behaviours to improve the naturalness of ECAs, and b) the believability of ECAs which is influenced by the consistency of an ECA's behaviour be it in terms of personality, cultural context and situation. Even though decisions on the form of generation (e.g. a particular facial expression or hand gesture) are taken on the basis of the aforementioned domains, it is the generation domain that provides the capability for a proof-of-concept implementation and testing of those concepts and essentially initiates interaction with the user and/or other ECAs.

3 Achievements

Work undertaken within WP6 has been directed toward the Definition of *Affective Interactive Embodied Conversational Agents.* To satisfy that description, ECAs should have several capabilities within the three considered domains. WP3 activities have systematically explored ways of providing the relevant capacities.

3.1 Perception Domain

Cognitive influence on action: We studied the cognitive influences on agent's actions. This capability is concerned with aspects related to cognition that may influence the actions of an agent, such as gaze behaviours, gestures and linguistic expression. In particular, it investigates attention shifts accounting for emotional stimuli and also the adaptation of an agent's politeness behaviours to an emotional user. These issues are related to WP7 themes and collaboration with WP7 has taken place. Here, we emphasise social aspects.

- ✓ *Emotion related attention shifts*: Emotion related attention shifts are concerned with the guidance and allocation of attention by an agent to emotional or potentially important stimuli within a scene for the purposes of interaction with the environment or social entities within that environment. With WP7 partners we have defined a design for computational novelty relating to the early stages of appraisal theory. We have also designed prototype elements of a general framework and scenario in which the novelty system can be embedded and evaluated. In addition, the framework has been designed so as to amalgamate aspects of emotion and attention into a single model, in order to perceive visually emotional and attentive stimuli from a virtual environment and to general emotional and attentive behaviours in a virtual agent towards such stimuli. In collaboration with WP4 partners, we have been working towards extending the framework design to allow it to process input from the real world as well as the virtual environment.

- ✓ *Adapt politeness behaviours to the user's emotional state*: Since emotions have an important impact on perceived face threat, we aimed to create an ECA that can adapt to the user by trying to mitigate such threats through use of gesture, mimics and speech and will also account for the causes of such emotions, rather than just their strengths. The corpus of the Gamble study has been further analysed to account for the reactions of the user towards emotional displays of the agent. These studies gave us further insights into which emotional displays of the agent were appropriate. We first focused on how to augment verbal politeness tactics by gestures and then extended this approach to additional modalities, such as facial displays (masking, inhibiting, fake, felt).

3.2 Interaction Domain

Create affective awareness: Here, we focused on creating connections between a user and an ECA with the aim to maintain user's engagement during an interaction. Sub-topics in this element are *creating affective bonds* between the user and ECA, and *imitation and adaptation* to generate or alter the behaviour of the ECA.

- ✓ *Creating Affective Bonds*: The creation of a bond between the user and the ECA is investigated through monitoring the users' level of engagement with the ECA and also by providing an explicit means for the ECA to try influencing the users' level of engagement in the interaction. We can roughly distinguish between the need for emotional displays to improve the user's awareness of the agent and the user's overall engagement in the interaction, and the need for suitable measurements of affective bonds. One measurement for affective bonds is the user's attention towards the agent which can be assessed by analysing user's gaze behaviours during the interaction. We also looked how agent's misbehaviour (insulting the user or overtly expressing the negative emotion of rage about the current state of the game) increases the impression of believability.

- ✓ *Imitation*: Imitation allows the creation of an affective "awareness" and coordination, and to alter affective bonds by allowing agents to be capable of mimicking certain aspects of another, so that synchrony of behaviour may in some circumstances also lead to a synchronisation of emotions. We have developed a tool called *Animation Score* that allows for the control of the agent not from high level specifications (e.g. the communicative functions of the APML tags), but at the behavioural level. One can specify the behaviours the agent should display through time. In relation with WP4, user's movement and expressivity parameters were extracted automatically. The extracted parameters were sent to the Animation Score that drove the animation of the agent. Several studies were conducted to test which parameters are pertinent to imitate to produce a sense of engagement, be gesture shape or expressivity parameters.

- ✓ *Adaptation*: In order for an agent to be adaptive, information about the emotional state of the user(s) must be obtained. Here, the user's expressive gestures and language usage are analysed to infer his/her emotional state – the inferred state is then used to alter the ECA's reactions. More precisely, we conducted perceptual tests towards the development of a computational model for emotion recognition by means of the analysis of the dynamics of movement and gestures. Part of the work has been done in collaboration in collaboration with WP4. We have then defined a mapping between user's movement expressivity and user's emotional state. Once the emotion is recognized, the agent, when interacting with a user, can plan affective response accordingly and can decide which facial expression to show.

- ✓ *User's engagement*: A single user test scenario was developed which actively reacts to user's gaze behaviour to increase his/her sense of awareness of the agent.

  We have also investigated how to act on user's engagement. In particular we tested the effectiveness of audio-visual feedback system in reproducing movement expressivity of subjects and the impact on user's engagement. A triangulation between methods was used to evaluate the system: notes of the experimenter, video recordings and interviews of the subjects were compared. Results show that the majority of the users perceived the changes in the audio-visual feedback generated by their movement and many of them understood the mapping we designed. Users' comments confirmed that our system provided strong sensations of participation, interaction and immersion.

Backchanneling: ECAs need to become interaction partners. A key feature of spoken dialogue is backchanneling by the listener, providing the speaker with immediate feedback about the

state of the communication. Backchannel utterances and signals can convey simple information of critical importance to the success of the communication attempt, e.g., whether the addressee is still listening, understands, or agrees with the spoken content. To this aim we sought to gain a better understanding of the mechanisms involved in backchanneling, and on finding the set of behaviours and determinants involved in the process. Also, we worked towards defining formal models, implementation and evaluation. Steps taken include: (1) data collection and analysis, (2) automatic detection and interpretation of relevant cues of the speaker, (3) designing a framework for defining rules for the agent, working on a feedback lexicon, (4) setting up an evaluation framework and (5) evaluating some aspects of the feedback lexicon.

3.3 Generation domain

Coordination of signs in multiple modalities: ECAs must show expressivity in a consistent and natural looking manner across modalities in order to be believable. Sub-topics in this element consider coordination of multimodal behaviour from corpora, relationships between signs of emotion in different modalities and gesture repositories for generating appropriate multimodal ECA behaviour.

a. *Copy-synthesis*: Nonverbal communication consists not only of the shape of multimodal behaviors but also of the expressive quality of the behavior movement. Thus, ECAs must show expressivity in a consistent and natural looking manner across modalities in order to be believable. Sub-topics in this element considered coordination of multimodal behaviour from corpora, relationships between signs of emotion in different modalities and gesture repositories for generating appropriate multimodal ECA behaviour. A model of the coordination of multimodal behaviours during non basic multimodal behaviours observed in non acted data, using multimodal corpora has been elaborated. A copy-synthesis approach was applied. It aims at replaying the annotated behaviours coded in video corpus by an expressive ECA. Such an approach proved be useful to validate the annotation and the representation, iteratively refine the annotation schemes and the expressive agents, and for the specification of blends of emotion in expressive agents.

b. *Gesticon*: We need to represent the relationships between the signs of emotion in different modalities, i.e., by designing a representation language for multimodal behaviours. This language, called Gesticon, encodes the mapping between "function" (i.e. a meaning, an emotional content, a pragmatic function etc.) and a certain "form" (i.e. a gesture, posture, facial expression). Gesticon has been designed to be player and graphics model independent. That is, the description of a behavior should not be specific to any particular model geometry nor model animation parametrisation. In addition, Gesticon, has been embedded within BML, Behavior Markup Language. BML is being defined by an international team.

c. *Context dependent emotional body gesture:* In this task we have developed a model that provides an ECA with the ability to modulate its body gestures (in particular upper-body: head, shoulders, arms, hands) according to the emotional state induced by the context – situation in a virtual environment.

Expressivity: Emotion is not simply expressed through a static facial expression or setting of vocal cords. Acoustic and visual expressions of emotion is dynamic; they evolve through time. The manner in which a movement is conducted provides relevant information on the affective state of the emitter. For this capability, we considered the aspects relating to the quality of behaviour.

a) *Behaviour expressivity*: Behavior expressivity refers to the qualitative aspects of co-verbal behaviour executed by ECAs. We have developed a model of behaviour expressivity based on perceptual studies. These studies describe expressivity along several dimensions. We have implemented six of them. These dimensions were used also in an analysis model where algorithms were derived to automatically extract the value of these six dimensions. Working on the analysis of expressive qualities of movement allows us to make discoveries about human movement quality as well as to refine our expressivity model.

b) *Speech expressivity*: Our aim was to better understand how emotions and related states are expressed through the voice, and how the expression through the corresponding acoustic parameters can be achieved using speech synthesis techniques. One major difficulty is in the ability to not only control prosody but also voice quality parameters in (high quality) speech synthesis. Several models were proposed. Signal processing tools were developed to change expressivity in unit selection synthesis; novel methods for creating and manipulating the databases (e.g. diphone-inventories) used in concatenative synthesizers were elaborated; copy synthesis allowed us to use parameters derived from natural speech. Perceptual tests were conducted to measure similarity and fidelity between each synthetic and target sequences, acceptability naturalness and preference.

4 Evidence of esteem

Numerous journal and conference papers have been published on WP6 works over the four years. They are listed in the final project report. We can also report other types of activity.

A special issue of the International Journal of Humanoid Robotics on “Achieving Human-Like Qualities in Interactive Virtual and Physical Humanoids” edited by Catherine Pelachaud and Lola Cañamero was published in 2006.

WP6 members have organised or co-organised several workshops hosted by major conferences, including:

AAMAS 2004 on “Balanced Perception and Actions in ECAs”

AAMAS 2005 on “Creating bonds with ECAs”,

ABUSE 2005, AISB 2005 on “Social Presence Cues for Virtual Humanoids”,

LREC 2004, “Multimodal Corpora: Models of Human Behaviour for the Specification and Evaluation of Multimodal Input and Output Interfaces”,

LREC 2006 and 2008 "Corpora for Research on Emotion and Affect"

HCI 2005 on “Human-animated characters interaction”.

These workshops have gathered many international researchers

Additionally, in September 2007, several HUMAINE members (Catherine Pelachaud, Jean-Claude Martin, Elisabeth André, Gerard Chollet, Kostas Karpouzis, Danielle Pelé) co-organized the 7th International Conference on Intelligent Virtual Agents held in Paris.

5 WP6 Deliverables (all of these are available on the portal)

Del. no	Deliverable name
D6a	Workshop proceedings: Interaction
D6b	Preliminary Plans for Exemplars - Interaction
D6c	potential exemplars: interaction
D6d	Proposal for exemplar and work towards it: Interaction
D6e	Report on representation language
D6f	Mid-term report on emotion in interaction exemplar progress
D6g	Building an Affective Interactive ECA