International Journal of Humanoid Robotics
Vol. 3, No. 3 (2006) 1–8

A DYNAMIC EMOTION REPRESENTATION MODEL WITHIN A FACIAL ANIMATION SYSTEM

EMMANUEL TANGUY*, PHILIP J. WILLIS† and JOANNA J. BRYSON‡

Department of Computer Science, University of Bath,
BA2 7AY, United Kingdom
**e.tanguy@cs.bath.ac.uk*
†p.j.willis@bath.ac.uk
‡j.j.bryson@cs.bath.ac.uk

Received 4 November 2005

This paper presents the *Dynamic Emotion Representation* (DER), and demonstrates how an instance of this model can be integrated into a facial animation system. The DER model has been implemented to enable users to create their own emotion representation. Developers can select which emotions they include and how these interact. The instance of the DER model described in this paper is composed of three layers, each representing states changing over different time scales: *behavior activations*, *emotions* and *moods*. The design of this DER is discussed with reference to emotion theories and to the needs of a facial animation system. The DER is used in our Emotionally Expressive Facial Animation System (EE-FAS) to produce emotional expressions, to select facial signals corresponding to communicative functions in relation to the emotional state of the agent and also in relation to the comparison between the emotional state and the intended meanings expressed through communicative functions.

Keywords: Emotion representations; facial animation; communicative functions.

1. Introduction

Facial expressions and emotions are linked in some undetermined loose fashion but intuitively people look for emotional signs in facial expressions. [1,2] This motivates the integration of an emotion model into a facial animation system: to select different communicative signals in relation to the emotional state of an agent producing emotionally expressive animations.

We have developed an Emotionally Expressive Facial Animation System (EE-FAS, pronounced "e-face"). This can be used to produce animations from a stream of text and XML tags. The tags carry information about communicative functions and emotional impulses. The use of communicative functions has the advantage of being flexible because they provide information about the meaning that should be communicated through the face, such as the emphasis of a word, and not *how* it should be communicated.[3] This flexibility enables us to define animation scripts that can produce different animations in relation to the characteristics of the agent

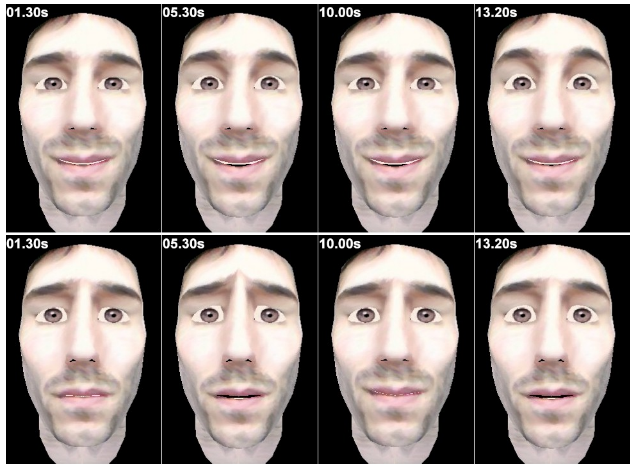

Fig. 1. Four screen-shots of two animations resulting from the same animation script played in two different emotional contexts. The top four pictures are in a context *happiness* whereas the four bottom pictures are in a context *sadness*.

such as its body, its current emotional states, its gender and its age.[3] An example of this flexibility is shown Fig. 1.

In the EE-FAS, signals corresponding to communicative functions are selected directly in relation to the emotional state, but they are also selected according to the difference between the intended communicative functions and the emotional state. By knowing the emotional state of an agent (through the emotion model) and the meaning that this agent wants to communicate (through communicative functions) it is possible to produce different facial expressions depending on whether the emotional state matches the intended meaning. For instance, if an agent wants to display a polite smile, the EE-FAS produces a different smile depending on whether the character is currently in a positive or negative emotional state. As far as we know, EE-FAS is the only facial animation system applying this technique.

To reach this objective, an emotion model is needed to represent the emotional state of the agent. Many emotion models already exist and they are used within facial animation systems,[4–6] as well as for influencing action selection within autonomous agents architectures.[6–12]

We propose that the term *emotion model* should be defined as containing two components: **mechanisms eliciting emotions** from external and internal stimuli, including potentially the agent's own goals, beliefs and standards; and **emotion representations** keeping track of the emotional states and their changes.

The distinction between these two parts is useful because the mechanisms eliciting emotions can produce an identical assessment of the same event but this event can have different emotional repercussions depending on the current state of the emotion representation. For instance, breaking a glass would generally be assessed as a negative event but the effect of this event would be different according to the emotional state of the person. For a person who is already angry, this event could infuriate him/her, whereas for a person having a good time, this event might just reduce his/her level of happiness.

This paper presents a framework to build Dynamic Emotion Representations (DER). Our work differs from existing work because it focuses on the emotion representation of emotion models and on the dynamics of emotions. The term *dynamic* is used to point out that emotions are ongoing, durative, constantly changing states and that emotions interact and influence each other. We have also implemented an instance of this model incorporating three types of emotional states, and have integrated this instance into a full animation system. This paper concentrates just on the basic structure.

2. A Dynamic Emotion Representation

2.1. *DER model: Overview of the working process*

This paper distinguishes between mechanisms eliciting emotions and emotion representations. *Emotional impulses* are generated from mechanisms eliciting emotions and fed to the emotion representations. We define *emotional impulses* with two components: a name representing an emotion type and an intensity. In existing work, vectors of emotional impulses are produced by mechanisms eliciting emotions; indexes of a vector specify types of emotions and values in the vector are impulse intensities.[5,6] Our definition of emotional impulse can be seen as a simplified version of the "emotional structure" definition described by Reilly, which contains information such as an intensity, an emotion type, and the cause of this emotion.[8]

The DER maintains lasting states, in particular, but not exclusively, emotional states which are influenced by emotional impulses. Emotional spaces can be represented as discreet spaces or as continuous spaces. The most common discreet representation is the six basic emotions.[13–18] Continuous spaces are also used to represent emotional states, such as those described by the dimensions *activation* and *evaluation*,[14] the dimensions *pleasure* and *arousal*,[2] or the dimensions *energy* and *tension*.[19]

The DER can represent any number of emotional spaces and emotional spaces can be represented by one or several dimensions. A *dimension* represents intensity variations of a variable. For instance, a dimension can represent the variation of anger intensity. In the DER, no distinction is made between the representations of discreet and continuous emotional spaces; both types of emotional space are represented as a list of dimensions.

Each emotional impulse can affect any dimensions represented by the DER. The effects of an emotional impulse are controlled by dynamic filters associated with each dimension. A *dynamic filter* is a function with parameters changing according to the intensity of the other dimensions in the DER. This mechanism enables the designer of a DER to control the effects of emotional impulses on each dimension, effects that depend on the current state of the DER. The architecture of the DER is configured using an XML file.[20]

2.2. *Overview of a three-layer DER architecture*

We typically design a DER with three layers: *behavior activations, emotions and moods*. Emotion types can be categorized according to the processes which elicit them, such as primary, secondary and tertiary emotions which are respectively generated by fast, innate reactive processes, deliberative processes and meta-management mechanisms providing awareness of internal states.[21] One characteristic of primary emotions is to trigger pre-organized behaviors. The first layer of the DER represents behavior activations which can be used to provoke fast reactions of virtual actors to changes in their environment. Affective states are also distinguished by their durations. For instance, emotions, which last from under a minute to few a minutes, are distinguished from moods, which last for hours, days or weeks.[19,22,23] The representation of these two affective states, emotions and moods, provides two emotional contexts with two different time scales. The second layer of the DER represents emotions as durative states changing on a short-to-medium time scale. The state represented by the second layer is used to influence action selection mechanisms and the expressivity of virtual actors. The third layer represents moods that vary on a medium-to-long time scale. As suggested by Picard,[24] in the DER moods influence the effects of emotional impulses on other emotions, e.g. other dimensions. For instance, "A bad mood can make it easier for negative-valenced emotion to be activated, while a good mood makes this more difficult" (Ref. 24, p. 155). A practical example in the DER is the effect of the mood state, among others, on the activations of behaviors. Emotional impulses might or might not activate their corresponding behaviors according to the mood state and to the intensity of other dimensions.

These three layers of the DER reflect some characteristics of affective states but they are not directly related to particular types of emotions or mechanisms eliciting emotions. Particular mechanisms produce certain types of emotional impulses which can affect any layers of the DER according to the configuration of the dynamic filters. For instance, emotional impulses related to primary emotions, generated by reactive mechanisms, should generally affect the first layer to trigger pre-organized behaviors but they can also produce durable states by affecting the second layer.

3. A Three-Layer DER in the EE-FAS

The Emotionally Expressive Facial Animation System (EE-FAS) includes a DER to produce and select facial expressions. This section presents in more detail the composition of the three layers of the DER used in the EE-FAS.

3.1. *Emotional and communicative facial expressions*

Two main types of facial expressions are distinguished: those produced by emotional events or episodes,[13,16–18] called emotional expressions, and those produced by communicative processes.[2,3] Bavelas and Chovil emphasize that expressions produced from emotional episodes are different from the expressions produced as emotional communicative functions, e.g. *personal reaction*. (Ref. 2, Chap. 15). The EE-FAS distinguishes and produces both types of facial expressions: emotional and communicative.

3.2. *Emotional expressions and behavior activation layer*

In the EE-FAS, emotional expressions are the six universally recognized facial expressions associated to six basic emotions.[17,18] Izard suggests that expressions of emotions are triggered at the beginning of an emotional episode, or at the "emotion activation" (Ref. 2, Chap. 3). To simulate this, the first layer of the EE-FAS DER includes six behavior activations corresponding to six basic emotions: *happy*, *angry*, *sad*, *surprised*, *afraid*, *disgusted*. Thus, when an emotional impulse corresponding to a basic emotion is elicited, a pre-organized behavior can be triggered, which will be displayed by the EE-FAS as one of the six universally recognized expressions.

Emotional impulses, such as those corresponding to basic emotions, can trigger behaviors but they can also produce durative states, emotions, represented by the second layer of the DER.

3.3. *Communicative functions and emotion layer*

The second layer of the DER represents emotions, which are durative states resulting from emotional impulses. In the EE-FAS, the six basic emotions[17,18] are represented in this layer, but more emotions or different emotions can be represented.

The representation of emotions in the second layer of the DER is used as an emotional context and a consistency mechanism to select facial signals corresponding to communicative functions. Communicative functions are meanings that are intentionally communicated. Such a meaning can be expressed by different facial signals, different visual representations. Depending on the virtual actor's emotional state, in particular the state of the second layer of the DER, different facial signals of a communicative function can be selected. This mechanism is one solution to the problem of mapping a small number of facial meanings to a large number of facial signals (Ref. 25, p. 1). By comparing the emotional state of the virtual actor and the communicative function intended, the EE-FAS can also distinguish contexts in which genuine or fake facial expressions should be displayed.

3.4. *Mood layer*

Mood is often represented by one dimension: good/bad.[8,24] However, in our default system, mood is represented by two dimensions, *calm/tense* and *energy/tiredness* —

a model based on the theory of mood described by Thayer.[19] Using these two dimensions, Thayer proposes four mood states: *energetic-calm* (high energy and low tension) is the optimum mood; *energetic-tense* (high energy and high tension) is a mood that enables people to be active and to do what has to be done; *tired-calm* (low energy and low tension) is a relaxed state such as before sleep but a person in this state is also very sensitive to tension; and *tired-tense*, (low energy and high tension) is the worst state when the energy is insufficient to do what has to be done.

Thayer's mood model is a more plausible model of moods in a natural system than models based on one dimension good/bad, because its two dimensions not only reflect variations of the emotional state but also variations of general physical and mental states. The term *energy* includes mental energy and physical energy, such as the extent to which a person is awake, her blood sugar level, or her health. The energy level follows a daily biological pattern.[19] The tension level varies according to the pressure of the environment, such as threats, task accomplishments and maybe reproduction instinct. Mood is described as an overview of a person's mental and physical state, not simply an overview of the emotional state.[19,22] Models based on one dimension good/bad are only a reflection of the emotional state.

By relating Thayer's model to existing work on computational emotion models[7,10], links can be built between drives, such as hunger or thirst, and energy and tension. For instance, hunger is related to the level of energy provided by food and the need for food produces some pressure on the system to fulfil this need, which can be interpreted as tension. Representing mood on the basis of Thayer's model would be closer to mood theories of natural systems and could also be integrated with existing autonomous virtual agent systems.

It is possible to reconcile Thayer's model with the suggestion of Picard to use positive or negative mood state to influence the effects of emotional stimuli.[24] Thayer's model does not represent directly a dimension good/bad or positive/negative but using his description of the four mood states two new dimensions appear:

- *pleasure/displeasure* represented by the extremes: *energy-calm* (high energy and low tension), and *tired-tense* (low energy and high tension);
- *sleep/arousal* represented by the extremes: *tired-calm* (low energy and low tension), and *energy-tense* (high energy and high tension).

A positive or negative value of mood can be computed as the difference between the level of energy and the level of tension. If the tension is superior to the energy the mood is qualified as negative, in the reverse case the mood is qualified as positive.

4. Conclusion

This paper introduces the basis of the structure of the Dynamic Emotion Representation (DER) model, and briefly describes an instance of it developed for a facial animation system. Animations from this system are available from the web and more detail on the system can be found elsewhere.[20] The design of a three-layer DER is inspired by emotion theories and guided by the needs of the facial animation system in which it is integrated. The first layer represents *behavior activations* and is used to produce coherent emotional expressions. The second layer, representing a short-to-medium term emotional context, influences the selection of facial signals corresponding to communicative functions. The *mood layer* is used as a medium-to-long term emotional context and also influences the effects of emotional stimuli on the representations.

The DER integrated within an Emotionally Expressive Facial Animation System (EE-FAS) enables the production of subtle differences in facial expressions in relation to the agent's emotional state. Facial signals corresponding to communicative functions are selected in relation to the emotional state of the agent but also in relation to the difference between the emotional state and the intended communicative functions. The DER provides for more varied emotional behavior than most AI emotional models because the agent's history determines its full emotional state and response, yet it also creates plausible, human-like emotional behavior.

Acknowledgments

All three authors wish to acknowledge that the first author (E. Tanguy) was solely responsible for the coding of the DER model and most of its design. This work is funded by a studentship from the Department of Computer Science, University of Bath and the ESPRC Grant GR/S/79299/01.

References

1. P. Ekman and W. V. Friesen, *The Repertoire of Non-Verbal Behaviour Categories, Origins Usage and Coding*, Semiotica 1 (Pergamon Press, 1969), pp. 49–98.
2. J. A. Russel and J. M. Fernández-Dols, *The Psychology of Facial Expression* (Cambridge University Press, 1997).
3. C. Pelachaud and M. Bilvi, Computational model of believable conversational agents, in *Communication in Multiagent Systems*, ed. M.-P. Huget (Springer, 2003), pp. 300–317.
4. I. Albrecht, *Faces and Hands: Modeling and Animating Anatomical and Photorealistic Models with Regard to the Communicative Competence of Virtual Humans*, PhD Thesis, Universität des Saarlandes (December, 2005).
5. A. Egges, S. Kshirsagar and N. Magnenat-Thalmann, Generic personality and emotion simulation for conversational agents, *Comput. Animat. Virt. W.* **15** (2004) 1–13.
6. T. Duy Bui, *Creating Emotions and Facial Expressions for Embodied Agents*, PhD thesis, University of Twente (2004).
7. D. Cañamero, Designing emotions for activity selection in autonomous agents, in *Emotions in Humans and Artifacts*, eds. R. Trappl, P. Petta and S. Payr (MIT Press, 2003), pp. 115–148.
8. W. Scott Neal Reilly, *Believable Social and Emotional Agents*, Phd Thesis, School of Computer Science, Carnegie Mellon University, Pittsburgh (1996).

9. J. Gratch and S.Marsella, Evaluating the modeling and use of emotion in virtual humans, in *Proc. 3rd Int. Joint Conf. Autonomous Agents and Multiagent Systems*, Vol. 1 (IEEE Computer Society, 2004), pp. 320–327.
10. J. D. Velásquez, Modeling emotions and other motivations in synthetic agents, in *AAAI/IAAI* (1997), pp. 10–15.
11. E. André, M. Klesen, P. Gebhard, S. Allen and T. Rist, Integrating models of personality and emotions into lifelike character, in *Proc. Workshop on Affect in Interaction — Towards a new Generation of Interfaces*, Siena, Italy, October (1999), pp. 136–149.
12. A. Paiva, J. Dias, D. Sobral, R. Aylett, P. Sobreperez, S. Woods, C. Zoll and L. Hall, Caring for agents and agents that care: Building empathic relations with synthetic agents, in *Proc. 3rd Int. Joint Conf. Autonomous Agents and Multiagent Systems*, Washington, DC, USA (IEEE Computer Society, 2004), pp. 194–201.
13. C. Darwin, *Expression of the Emotions in Man and Animals* (Friedmann, 1979).
14. R. Plutchik, A general psychoevolutionary theory of emotion, in *Emotion: Theory, Research, and Experience*, eds. R. Plutchik and H. Kellerman (Academic press, 1980), pp. 3–33.
15. C. E. Izard, *Human Emotions* (Plenum Press, 1977).
16. S. S. Tomkins, Affect as amplification: Some modification in theory, in *Emotion: Theory, Research, and Experience*, eds. R. Plutchik and H. Kellerman (Academic Press, 1980), pp. 3–33.
17. P. Ekman, Facial expressions of emotion: New findings, new questions, *Psychol. Sci.*, **3** (1992) 34–38.
18. C. E. Izard, *The Face of Emotion* (Appleton-Century-Crofts, 1971).
19. R. E. Thayer, *The Origin of Everyday Moods* (Oxford University Press, 1996).
20. E. A. R. Tanguy, J. J. Bryson and P. J. Willis, A dynamic emotion representation model within a facial animation system, Tech. Rep. CSBU-2005-14, Department of Computer Science, University of Bath (November, 2005), www.cs.bath.ac.uk/ cspeart/publication/techReport/DERModel_tecRe p.pdf.
21. A. Sloman, How many separately evolved emotional beasties live within us?, in *Emotions in Humans and Artifacts* (MIT Press, 2003), pp. 35–114.
22. L. Sizer, Towards a computational theory of mood, *British J. Phil. Sci.* **51**(4) (2000) 743–769.
23. K. Oatley and J. M. Jenkins, *Understanding Emotion* (Blackwell Publishers Ltd, 1996).
24. R. W. Picard, *Affective Computing* (MIT Press, Cambridge, Massachusetts, 1997).
25. Aldo Paradiso, An algebra for combining MPEG-4 compliant facial animations, in *Proc. Int. Workshop Lifelike Animated Agents: Tools, Affective Functions, and Applications*, 2002. Tokyo, Japan, 19 August, ed. H. Prendinger (2002) [Online].