Emotional agents for collaborative e-learning

Mohamed Ben Ammar, *University of Sfax, Sfax,Tunisia*
Mahmoud Neji, *University of Sfax, Sfax, Tunisia*
Adel M. Alimi, *Université de Pau et des Pays de l'Adour, France*
Guy Gouardères, *Université de Pau et des Pays de l'Adour, France*

Received: September 17th, 2007
Revised: September 30th, 2007

ABSTRACT. In this paper, we discuss the social intelligence that renders affective behaviours of intelligent agents, and its application to a collaborative learning system. We argue that socially appropriate affective behaviours would provide a new dimension for collaborative learning systems. The description of a system to recognize the six universal facial expressions (happiness, sadness, anger, fear, surprise, and disgust) using an agent-based approach is presented. Then, we describe how emotions can efficiently and effectively be visualized in Collaborative Virtual Environments, or CVEs, with an animated virtual head (Emotional Embodied Conversational Agent, or EECA) that is designed to express and act in response to the "universal facial expressions". The objective of the paper is to present the emotional framework EMASPEL, that is Emotional Multi-Agents System for Peer to peer E-Learning, based on the Multi-Agents Architecture approach.

KEYWORDS: *Agent-based approach, Collaborative learning systems, Collaborative virtual environments, Emotional agents, Peer-to-peer e-learning*

Problem formulation

The field of affective computing was proposed and pioneered by Rosalind Picard from the MIT Media Laboratory. Her definition of affective computing is: "computing that relates to, arises from, or deliberately influences emotions" (Picard, 1997). Her argument for putting emotions or the ability to recognize emotions into machines is that neurological studies have indicated that emotions play an important role in our decision-making process. Our "gut feelings" influence our decisions. Fear helps us survive and avoid dangerous situations. When we succeed, a feeling of pride might encourage us to keep on going, and push ourselves even harder to reach even greater goals. Putting emotions into machines makes them more human, and should improve human-computer communication. Also exploiting emotions could lead to a more human decision-making process. Consequently, in this paper, the collaborative affective e-learning framework aims at reintroducing emotional and social context into distance learning, while offering a stimulating and integrated framework for affective conversation and collaboration. Learners can become actively engaged in interaction with the virtual world. Further, the use of avatars with emotionally expressive faces is potentially highly beneficial to communication in Collaborative Virtual Environments (CVEs), especially when they are used in a distance e-learning context. However, little is known about how or indeed whether emotions can effectively be transmitted through the medium of CVEs. Given this, an avatar head model with human-like expressive abilities was built, designed to enrich CVEs affective communication. This is the objective of introducing the Emotional Embodied Conversational Agent (EECA) (Ben Ammar et al., 2006). We are arguing, then, that the use of peer-to-peer network, in combination with collaborative learning, is the best solution to the e-learning environments. Peer-to-peer (P2P) technology is often suggested as a better solution, because the architecture of peer-to-peer networks and collaborative learning are similar (Biström, 2005). This paper explores CVEs as an alternative communication technology, potentially allowing interlocutors to express themselves emotionally, in an efficient and effective way. Potential applications for such CVEs systems are all areas, where people cannot come together physically, but wish to discuss or collaborate on certain matters, for example in e-learning, based on the affective communication.

There are several novel elements to this research. Firstly, although CVEs, as a technology, have been available for more than a decade, user representations are still rudimentary, and their potential is not well explored, particularly the avatar as a device for social interaction. Secondly, the use of emotions to complement and indeed facilitate communication in CVEs is equally under-explored. This is partly because early CVEs research was mainly technology-driven, leaving aside the social and psychological aspects, and partly because the required computing, display and networking resources became available only recently. Thirdly, design guidelines for an efficient, effective, emotionally expressive avatar for real-time conversation did not exist, prior to this research. The multi-agent methodology can certainly bring several advantages to the development of e-learning systems, since it deals well with applications, where such crucial issues (distance, co-operation among different entities, and integration of different components of software) are found. As a result, multi-agent systems, combined with technologies of networking and telecommunications, bring powerful resources to develop e-learning systems. In this research work, we propose emotional framework for an intelligent emotional system. This system is called EMASPEL, that is Emotional Multi-Agents System for Peer to peer E-Learning, based on a multi-agents architecture (Ben Ammar et al., 2006).

Related works

Several projects implement learning systems, based on multi-agents architectures. Some of them work on a generic platform of agents (Silveira et al., 2000). For example, JTS is a Web-based environment for learning Java language, based on a CORBA platform, and using Microsoft agents. In this environment, learners have access to their learner model, and they are able to change it, in the case they do not agree with the information represented (Zapata-Rivera, Greer, 2001). Another example is I-Help, a Web-based application that allows learners to locate human peers and artificial resources, available in the environment, to get help during learning activities. I-Help is an example of a large-scale multi-agent learning environment (Vassileva et al., 2001). Moreover, interesting results have been achieved by pedagogical agents regarding the learner motivation (Johnson et al., 2000), and companion agents acting sometimes as mediators of the learning process (Conati, Klane, 2000). Finally, tutor agents are usually related to learner modelling and didactic decision making. A Virtual Learning Environment (VLE) provides an environment based on network, and the resources in the network are free to share. Therefore, the study process can enhance the collaboration among learners. The VLE can help learners do co-operative study, and make necessary interactions between each other. For example, ExploreNet is a general-purpose, object-oriented, distributed, two-dimensional graphic-based computational environment, with features to support role-playing games for educational purposes, and co-operative learning of many kinds. Among the research works, Prada, in his BeLife system, implements an explicit collaborative mode of utilization. The collaborative mode will challenge learners to share a virtual greenhouse. Each group will be able to alter

some environmental conditions parameters to maximize their crop (Prada et al., 2004). Liu focuses on Web-based collaborative Virtual Learning Environment. He provides a cognitive learning architecture, based on constructivism. The prototype learning system is effective, based on the evaluation. Interaction can promote learners' active learning. During the studying experience, interaction can offer the learner various controls, such as interacting with the virtual environment, and manipulating characters or objects in the virtual environment (Liu, 2005). VLE technologies can address a wide range of interaction capabilities. Picard and her affective computing groups describe affective wearable computers, that are with users over long periods, like clothing, and that are therefore potentially able to build up long-term models of users' expressions and preferences (Ahn, Picard, 2005). The affective wearable offers new ways to augment human abilities, such as assisting the speaking-impaired, and helping remember important information that is perceived.
Synthetic characters, with their significant feature of interactive and intelligent behaviour, can provide a potential tool for learning in VLE. A number of worthy systems and architectures for synthetic character behaviour control have been proposed. For example, Blumberg and his synthetic character groups focus on developing a practical approach for real-time learning, using synthetic characters (Blumberg et al., 2002). Their implementations are grounded in the techniques of reinforcement learning, and informed by insights from animal training (Burke et al., 2001). Similarly, USC/ISI developed a pedagogical agent, called Steve, that supports the learning process. Agent Steve can demonstrate skills to students, answer students' questions, watch the students as they perform tasks, and give advice if students run into difficulties. Multiple Steve agents can inhabit a virtual environment, along with multiple students. Therefore, it helps make it possible to train students on team tasks. At the same time, giving synthetic characters with emotions and personality has a powerful ability to capture and hold students' attention.
Psychologists and pedagogues have pointed out the way that emotions affect learning. According to Piaget, it is incontestable that the affection has an accelerating or perturbing role in learning. A good part of the learners who are weak in mathematics fails due to an affective blockage (Piaget, 1989). Coles suggests that negative

emotions can impair learning, and positive emotions can contribute to learning achievement (Coles, 2004). Some educational systems have given attention to generation of emotion in pedagogical environments - emotion expression and emotion synthesis (Lester et al., 1999), and to the learner's emotion recognition (Conati, Zhou, 2002), pointing out the richness presented in affective interaction between learner and tutor.

We argue that socially appropriate affective behaviours provide a new dimension for collaborative learning systems. The system provides an environment, in which learning takes place through interactions with a coaching computer agent and a co-learner, an autonomous agent that makes affective responses. There are two kinds of benefits for learning in the collaborative learning environment: one is what is often called "learning by teaching", in which one can learn given knowledge by explaining it to another learner; the other benefit is often called "learning by observation", in which one can learn given knowledge by observing another learner working on problem solving, teaching other learners, and so on. While these approaches of collaborative learning are primarily knowledge-based, goal-oriented, and rational, social intelligence might only be utilized as a side effect.

Affection in our framework is considered from various angles, and on different levels:

- The emotional state of the learner will be modelled by an event appraisal system.
- The emotional state of the tutor is modelled as well, including values for emotions, and parameters such as satisfaction, disappointment, and surprise.
- The dialogue acts come in different forms, with variation in affective values.
- Various affective parameters are used in determining which tutoring strategy to use, and which instructional act to perform (sympathizing or non-sympathizing feedback, motivation, explanation, steering, etc.).

Emotional concept ontology

A verbal dictionary can be described as a tool that aims at providing a partial solution to the problem, where two persons neither understand the language the other is speaking, but still want to communicate. One can just look up the meaning of the words of another language. A non-verbal dictionary has the same concept of a verbal dictionary, but it differs in the type of information that is stored. Instead of words, a non-verbal dictionary contains information about all the ways people communicate with one another non-verbally, such as facial expressions, in our case, to construct the emotional ontology. It is well accepted that a common ontology holds the key to fluent communication between agents; most researchers believe that the common ontology is domain ontology. This assignment can be considered as the extension of a previous work, named FED - an online Facial Expression Dictionary (De Jongh, 2002), concerning a non-verbal dictionary. Before we define our research question and objectives, we summarize the idea of a specific part of FED. We only focus on that part of FED, which allows the user to send a picture. This image input will be labeled by emotional word (happiness, sad, and so on). FED requires that the user manually locates the face and Facial Characteristic Points (FCPs). The FCPs are predefined, conform to the Kobayashi and Hara face model. After manually selecting and submitting the points, an emotional word will be output. Thus, FED lacks the ability of automatic extraction of Facial Characteristic Points, that are needed for the facial expression recognition process. In the current situation, user interaction is needed to complete the whole procedure. In our system, the emotional ontology covers the major role that helps the emotional agents distinguish emotions (Cowie et al., 2005). These knots represent features of a current emotion, for example, labels or distances - that is the case of APML-Affective Presentation Markup Language (VHML, 2001).

Update of Emotional Markup Language (EML)

What exactly do we mean by "emotion"? There is much disagreement on this, but one of the most useful definitions, by psychologist Magda Arnold, draws a careful distinction between states and behaviours. In Arnold's theory, emotional experience proceeds in three steps:

1. Perception and appraisal (external stimulus is perceived and judged as good, bad, useful, harmful, etc., mostly based on learned associations).

2. Emotion (internal state of arousal or “feeling” arises, involving physiological effects).
3. Action (specific behaviour, such as approach, avoidance, attack, or feeding, depending on emotional intensity, learned behavioural patterns, and other motivations simultaneously present). In this view, emotion is an internal state, not a behaviour or a perception of external reality (Freitas, 1984).

The EML-Emotional Markup Language defines the emotion elements that affect the VH-Virtual Human, regarding voice, face, and body. The speech and facial animation languages therefore inherit these elements (VHML, 2001).
We have realized some modifications to the APML-Affective Presentation Markup Language (Bilvi et al., 2002), in order to allow the EECA to communicate a wider variety of facial expressions of emotion, as well as to allow for a more flexible definition of facial expressions, and their corresponding parameters. These modifications refer mainly to facial expressions timings, as well as to their intensity; intensity corresponds to the amplitude of facial muscles movements. For each APML tag, we have introduced some new attributes like frequency. The facial expression of an emotion has a limited duration (1/2 to 4 seconds), and the facial muscles cannot hold the corresponding expression for hours, or even minutes, without cramping.

Attributes of EML

EML is an XML-eXtensible Markup Language (W3C, 2002). This implies that it conforms to a standard for the World Wide Web, and hence it can be used with (sufficiently powerful) Web browsers.

Frequency: the number of times an emotion is felt.
Duration: the time taken in seconds or milliseconds of the emotion existence in the human being.
Intensity: the intensity of this particular emotion, either by a descriptive value or by a numeric value.
Wait: a pause in seconds or milliseconds, before continuing with other elements of EML, such as <angry> <disgusted> <neutral> <surprised> <happy> <sadness> <fear>.

Update of EML elements

In our framework, we propose the EmotionStyle language, designed to define style, in terms of multimodal behaviour, and make an EECA display and recognize emotion accordingly. A new feature was added to the EML language, that is distances and frequency attributes to EML, in order to describe more carefully the facial expression. For each, we have introduced some distances like D1 to D6.

<Neutral>

The neutral face represents the reference emotion. The concept of the neutral face is fundamental, because all the distances describe displacements with respect to the neutral face.
Description: facial expression; {D1-D6}=initialized.
Attributes: default EML attributes.
Properties: all face muscles are relaxed, the eyelids are tangent to the iris, lips are in contact, the mouth is closed, and the line of the lips is horizontal.

<Angry>

Description: facial expression; {D2 decreases}, {D1 increases}, {D4 either decreases or increases}.
Attributes: default EML attributes.
Properties: the internal corners of the eyelids decrease together, the eyes are open largely, the lips join each other or they are open to make the mouth appear.

<Disgusted>

Description: facial expression; {D3 increases and D4 increases}, {the other distances remain constant}.
Attributes: default EML attributes.
Properties: the superior lip gets up and is stretched in asymmetrical manner, the eyelids are deconstructed.

<Surprised>

Description: facial expression; {D2 increases}, {D1 increases}, {D4 increases}, {the other distances remain constant}.
Attributes: default EML attributes.
Properties: the eyelids get up. the mouth is open.

<Happy>

Description: facial expression; {D4 increases}, {D3 decreases and D6 decreases}, {the other distances remain constant}.
Attributes: default EML attributes.

Properties: (the mouth is open), the commission is stretched back to the ears, the eyelids are stretched; the eyelids get up, the mouth is open.

<Sadness>

Description: facial expression; {D2 increases and D5 decreases}, {D1 decreases}, {the other distances remain constant}.
Attributes: default EML attributes.
Properties: the internal corners of the eyelids go to the height; the eyes are closed slightly; the mouth is stretched.

<Fear>

Description: facial expression; {D2 increases and D5 increases but more than D2}.
Attributes: default EML attributes.
Properties: the eyelids get up together, and their internal parts go to the height; the eyes are contracted and in alert.

Temporal facial expression features

The facial expression can be defined in relation with the time of changes in the facial movement, and can be described according to these three temporal parameters:

1. Duration of onset: how much time is necessary for the emotion to appear?
2. Duration of apex: how much time does the expression remain in this position?
3. Duration of offset: how much time is necessary so that the expression will disappear?

EMASPEL framework

Architecture

Figure 1 illustrates the architecture of a peer in our P2P e-learning system: in order to promote a more dynamic and flexible communication between the learner and the system, we integrate five kinds of agents.

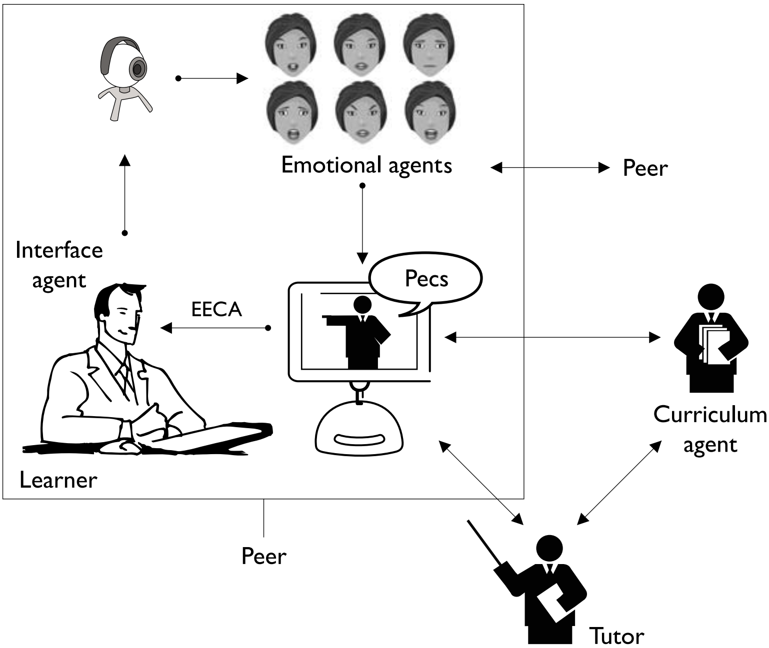


Figure 1. EMASPEL architecture

Interface agent

Transmits the facial information coming from the learner to the other agents of the Multi-Agents System (MAS).
Assigns the achieved actions and information, communicated by the learner, to agents Curriculum, EECA, and the other agents of the MAS.

Emotional embodied conversational agent

Motivation

Agents cannot be content to be intelligent, but must be endowed also with emotions and personality. In the same way, the communication in natural language is not enough, it must be doubled by non-verbal communication. Agents are able to give natural responses, and therefore come across as believable and even interesting conversational partners (Prendinger, Ishizuka, 2004). Animated pedagogical agents are "lifelike autonomous characters that cohabit in learning environments with learners to create rich, face-to-face learning interactions". Animated agents carry a personal effect, that is the presence of a lifelike character which can strongly influence learners to perceive their learning experiences positively (Alexander et al., 2006).

It is widely accepted that animated agents, capable of emotion expression, are crucial to make the interaction with them more enjoyable and compelling for users (Lester et al., 1999). Emotional behaviour can be conveyed through various channels, such as facial display (expression). The so-called "basic emotions" approach distills those emotions that have distinctive (facial) expressions associated with them, and seems to be universal: fear, anger, sadness, happiness, disgust, and surprise. More accurately, Ekman prefers to talk about (basic) emotion families. Thus, it consists in having many expressions for the same basic emotion. The characteristics of basic emotions include quick onset (emotions begin quickly) and brief duration, which clearly distinguish them from other affective phenomena, such as moods, personality traits or attitudes. Note that only enjoyment and possibly surprise are "positive" emotions. The enjoyment family covers amusement, satisfaction, sensory leisure, pride, thrill of excitement, and contentment. Interestingly, the positive emotions do not have a distinct physiological signal. Ekman explains this by referring to the minor relevance of positive emotions in evolution (Ekman, 1999).

Description

In the construction of embodied agents, capable of expressive and communicative behaviours in the e-learning environment, an important factor is the ability to provide modalities and conversational facial expressions on synthetic faces. For example, an animated interface agent is now being used in a wide range of application areas, including personal assistants, e-commerce, entertainment, and e-learning environments. Amongst our objectives is the creation of a model to generate and visualize emotions on an embodied conversational agent. The emotions are particularly important for a conversational agent, since they reveal an essential share of the speech through non-verbal signals. William James perceives the emotions like a direct response to the perception of an event contributing to the survival of the individual, and insists on the changes induced on the body behaviour of the individual. The body answers initially in a programmed way to this change, and so constitutes what one calls the emotions. The feedbacks of the body by the nervous system contribute largely to the experiment of the emotions. Research proved that the emotions succeed the facial expressions. During the learning process, and when interacting with the learner, some tutoring agents may want to express affects. Thus, they use EECA, which is able, within a specific activity, to translate, through a character, the emotions of the tutoring agent. It has to be aware of the concerned task, and of the desired emotional reaction (by the designer or the concerned tutoring agent). The emotional state of EECA is a short-term memory, which represents the current emotional reaction. To be able to compute emotion, a computational model of emotion is required. Our approach was built on Frijda model (Frijda, 1986).

Design of EECAs

This sub-section is about the design of Emotional Embodied Conversational Agents, or EECAs. In this field of Human-Computer Interaction (or HCI), and Artificial Intelligence (or AI), the design of EECAs or "virtual humans", and the communication between those agents and human users is the main object of research. A lot of effort was put into making ECAs more natural and believable, and making communication with ECAs more affective, efficient and fun. One way to improve communication in this way is to make the agent more actively involved in building a relationship with the user. An agent that is able to observe the user, and with its personality, appearance, and behaviour is able to respond to the (implicit) likes and dislikes of the user, in such a way that it can become acquaintances with the user, and create an affective interpersonal relationship: such an agent could have additional benefits over a "normal" Embodied Conversational Agent, in areas such as e-learning. EECA is made of three layers, or modules (figure 2). The first one (perception layer) captures and extracts the facial expressions (image acquisition and face tracking) and proceeds to its categorization (classification). The second one (cognition layer) analyses and diagnoses the perceived learner's emotional state. The third one (action layer) makes decision on remedy pedagogical actions, to carry out in response to the actual emotional state. Tutoring agents are then informed, and may access information in the new affective state to adapt the current tutoring flow accordingly. The cognitive layer includes two main processes named analysis and diagnosis. The analysis of an emotional state recognized by the perception layer makes it possible to translate the meaning of this emotion in the learning context. It is carried out by taking into account several elements, including the recognized emotion, the current affective profile, the history of the actions realized before the emotion expression, the cognitive state of the learner, the emotion evolution, and the social context (if it corresponds to a social or collaborative learning). Agents are virtual human beings. They are designed to imitate or model human behaviour. Human behaviour is complex and many-sided. Nevertheless, it is possible to argue that human behaviour can, within limits, be modelled and can thus be made comprehensible and predictable. Physical, emotional, cognitive, and social factors occur in all forms of human behaviour. Approaches, which regard human beings exclusively as rational decision makers, are of limited value. The modelling of human behaviour plays an important role in all areas, in which action planning, decision making, social interacting, and such like play a part. These areas include, in particular, Sociology Teaching and Education. Consequently, the internal state of the EECA agent is based on the PECS architecture proposed by Schmidt. The PECS architecture is a model of agent that aims at simulating the human behaviour in a group. PECS stands for Physical conditions, Emotional state, Cognitive capabilities and Social status (Schmidt, 2000). These are the four main building blocks of a particular PECS agent architecture, adding a Sensor-Perception module and a Behaviour-Actor module (figure 2). The PECS reference model aims at replacing the so-called BDI-Belief, Desire and Intention architecture (Rao, Georgeff, 1995). Architectures, such as BDI, which conceive of human beings as rational decision makers, are sensible and useful to a very limited degree only. Restriction to the factors of belief, desire and intention is simply not appropriate for sophisticated models of real systems, where human factors play an important role.

PECS's agent model consists of three horizontal layers:

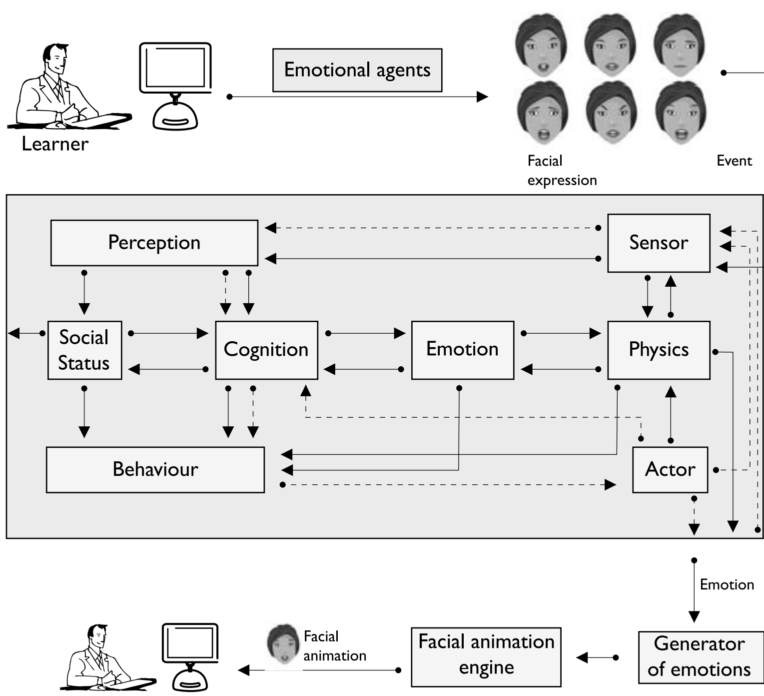
1. Information input layer: this layer processes the input taken from the agent environment, and consists of two components: the Sensor and Perception components. The Sensor component takes the external data by means of a set of sensors, and the Perception component filters the non-relevant data, and processes the information. The perceptions are used to update the mental state of the agent, or for learning purposes.

2. Internal components layer: the personality of the agent is modelled at the second layer. Thus, the parameters of this second layer are crucial to find out the response of the agent to the input taken by the information layer. They consist of four components: Physics, Cognitive, Emotional, and Social Status. The physical and material properties of the agents are modelled in the Physical component. The emotions that can affect the behaviour of the agent are modelled as part of the Emotional component. The agent's experience and knowledge are part of the Cognitive component; and finally, the social features of the agent (for instance, whether the agents like to socialize, or they prefer to be alone) are described in the Social Status component.

3. Agent output layer: this layer is in charge of modelling the set of possible actions and the selection process, and thus it produces the response of the agent, and consists of two components: the behaviour, and actor components. The behaviour component selects the actions that are associated with the input information that reaches the agent, and the agent's response is based on its internal parameters. The actor component takes the actions and executes them. The PECS architecture is not based on any social or psychological theory. The architecture is mainly an integrated model, in which several fundamental aspects to human behaviour and decision-making process are taken into account (Martínez-Miranda, Aldea, 2005). The purpose of the emotional agents consists in extracting the facial expressions (acquisition and facial alignment) and subsequently categorizing them, using the temporal evolution of distances D, like it is demonstrated in table 1. The analysis of table 1 shows that it will be possible to differentiate between different emotions, while being interested in priority in the distances D, which undertake significant modifications. Indeed, there is always at least one different evolution in each scenario. The EECA first carries out an analysis of the emotional state of the learner. The purpose of this analysis is to translate the meaning of the emotion into the training context. It is achieved based on several factors, such as: the emotion sent by the emotional agents, the current emotional profile, the history of the actions carried out before the appearance of the emotion, the cognitive state, the evolution of the emotion, and the social context (if it is about a social training or collaborative).

The expressions in input are "joy", "fear", "dislike", "sadness", "anger", "surprise", and the analysis makes it possible to conclude whether the learner is in state of "satisfaction", "confidence", "surprise", "confusion" or "frustration". The interpretation of the analyzed emotional state is then established. It will consequently determine the causes having led to this situation (success/failure with an exercise, difficulty of work, missed knowledge, etc.), while being based again on the cognitive state of the learner, and thus making it possible for the tutor to take, if it is necessary, the suitable teaching actions. The role of the action layer is to define, even if necessary, a whole set of tasks allowing to remedy at the observed emotional state, in order to bring the learner in a state more propitious to knowledge assimilation (figure 2). For this reason, a collaborative reciprocal strategy in ITS can gain advantage from "mirroring", and then assessing emotions in P2P learning situations.

Figure 2. The EECA-Emotional Embodied Conversational Agent architecture



The emotional agents

Integrated into a learning environment, they aim at capturing and managing the emotions expressed by the learner, during a learning session. They currently capture emotions only through facial expression analysis, and they are in charge of learner emotion detection. They recognize the learner's emotional state by capturing emotions that he or she expressed during learning activities (Nkambou, 2006). For making the affective communication between an EECA and a learner, they need to be able to identify the other's emotional state through the other's expression, and we call this task emotion identification, established by the emotional agents. Extracting and validating emotional cues through analysis of users' facial expressions is of high importance for improving the level of interaction in man-machine communication systems. Extraction of appropriate facial features and consequent recognition of the user's emotional state is the topic of these emotional agents.

Recognition and interpretation of facial expression

The recognition of the facial expressions by the emotional agents is generally done according to the following stages: detection of the face, the automatic extraction of contours of the permanent features of the face, the eyes, the eyebrows, and the lips. Extracted contours being sufficiently realistic, we then use them in a system of recognition of the six universal emotions on the face.

Detection/localization of the face

For this first step of the recognition of expression, we have one approach: this detects and localizes a face in a fixed coloured image (figure 3), being based on the space of colours YCbCr.
The space of colour YUV, which is equivalent to YCbCr, is well adapted to the recognition of the colour of the skin. This space derives from the RGB. U and Y are called signals of different colour, and contain the rest of the information estimated to the colour monitors (U=B-Y is called the chrominance, and V=R-Y is called the saturation, B and R designate, respectively, the Blue and the Red). The format YCbCr is a version of YUV, to which we added a ladder factor, and an offset. The components Cb and Cr offer the advantage of being a little sensitive to the variations of luminosity, and to the colour code. So, I decided to use the space of colours YCbCr for the rest of the development (Figure 3).

Figure 3. Face detection

Figure 4. Mouth localization

The detection of the eyes and the mouth is a precision tool, which helps localize the face. Many methods have been suggested: pattern matching, correlation, structural approach, and neural networks. The method, which we have adopted, is based on the geometric form.

The mouth case

To detect the form of the human mouth in the input image. For this one, we have used a code written under Matlab, which guarantees the search of the mouth, then returns the result found toward the platform (figure 4).
Explanation of Matlab code:
Our Matlab code must extract from a fixed image, and in following phases of pre-processing, the characteristic points of the mouth. These points have the form of an ellipsis that surrounds the mouth. Therefore, it gives the two extreme points, and their coordinates (figure 4):

- Acquirement of a coloured image
- Filtering the image

Conversion of the RGB space toward YIQ by the following formula:

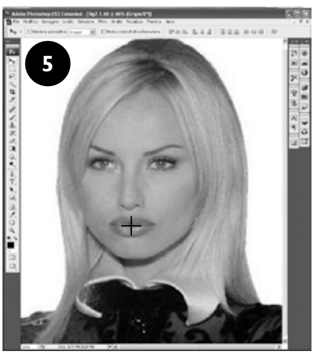
$$Y = 0{,}30\ R + 0{,}59\ G + 0{,}11\ B$$
$$I = 0{,}60\ R - 0{,}28\ G - 0{,}32\ B$$
$$Q = 0{,}21\ R - 0{,}52\ G + 0{,}31\ B$$

R: red; G: green; B: blue; Y: luminance information; I & Q: chrominance; they are the linear combinations of R, G and B intensities.

Filtering the resulting image, following this algorithm:

```
GR= image_result;
for i=1:info.Height
    for j=1:info.Width
        if GR(i,j)<48
           GR(i,j)=255;
        end
    end
end
```

Following another mode of filtering on the 9ppv to suppress all manners of stigmas and points that don't make part of the mouth (Figure 5).
Research of an ellipsis shape on the filtered image, and displaying its centre coordinates.



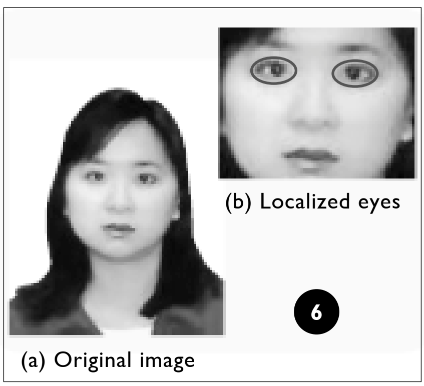


Figure 5. Mouth detection

Figure 6. Eyes detection

The eyes case

The approach is based on the analysis of the global texture of a face, and the geometrical approach. The knowledge of these features enables us later to normalize the geometry of the studied face (Figure 6).
To detect the form of the human eyes in the picture in process. For this one, we have used a code written under Matlab, which guarantees the search of the mouth then returns the result found toward the platform.

Explanation of Matlab code:
Our Matlab code is based on the analysis of the global texture of a face and the geometric approach (figure 7):

- Acquirement of a coloured image
- Converting to YCbCr space

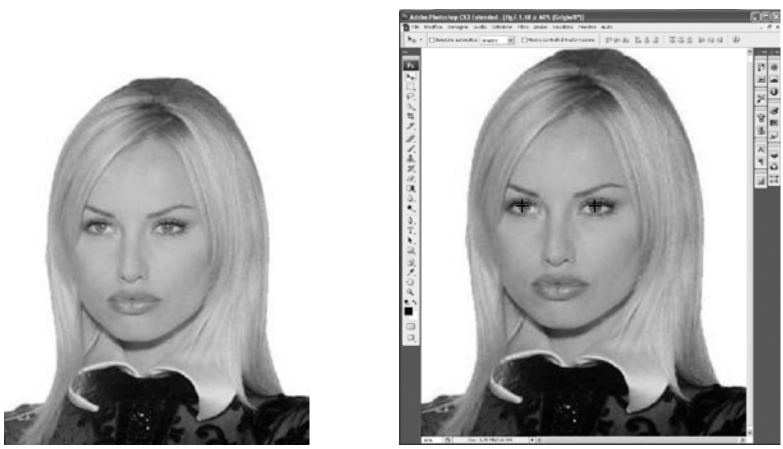
Filtering according to this formula:

$$r=[1/3*(Cb^2+(1-Cr)^2+Cb/Cr]^2$$

```
for i=1:info.Height
    for j=1:info.Width
        if r(i,j)<56
            r(i,j)=255;
        end
    end
end
```

Filtering according to the contour of the border and 9ppv.
Research of a pixel zone that forms a circle, and to get up its centre coordinates.

Figure 7. Eyes detection



Input image Eyes detection

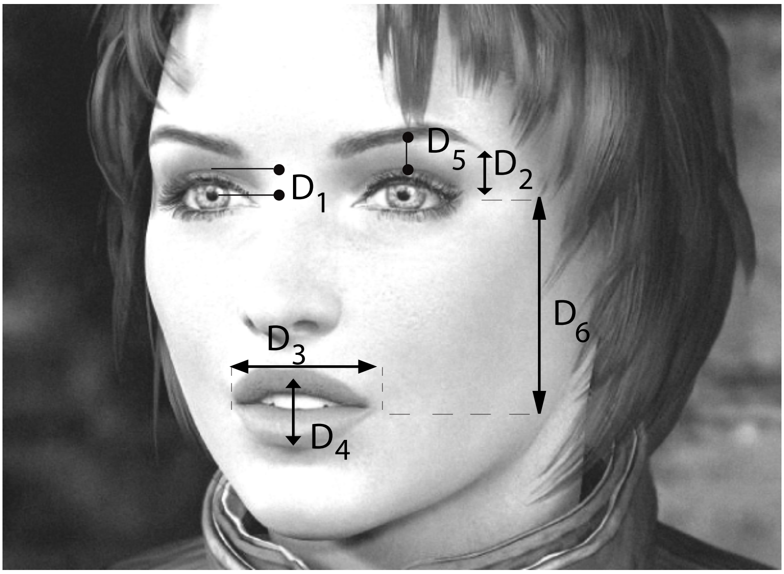


Figure 8. Definition of the distances D

The classification is based on the analysis of the distances computed on face skeletons. The distances considered make it possible to develop an expert system (for classification), which is compatible with the description MPEG-4 of the six universal emotions. Contours of the eyes, the eyebrows and the mouth are extracted automatically, by using the algorithms described in Ben Ammar et al., 2004 and 2005. The segmentation leads to obtain what we call skeleton of expression. Six distances were defined (Figure 8):

D1: opening of the eye
D2: outdistance between the interior corner of the eye and the eyebrow
D3: opening of the mouth in width
D4: opening of the mouth in height
D5: outdistance between the eye and eyebrow
D6: outdistance between the corner of the mouth and the external corner of the eye

Joy: {D4 increases}, {D3 decreases and D6 decreases}, {the other distances remain constant}.
Sadness: {D2 increases and D5 decreases}, {D1 decreases}, {the other distances remain constant}.

Anger: {D2 decreases}, {D1 increases}, {D4 either decreases or increases}.
Fear: {D2 increases and D5 increases, but more than D2}.
Disgust: {D3 increases and D4 increases}, {the other distances remain constant}.
Surprise: {D2 increases}, {D1 increases}, {D4 increase}, {the other distances remain constant}.
Table 1 gives a script of evolution of the distance D for the six emotions (↑ means increase, ↓ means decrease, and = translates the absence of evolution). Notice that for the Fear, we do not make any hypothesis on the evolution of D1, because we do not know how to translate the condition {eyes are contracted, and in state of alert} (Figure 9).

Table 1. D evolution for every emotion

	D1	D2	D3	D4	D5	D6
Joy	=	=	↑	↑	=	↓
Sadness	↓	↑	=	=	↓	=
Anger	↑	↓	=	↑ or ↓	=	=
Fear	?	↑	=	=	↑	=
Disgust	=	=	↑	↑	=	=
Surprise	↑	↑	=	↑	=	=

The classification of an emotion is based on the temporal evolution of the information, contained in the "skeleton" resulting from this stage of segmentation (temporal evolution of six characteristic distances). For example, Joy and Disgust differ by the evolution of the distance D6. One notes that emotions (Joy and Surprise) differ by the evolution of distances D1, D2, D3 and D6. This permits a distinction between these two emotions.

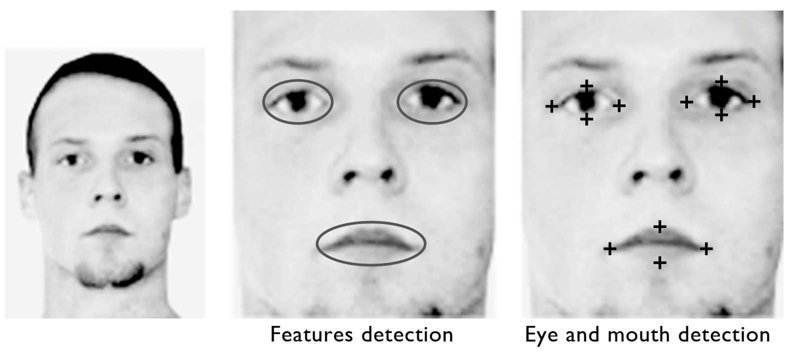


Figure 9. Eye and mouth localization

Curriculum agent

Saves the history of the learner's progression in the exercise. While analyzing the profile of the learner, this agent proposes sessions of activities subsequently to apply.
The agent curriculum keeps the trace of:

- the evolution of the interacting system with the learner;
- the history of the learner's progression in the exercise.

The agent curriculum carries out the following operations:

- to manage the model of learner throughout the training;
- to initialize the session of training, by communicating the exercise to the learners, according to their courses;
- to be the person in charge for the individualization of the training;
- to carry out the update of the history of the learner's model;
- to record, on the basis of errors, the gaps met (errors made by the learner) to help the tutor be useful itself to direct its interventions on this basis.

Tutoring agent

The tutor's role is:

- to ensure the follow-up of the training of each learner;
- to support learners in their activities;
- to support human relations and contacts between learners;
- to seek to reinforce the intrinsic motivation of the learner through its own implication as guide who shares the same objective - these interventions aim at the engagement and the persistence of the learner in the realization of training;
- to explain the method of training, and to help exceed the encountered difficulties.
- to help the learner how he/she can evaluate his/her way, needs, difficulties, rhythm and preferences.

Implementation

The interaction among agents

The interaction among human agents is not restricted to the proposed computational model. On the contrary, the computational interaction among the artificial agents aims at contributing even more to the communication and the exchange among the human agents. The interaction will be one of the main objectives of this model, because the proposal is about a model of collaborative learning. The several interaction forms, involved in the model, are: interaction among artificial agents; interaction among artificial and human agents; and interaction among human agents. In respect to communication among the human agents, the system offers tools (synchronous or asynchronous), when physical presence is not possible (for example, in the case of virtual classes).

The organizational model

Our organizational model is based on the Agent, Group and Role Meta Model - AGR for short (Ferber et al., 2003). This Meta Model is one of the frameworks proposed to define the organizational dimension of a multi-agent system, and it is well appropriate to the e-learning context. According to this model, the organization of the system is defined as a set of related groups, agents, and roles. There are several reasons, which justify the importance of this Meta Model. The main reasons are the following:

1. it is possible to construct secure systems, using groups, viewed as "black boxes", because what happens in a group cannot be seen from agents that do not belong to that group;
2. it is possible to construct dynamically components of a system, when we view system as an organization, where agents are components. Adding a new group or playing a new role may be seen as a plug-in process, where a component is integrated into a system;
3. semantic interoperability may be guaranteed using roles, because a role describes the constraints (obligations, requirements, and skills) that an agent will have to satisfy.

Implementation

We programmed agents used in the EMASPEL Framework (figure 10) with the MadKit Platform (Ferber, Gutknecht, 1998). MadKit is a modular and scalable multi-agents platform, written in Java, and built upon the AGR (Agent/Group/Role) organizational model: agents are situated in groups, and play roles. MadKit allows high heterogeneity in agent architectures and communication languages, and various customizations. In fact, MadKit does not enforce any consideration about the internal structure of agents, thus allowing a developer to freely implement his/her own agent architecture. Communication among agents is implemented by a set of communication primitives, which is a subset of FIPA-ACL (FIPA, 2004), extended with specific primitives. We used the JXTA Framework (Gong, 2002) to build an open source P2P network.

Figure 10. EMASPEL Framework

Conclusions and further work

The emotion analysis may reveal whether the learner feels "satisfaction", "confidence", "surprise", "confusion", or "frustration". These states are more precise in educational context, and appropriate pedagogical actions can be taken, in order to influence those emotions. Another important process is the diagnosis of the analyzed emotional state. This process determines the possible causes, which have led to this situation (success/failure in an exercise, difficulty of the tasks, lack of knowledge, incorrect command of the knowledge, etc.). This is done using the learner's cognitive state, and the history of his/her actions.

Showing emotions, empathy and understanding, through facial expressions and body language, is central to human interaction. More recently, emotions have also been linked closely with decision making, problem solving and intelligence in general. We therefore argue that computer-based communication technologies ought to emulate this in some way. We have conducted an experimental study on visualization, and recognition of emotion, in the human face, and an animated face. The study used six "universal" facial expressions of emotion, as established by Ekman: happiness, surprise, anger, fear, sadness, disgust, together with the neutral expression (Ekman, 1999). Results show that emotions can be visualized with a limited number of facial features, and build a potentially strong basis for communication in collaborative environments. To further establish the possible role emotions can play in collaborative environments, we are currently concentrating on real-time interaction. A group of people enters the virtual space, and is assigned a task to complete.

The main objective of the experiment is to investigate how the perception of emotion affects individual and group experience in the virtual world. From the real world, we know that emotions of others influence us in our decisions, and in our own emotional state. Emotions can motivate, encourage, help us achieve things. But they can also change our mood to the negative, make us feel angry or sad, when someone else feels that way. Emotions are contagious, and their contagious nature in the real world is potentially transferable, and beneficial to the virtual world.

The proposed framework mainly includes the peer-to-peer based network platform; for further work, we would like to:

- integrate the peer-to-peer based network platform into a grid system: the newly emerged Peer to Peer (P2P) and grid computing will serve as the key-driven forces to bring revolutionary impact on the future society;
- standardize the e-learning materials: we will implement the SCORM specification (ADL Technical Team, 2004) to describe the educational resources in EMASPEL, which will provide the interoperability with other systems, as well as the reusability of learning materials.

References

All URLs checked December 2007

ADL Technical Team (2004), *SCORM-Sharable Content Object Reference Model*

Ahn Hyung-il, Picard Rosalind W. (2005), *Affective-cognitive learning and decision making: a motivational reward framework for affective agents*, 1st International Conference on Affective Computing and Intelligent Interaction, Beijing, China, 22-24 October

Alexander Samuel T. V., Sarrafzadeh Abdolhossein, Hill Stephen (2006), *Easy with Eve: a functional affective tutoring system*, in Rebolledo-Mendez Genaro, Martinez-Miron Erika (eds.), *Workshop on Motivational and Affective Issues in ITS*. Proceedings of the 8th International Conference on ITS, pp. 38-45

Bailenson Jeremy N., Beall Andrew C., Blascovich Jim, Raimundo Mike, Weishbush Max (2000), *Intelligent agents who wear your face: user's reactions to the virtual self*, Research Center for Virtual Environments and Behavior, Department of Psychology, University of California, Santa Barbara, CA, USA [Technical Report]

Ben Ammar Mohammed, Neji Mahmoud, Gouardères Guy (2006), *Affective e-learning framework*, in Reeves T. C., Yamashita S. F. (eds.). Proceedings of World Conference on E-learning in Corporate, Government, Healthcare and Higher Education, Chesapeake, VA, USA, AACE, pp. 1595-1602

Ben Ammar Mohammed, Neji Mahmoud, Gouardères Guy (2006), *Conversational embodied peer agents in affective e-learning*, in Rebolledo-Mendez Genaro, Martinez-Miron Erika (eds.), *Workshop on Motivational and Affective Issues in ITS*, 8th International Conference on ITS, pp. 29-37.

Ben Ammar Mohammed, Neji Mahmoud, Alimi Adel M. (2005), *The integration of an emotional system in the Intelligent Tutoring System*. Proceedings of the 3rd ACS/IEEE International Conference on Computer Systems and Applications, Cairo, Egypt, 3-6 January, pp. 145-vii

Ben Ammar Mohammed, Neji Mahmoud, Alimi Adel M. (2004), *Emotion expression function in human-computer interaction*, Medianet, Tozeur, Tunisia

Bilvi Massimo, De Carolis Berardina, Carofiglio Valeria, Pelachaud Catherine (2002), *APML: a mark-up language for believable behavior generation,* in *Embodied conversational agents - let's specify and evaluate them!* Proceedings of the 1st International Joint Conference on AAMAS-Autonomous Agents and Multi-Agent Systems Workshop, Bologna, IT, 15-19 July

Biström Johnny (2005), *Peer-to-peer networks as collaborative learning environments*, Seminar on Internetworking http://people.arcada.fi/~johnny/collp2p.pdf

Blomqvist Ulf, Sjödin Christian (2006), *From peer learning to P2P learning - new life or obsolescence for the traditional learning institutions?*, in *Exploiting the knowledge economy: issues, applications, case studies*, Cunningham Paul, Cunningham Miriam (eds.), Amsterdam, IOS Press

Blumberg Bruce, Downie Marc, Ivanov Marc, Berlin Matt, Johnson Michael Patrick, Tomlinson Bill (2002), *Integrated learning for interactive synthetic characters*. Proceedings of ACM SIGGRAPH, San Antonio, TX, USA, 21-26 July, pp. 417-426

Burke Robert, Isla Damian, Downie Marc, Ivanov Yuri, Blumberg Bruce (2001), *CreatureSmarts: the art and architecture of a virtual brain*. Proceedings of the Game Developers Conference, San José, CA, USA, pp. 147–166

Chan W., Wong W.S. (2004), *Literature survey: relevance vector machine for face detection and face model feature extraction*, Delft University of Technology, Delft, NL

Coles Gerald (2004), *Reading lessons: the debate over literacy*, New York, Hill & Wang

Conati Cristina, Zhou Xiaoming (2002), *Modelling students' emotions from cognitive appraisal in educational games*. Proceedings of 6th International Conference on ITS-Intelligent Tutoring Systems, Biarritz, France, San Sebastian, Spain, 2-7 June, pp. 944-954

Conati Cristina, Klawe Maria (2000), *Socially intelligent agents to improve the effectiveness of educational games*. Proceedings of AAAI Fall Symposium "Socially intelligent agents: the human in the loop", North Falmouth, USA, AAAI Press, pp. 29-34 http://www.cs.ubc.ca/~conati/publications.html

Cowie Roddy, Douglas-Cowie Ellen, Taylor John T., Ioannou Spiros, Wallace Manolis, Kollias Stefanos (2005), *An intelligent system for facial emotion recognition*, IEEE ICME-International Conference on Multimedia and Expo, Special Session on Multimodal Emotion Recognition and Expressivity Analysis, National Technical University of Athens, Greece, 6-8 July

Darwin Charles (1872/1998), *The expression of the emotions in man and animals*. 1st ed. 1872, New York, Philosophical Library. 3rd ed. 1998, Paul Ekman (ed.), London, Harper Collins, New York, Oxford University Press

DeCarolis Berardina, Carofiglio Valeria, Pelachaud Catherine (2002), *From discourse plans to believable behavior generation*. Proceedings of International Natural Language Generation Conference, New York, NY, USA, 1-3 July

De Jongh Edwin J. (2002), *FED: an online facial expression dictionary as a first step in the creation of a complete non-verbal dictionary*, Delft University of Technology, Faculty of Information Technology and Systems, Delft, NL [Master's thesis]

ELeGI-European Learning Grid Infrastructure (2004), http://www.elegi.org

Ekman Paul (1999), *Basic emotions*, in Dalgleish T, Power T. (eds.), *The handbook of cognition and emotion*, Sussex, UK, John Wiley & Sons Ltd., pp. 45-60

Fabri Marc, Moore David J., Hobbs Dave J. (2004), *Mediating the expression of emotion in educational collaborative virtual environments: an experimental study*, "International Journal of Virtual Reality", V. 7, n. 2, pp. 66-81

Ferber Jacques, Gutknecht Olivier, Michel Fabien (2003), *From agents to organizations: an organizational view of multi-agent systems*, in Giorgini P., Müller Jörg, Odell James (eds.), *Agent-Oriented Software Engineering IV*, 4th International Workshop on Agent-Oriented Software Engineering, Melbourne, AUS, 15 July, LNCS, V. 2935, Springer-Verlag, pp. 214–230

Ferber Jacques, Gutknecht Olivier (1998), *A meta-model for the analysis and design of organizations in multi-agent systems*. Proceedings of the 3rd International Conference on Multi-Agent Systems, Paris, France, 3-8 July, pp. 128–135

FIPA-Foundation for Intelligent Physical Agents (2004), *FIPA Agent Management Specifications*, http://www.fipa.org/specs/fipa00023/

Foster Ian, Kesselman Carl (1999), *The Grid: blueprint for a new computing infrastructure*, San Francisco, CA, USA, Morgan Kaufmann Publishers

Freitas Jr Robert A. (1984), *Xenopsychology*, "Analog Science Fiction/Science Fact", V. 104, n.1, pp. 41-53

Frijda Nico H. (1986), *The emotions*, Cambridge, UK, Cambridge University Press

Gong Li (2002), *Project JXTA: A Technology Overview*, Sun Microsystems Inc. [Technical Report]

Gouardères Guy, Conté Emilie, Mansour Saber, Razmerita Liana, Antipolis Sophia (2005), *Ontology-based user modelling for personalization of grid learning services*, 1st International ELeGI Conference on Advanced Technology for Enhanced Learning, Vico Equense, Naples, IT, 15-16 March http://www.bcs.org/upload/pdf/ewic_el05_s1paper10.pdf

Gratch Jonathan, Marsella Stacy (2001), *Tears and fears: modelling emotions and emotional behaviors in synthetic agents*. Proceedings of the 5th International Conference on Autonomous Agents, Montreal, Quebec, CND, 28 May - 1 June, pp. 278-285

Hai Jin, Zuoning Yin, Xudong Yang, Fucheng Wang, Jie Ma, Hao Wang, Jiangpei Yin (2004), *APPLE: a novel P2P-based e-learning environment*, 6th International Workshop on Distributed Computing, Indian Statistical Institute, Kolkata, India, 27-30 December, LNCS, V. 3326, Springer-Verlag, pp. 52-62

Johnson W. Lewis, Rickel Jeff W., Lester James C. (2000), *Animated pedagogical agents: face-to-face interaction in interactive learning environments*, "International Journal of Artificial Intelligence in Education", V. 11, pp. 47-78

Lester James C., Towns Stuart G., FitzGerald Patrick J. (1999), *Achieving affective impact: visual emotive communication in lifelike pedagogical agents*, "International Journal of AI in Education", V. 10, n. 3-4, pp. 278-291

Liu G. (2005), *Design and implementation of a collaborative virtual learning system*, "Microcomputer Development", V. 15, n. 2, pp. 87-90

Martínez-Miranda Juan, Aldea Arantza (2005), *Emotions in human and artificial intelligence*, "Journal Computers in Human Behavior" V. 21, n.2, pp. 323–341

Nkambou Rogers (2006), *Towards affective intelligent tutoring system*, Rebolledo-Mendez Genaro, Martinez-Miron Erika (eds.), *Workshop on Motivational and Affective Issues in ITS*. Proceedings of the 8th International Conference on ITS, pp. 5-12

Piaget Jean (1989), *Les relations entre l'intelligence et l'affectivité dans le développement de l'enfant*, in Rimé Bernard, Scherer Klaus R. (eds.), *Les émotions*, Neuchâtel-Paris, Delachaux & Niestlé, pp. 75-96

Picard Rosalind W. (1997), *Affective computing*, Cambridge, MA, USA, MIT Press

Prada Ruhi, Otero Nuno, Vala André, Paiva Ana (2004), *BeLife: teaching greenhouse management using an agent based simulator*, 5th ABS-Agent Based Simulation Workshop, Lisbon, P, 3-5 May

Prendinger Helmut, Ishizuka Mitsuru (eds.) (2004), *Life-like characters, tools, affective functions, and applications*, Cognitive Technologies Series, Berlin-Heidelberg, Springer Verlag

Rao Anand S., Georgeff Michael P. (1995), *BDI-Agents: from theory to practice*. Proceedings of the 1st International Conference on Multi-Agent Systems, San Francisco, CA, USA, 12-14 June, Lesser Victor, Gasser Les (eds.), pp. 312-319

Schmidt Bernd (2000), *The modelling of human behaviour*, Ghent/Gand, SCS Publication

Silveira Ricardo Azambuja, Bica F., Viccari Rosa Maria (2000), *JADE-Java Agents for Distance Education Framework*. Proceedings of the 3rd Iberoamerican Workshop on Distributed Artificial Intelligence and Multi-Agent Systems, Atibaia, Brasil, pp.112-122

Tubbs Stewart L., Moss Sylvia (1994), *The non-verbal message, in Human communication*. 7th ed. Chapter 4, pp. 101-139, New York, McGraw-Hil

Vassileva Julita, Deters Ralph, Greer Jim, McCalla Gordon, Bull Susan, Kettel Lori (2001), *Lessons from deploying I-Help*. Proceedings of the Workshop Multi-Agent Architectures for Distributed Learning Environments, AIED, San Antonio, TX, USA, pp.3-11

VHML-Virtual Human Markup Language (2001), http://www.vhml.org

W3C-The World Wide Web Consortium (2002), http://www.w3c.org

Zapata-Rivera Juan Diego, Greer Jim (2001), *SMODEL server: student modelling in distributed multi-agent tutoring systems*, in Moore Johann D., Redfield Carol Luckhardt, Johnson Lewis W. (eds.), *Artificial Intelligence in Education*, Amsterdam, IOS Press, pp.446-455

Sintesi

L'interazione umana produce emozioni che influenzano le nostre decisioni e i nostri comportamenti, come evidenziato da numerosi studi a livello neurologico e psichiatrico. Conseguentemente, risulta possibile utilizzare le emozioni in campo educativo per facilitare l'apprendimento e stimolare la motivazione dei discenti. Quando l'apprendimento è supportato dalle nuove tecnologie di comunicazione, si parla di Affective Computing, "Computing relativo a, derivato da o che influenza deliberatamente le emozioni", in base alla definizione di Rosalind Picard del MIT Media Laboratory, tra i pionieri nello studio di questa disciplina.
Negli ultimi anni gli studi e le sperimentazioni per facilitare l'interazione tra utenti/studenti e agenti intelligenti attraverso l'uso di agenti emozionali si sono moltiplicati notevolmente. L'Emotional Embodied Conversational Agent (EECA) è un agente intelligente emozionale di nuova generazione, realizzato in via sperimentale allo scopo di arricchire l'utilizzo dell'emotional computing negli ambienti virtuali di apprendimento (Collaborative Virtual Environments o CVEs), con particolare attenzione a sistemi di peer-to-peer network e di collaborative learning. La ricerca sperimentale sugli EECA permette di realizzare e proporre le linee guida per avatar capaci di esprimere emozioni in modo efficace in conversazioni in tempo reale, esemplari ancora oggi non presenti nel panorama delle tecnologie avatar. Tali agenti sono inseriti in un sistema complessivo chiamato EMASPEL (Emotional Multi-Agents System for Peer to peer E-Learning). Nell'ambito di questo studio, il concetto di emozione è trattato in base alla definizione della psicologa Magda Arnold che distingue tra stati interiori (l'emozione vera e propria) e comportamenti che ne derivano. L'esperienza emotiva, quindi, si sviluppa in tre fasi: percezione e valutazione; emozione (interiore); azione. L'Emotion Markup Language o ELM, complemento del linguaggio XML, viene utilizzato per definire gli elementi della voce, viso e corpo che esprimono le emozioni e che devono essere tradotti in linguaggio informatico APML (Affective Presentation Markup Language). Alcune modifiche al linguaggio APML permettono agli agenti EECA di comunicare un maggior numero di informazioni in base all'intensità e durata delle espressioni del viso. Le espressioni del viso che in questo modo riescono ad essere tradotte ed interpretate con maggiore efficacia sono: rabbia, disgusto, sorpresa, felicità, tristezza e paura, più un elemento di espressione neutra che funge da riferimento per le altre. Tali espressioni corrispondono alle 6 emozioni di base individuate da Ekman alle quali è possibile far corrispondere una mimica riconoscibile del volto. Sono definiti quindi tre parametri temporali associati a ciascuna emozione: Onset (il tempo necessario perché l'emozione si manifesti); Apex (quanto dura l'emozione); Offset (dopo quanto tempo l'emozione scompare).
L'architettura studiata per accogliere agenti intelligenti emozionali è un network di peer-to-peer e-learning in cui sono integrate cinque tipologie di agenti. L'Interface Agent si occupa di trasmettere le informazioni che provengono dagli utenti/studenti agli altri agenti nel sistema. Un ruolo fondamentale è quello dell'innovativo EECA - Emotional Embodied Conversational Agent. Si tratta di un

"virtual human" in grado di esprimere le 6 emozioni di base e che, utilizzando le informazioni provenienti dall'ambiente e dall'utente, è in grado di creare una relazione con l'utente e rispondere alle sue espressioni emotive. Un simile agente si sviluppa in base a sei layers di programmazione che riproducono essenzialmente il processo di percezione e categorizzazione degli input, analisi e comprensione dello stato emotivo dell'utente e azione di risposta.

La costruzione di EECA è basata sull'architettura informatica PECS elaborata da Schmidt, che sostituisce l'architettura BDI (Beliefs, Desire, Intentions), precedentemente utilizzata per supportare agenti intelligenti. Anche in questo caso i livelli della struttura sono tre, relativi alla rilevazione dei dati esterni, alla definizione delle caratteristiche "interne" dell'agente e alla modellizzazione delle possibili risposte dell'agente.

Il terzo tipo di agente inserito nel sistema è denominato Emotional Agent e ha il compito principale di interpretare e gestire le emozioni espresse dallo studente durante le sessioni di apprendimento attraverso l'analisi delle espressioni del suo viso (emotion identification). Le fasi del riconoscimento includono la localizzazione del volto dello studente (attraverso la localizzazione dei contorni, poi della bocca e degli occhi) e la definizione delle espressioni in base ai rapporti di distanza e forma di bocca e occhi; tali rapporti sono catalogati in 6 categorie corrispondenti alle 6 emozioni base universali. Gli ultimi due agenti previsti nel sistema sono il Curriculum Agent, che analizza i risultati ottenuti dagli studenti nel tempo e li salva in una loro "storia", e il Tutoring Agent, i cui compiti consistono nel supporto, stimolo, rinforzo allo studente, in base soprattutto alle informazioni che provengono dalle attività di monitoraggio del Curriculum Agent.

In questo modo, l'analisi delle emozioni permette di rilevare i sentimenti degli studenti in merito alle proprie prestazioni e ai contenuti del corso e di rispondere a tali stati emotivi attraverso agenti intelligenti in grado di operare a livello emotivo di rinforzo o di modificazione della condotta degli studenti.

Successivi obiettivi di ricerca saranno rivolti ad integrare il sistema in un grid system e nella standardizzazione dei materiali nella codifica SCORM, allo scopo di garantire l'interoperabilità.