IGI PUBLISHING

701 E. Chocolate Avenue, Hershey PA 17033-1240, USA
Tel: 717/533-8845; Fax 717/533-8661; URL-http://www.igi-global.com

ITJ 5588

This paper appears in the publication, International Journal of Synthetic Emotions, Volume 1, Issue 1
edited by **Jordi Vallverdú © 2010, IGI Global**

Simplifying the Design of Human-Like Behaviour:

Emotions as Durative Dynamic State for Action Selection

Joanna J. Bryson, University of Bath, UK

Emmanuel Tanguy, University of Bath, UK

ABSTRACT

Human intelligence requires decades of full-time training before it can be reliably utilised in modern economies. In contrast, AI agents must be made reliable but interesting in relatively short order. Realistic emotion representations are one way to ensure that even relatively simple specifications of agent behaviour will be expressed with engaging variation, and those social and temporal contexts can be tracked and responded to appropriately. We describe a representation system for maintaining an interacting set of durative states to replicate emotional control. Our model, the Dynamic Emotion Representation (DER), integrates emotional responses and keeps track of emotion intensities changing over time. The developer can specify an interacting network of emotional states with appropriate onsets, sustains and decays. The levels of these states can be used as input for action selection, including emotional expression. We present both a general representational framework and a specific instance of a DER network constructed for a virtual character. The character's DER uses three types of emotional state as classified by duration timescales, keeping with current emotional theory. We demonstrate the system with a virtual actor. We also demonstrate how even a simplified version of this representation can improve goal arbitration in autonomous agents.

Keywords: Durative States, Dynamic Emotion Representation, Goal Arbitration

INTRODUCTION

Emotion is a popular topic in AI research, but most existing work focuses on the appraisal of emotions or mimicking their expression for HCI (see review below). Our research is concerned with their role in evolved action-selection mechanisms. In nature, emotions provide decision state which serves as a context for limiting the scope of search for action selection (LeDoux, 1996). This state is sustained more briefly than traditional (life-long) learning, but longer than simple reactive responses.

Improving the realism of emotion representations can allow us to not only improve the realism of intelligent virtual actors, but also to make programming them easier. Rather than needing to describe the exact details of a

DOI: 10.4018/jse.2010101603

facial expression, a behaviour script can simply specify abstract concepts like *emphasis* or intentional, communicative expression gestures such as *smile in greeting*. In real-time, the agent can then interpolate these instructions with its current emotional state. This latter in turn reflects the agent's recent experiences. For example, a FAQ agent that has just been accessed might respond with more apparent enthusiasm than one that has been interacting with a client and receiving verbal abuse (Brahnam & De Angeli, 2008). For commercial applications this is often desirable, since companies do not want to be represented by "stupid" agents. Another place where real-time emotion tracking is useful is for home assistance agents. Instructions (e.g., to take medication or remind the user that the stove is on) need to be reliable and clear, yet they cannot be always presented identically or else even patients with severely compromised short-term memory can become habituated. Using recent interaction history as a seed to vary delivery style is one mechanism for maintaining variation in presentation style, as well as potentially increasing user engagement.

To this end, we have developed mechanisms for modelling both the temporal course of emotional state and the interactions between such states. We have an elaborate model for complex, human-like emotions for generating realistic facial expressions, the Dynamic Emotion Representation (DER) (Tanguy, Willis, & Bryson, 2003; Tanguy, 2006). For applications with less demand for emotional complexity, we also present a simplified system called Flexible Latching. This provides basic goal arbitration as a part of an action selection mechanism without requiring as much programming. Both systems track systems of emotion and/or drive intensities which change and interact over time. The actions and emotional response of agents containing such durative-state models depends on their recent history as well as their individual priorities or personality and their environment.

Our durative-state systems assume other independent mechanisms for appraising the agent's situation and expressing the emotional responses. Developers using our representations can specify and describe both the number and the attributes of fundamental emotions and express how they interact. In this respect, these emotion representation systems are similar to spreading activation action-selection systems (e.g., Maes, 1991). They are designed to be the root of an agent's action selection, determining the current goal structure. Note that in this fully modular system, additional "higher order" emotions may either be interpreted as emerging from the interactions of fundamental emotions, or they can be introduced with explicit representations—the choice is left to the developer.

We begin this article with a review of the concepts and literature. We next give a detailed description of the relatively complex mechanism, the DER, capable of producing biomimetic human-like emotions. We then describe for the more basic action-selection aspects of goal arbitration Flexible Latching. Finally, we describe full implementations of each mechanism demonstrating their roles as parts of complete systems.

BACKGROUND

Action Selection and Durative State

Action selection is one of the fundamental problems of intelligence (Prescott, Bryson, & Seth, 2007). For an agent (whether biological or artificial), *action selection* is the ongoing problem of choosing what to do next. For a developer, action selection presents two problems:

- designing the agent's action-selection process, and
- determining the level of abstraction at which the process will operate.

A physical agent must ultimately perform precise control of motors or muscles, but this is almost certainly not the level at which decision making should occur. We know animals that have lost their higher cognitive capacity (e.g., lost their forebrains) can still perform species-typical behaviours, although they may not be performed in appropriate contexts (Carlson, 2000). Also in animals, complex actions such as grasping or moving a hand to a mouth can be triggered by stimulating individual nerve cells in associative cortecies (Graziano, Taylor, Moore, & Cooke, 2002), while targeted scratching of irritants can be controlled from the spine (Bizzi, Giszter, Loeb, Mussa-Ivaldi, & Saltiel, 1995).

Artificial action selection also normally operates on a set of pre-defined primitive acts (Bryson, 2000). Even where these actions are acquired through machine learning, they still tend to be segmented into discrete actions, gestures or target postures which can then be reassembled in a desired order (Brand, 2001; Schaal, Ijspeert, & Billard, 2004; Whiteson, Taylor, & Stone, 2007; Wood & Bryson, 2006).

To date, *durative state* like emotions and drives has not been adequately incorporated in standard action-selection architectures (Bryson, 2008). We call this state “durative” in contrast to long-term learning and to the transient state associated with real-time decisions. Certainly many (if not most) current AI architectures do address emotions and/or drives in some way. However, few current best-practice techniques of action selection include fully integrated emotion systems. Some systems use emotions as relatively isolated systems, essentially social effectors for human-robot interaction (Breazeal & Scassellati, 1999; De Rosis, Pelachaud, Poggi, Carofiglio, & De Carolis, 2003; Velásquez & Maes, 1997). An outstanding exception from within this general approach is from Marcella and Gratch (2002), whose emotional system does affect action selection by changing their military-training system’s basic reasoning to better simulate civilian reactions to troops.

Most AI research that does postulate agent-centric utility for affect has focussed on it as a sort of reward accumulator. Emotions serve as additional pieces of internal state to assist in learning and applying action selection policies (Broekens, Kosters, & Verbeek, 2007; Gadanho, 1999; Hiller, 1995; Zadeh, Shouraki, & Halavati, 2006). Some of this research has been inspired partly by the controversial somatic marker hypothesis (Tomb, Hauser, Deldin, & Caramazza, 2002).

Several elaborate architectures have been proposed but not yet constructed which postulate similar referential or marker roles for emotional systems, but operating at a higher self-referential level (Minsky, Singh, & Sloman, 2004; Norman, Ortony, & Russell, 2003; Sloman & Croucher, 1981). Such elaborate theories of the role of emotion in metacognition are beyond the scope of this article. We discuss basic action selection, not reflective reasoning.

We believe emotions and drives are absolutely integral to action selection, determining the current focus of behaviour attention. A similar perspective is taken by Morgado and Gaspar (2005), but these mechanisms are not incorporated into a full architecture. At the other extreme of implementation detail, Breazeal (2003) like us treats both moods and drives as essential mechanisms for maintaining homeostatic goals. She has an extremely elaborate robotic system built around her architecture. Here we will emphasise usability considerations as a general-purpose element of animal-mimetic AI systems.

Representing Emotional State

Computational emotion models should include two parts:

- mechanisms eliciting emotions from external and internal stimuli, including potentially the agent’s own goals, beliefs and standards;
- emotion representations keeping track of the emotional states and their changes over time.

In the design of emotion models the distinction between mechanisms eliciting emotions and emotion representations is useful; the assessment of an emotional event can be the same but its impact on the actions and future emotional state of the virtual actor can vary according to its *current* emotional state. For instance the event of knocking over a cup of tea might make somebody already angry lose their temper, whereas if this person was happy in the first place this negative event might have little impact, just a slight reduction of happiness. An appropriate emotion representation can enable programmers to reduce the complexity of mechanisms eliciting emotional responses by allowing them to assess an identical event in the same way, regardless of context.

Most existing emotion theories are concerned with mechanisms eliciting emotions (Frijda, 1986; Izard, 1993; Lazarus, 1991; Plutchik, 1980; Sloman, 2003). In contrast, the duration of emotions and the way different emotions interact are not the focus of much research. The same imbalance is found in computational models of emotions—the focus is on the mechanisms eliciting emotions and on their expression, but their representation is typically trivial.

But if we want to use emotions and other durative state in a biomimetic way—as a key part of action selection—we need something more. Thus we recommend a modular and temporal approach to representing durative state.

- Since emotions and drives are integral to an agent's motivation system, their number should be flexible and correlated to the tasks the agent needs to perform. This includes social goals such as achieving agreement with a user. Thus a good durative state system should be able to represent any number of persisting states, such as moods, emotions and drives.
- A single event may have multiple different consequences for different goals and therefore different emotions. Thus an appraised emotional impulse should be able to affect more than one state, and do so positively or negatively depending on their interactions.
- Similarly, the state of some goals may have ramifications on others, so they should also be able to have impact on others. At the same time, we do not necessarily want to fully connect all durative state, since that could lead to an impossibly complex system to develop. Rather, connections between durative state modules should be optional and configurable.
- Finally, a complex agent requires not only AI, but also the easy expression of developer intelligence (Bryson & Stein, 2001). Thus our approaches always focus on usability. For example the DER (described in the next section) is configured by using an XML file.

The Dynamic Emotion Representation

We describe here an elaborate and powerful system for creating realistic emotions for real-time, human-like agents such as robots or VR avatars. This is the Dynamic Emotion Representation (DER). Each emotion in the DER is described with characteristic intervals of onset, sustain and decay, and each emotion may either excite or inhibit any other.

Besides simplifying the emotion elicitation process, the DER makes it easier to generate variety in the behaviour of real-time interactive character-based AI systems, since the same stimuli can result in significantly different (but locally coherent) responses, depending on the agent's emotional state. The DER also greatly simplifies scripting for virtual actors by decomposing the problems of specifying the emotionally salient events from describing the agent's individual reaction to or expression of characteristic emotional states. For instance, a script could specify the general action of grabbing an object. Detailed characteristics of this action, such as the location and effort of the grasp (Badler, Allbeck, Zhao, & Byun, 2002; Blumberg, 1996), could be modified per the current DER state, producing different animations and implying different interpretations. The DER presents a powerful mechanism for specifying virtual agent personalities, by allowing developers to (for example) specify characteristic moods or other emotion-related attributes such as tension level.

The DER was inspired by a variety of appraisal theories (Izard, 1993; Frijda, 1986; Lazarus, 1991; Ortony, Clore, & Collins, 1990; Plutchik, 1980), and the pioneering work of several influential researchers (Picard, 1997; Sloman, 2003). Many other AI systems use appraisal to influence the decision processes and behaviours of virtual actors (André, Klesen, Gebhard, Allen, & Rist, 1999; Cañamero, 2003; Delgado-Mata & Aylett, 2004; Gratch & Marsella, 2004). Relatively few systems provide dynamic emotion representations on multiple time scales like the DER, and none of these are as configurable as our system. We present a more thorough review of the most closely related systems below, but first we present the DER.

Three Responses to One Series of Events

Before explaining the details of the DER we will first clarify its utility with a concrete example of its use. This example involves a specific instance of a DER system (described below) where there are representations corresponding to three different time courses:

- *behaviour activations*: action-selection impulses corresponding to states such as "happy" or "angry." These are elicited by events (either internal or perceived) and result in basic emotion-related behaviours such as smiling. Behaviour activations *trigger* pre-organised behaviours, which follow their own time course after their activations. Activations can be associated with intensity levels.
- *emotions*: such as happiness or anger. This is durative state, which varies gradually in response to events and the passage of time. Emotions provide a context influencing the current actions. The interaction between emotions and behaviour activations is complex—the emotional impulses that elicit behaviour activations can also increase compatible emotion levels and decrease incompatible ones, while the state of the emotions can influence the appraisal which elicits the activations.
- *moods*: longer-term durative state. Moods are similar to emotions, but change much more slowly. In the demonstration, we use tension and energy as interacting mood components. During the short period of the demonstration described next, they are essentially fixed environments that differentiate the various conditions. However, moods too alter as a consequence of events, though slowly, doing so in response to emotion or other internal stimuli.

We illustrate this system embedded in a facial avatar system called EE-FAS (described further below) (Tanguy, Willis, & Bryson, 2006). EE-FAS provides the actual behaviour primitives or gestures, such as the capacity for smiling; the DER determines when and how much to express smiling.

Figure 1 shows a DER in three different mood contexts responding to the same series of six elicited emotional impulses coming from an appaisal mechanism: three happiness-elicited impulses followed by three anger-elicited impulses. Graph d shows these impulses, while graph a shows the response in a negative mood (high tension, low energy), b a neutral mood, and c a positive mood (high energy, low tension). The Happy and Angry graphs have $+$ symbols on them when they result in a behaviour activation—a signal being sent to the EE-FAS system to generate a facial expression (see Figure 2). The EE-FAS contains a muscle model which results in realistic transitions between target expressions.

Figure 2 shows the EE-FAS output. The intensities of the behaviour activations determine the strength of the displayed facial expressions. However the durations of facial expressions are innate characteristics and are not limited to the duration of behaviour activations. Between each screen shot shown in Figure 2, expressions decay

Figure 1. Changes of DER states due to six emotion impulses in three contexts. Graphs a, b and c show state changes in the contexts of a Negative Mood, a Neutral Mood and a Positive Mood, respectively. Graph d shows the emotion impulses initially sent to the DER in all three contexts.

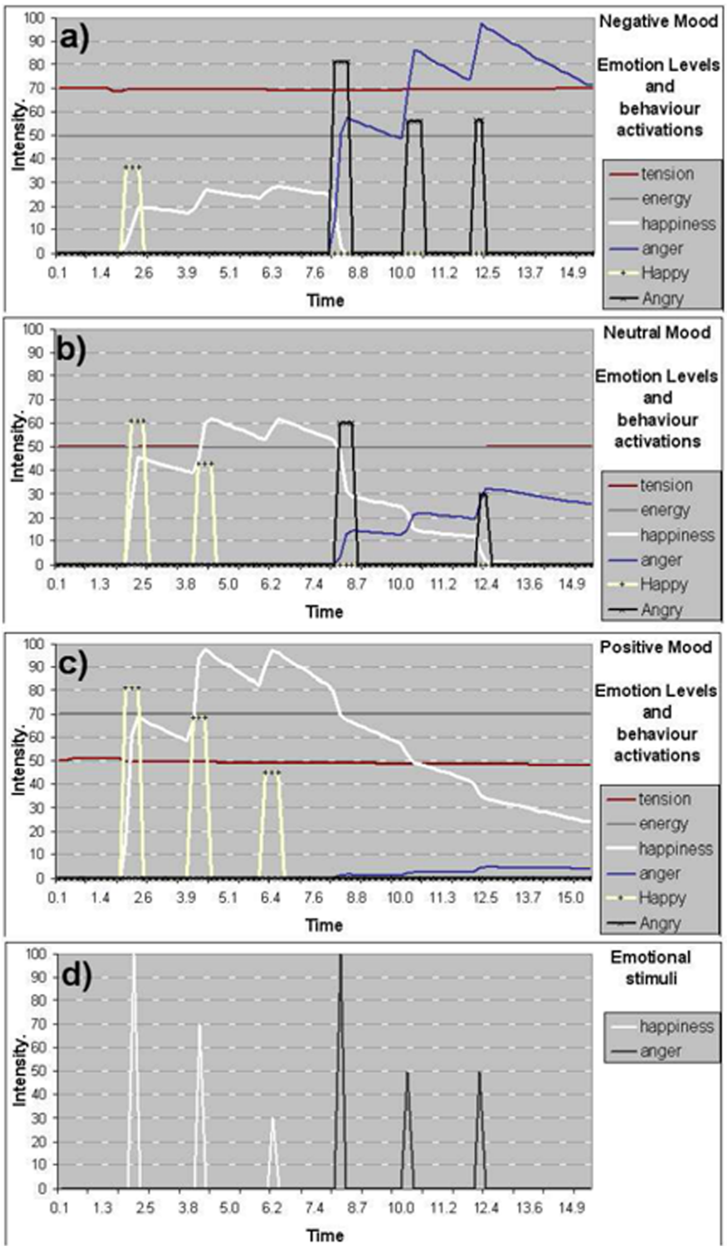
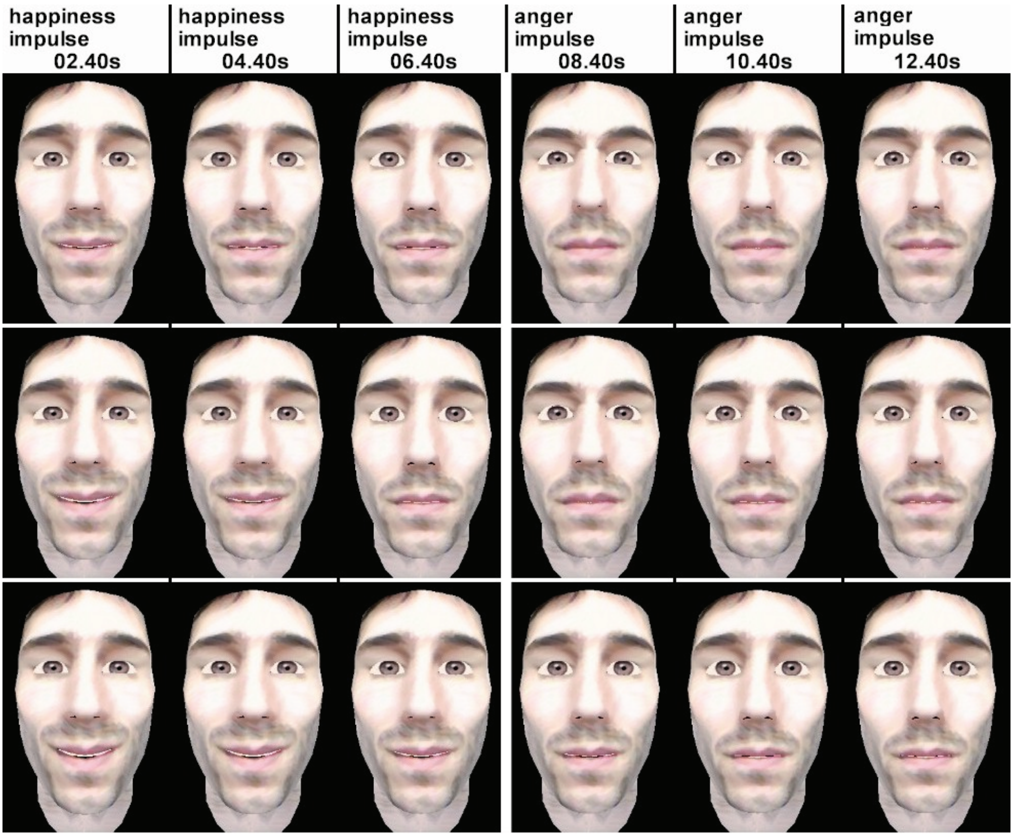


Figure 2. The top through bottom rows show EE-FAS screenshots of the Negative, Neutral and Positive Mood, corresponding to charts a , b and c of Figure 1. Columns correspond to the time when a impulses are sent to the DER, as per d of that figure.



slowly to return to its neutral position or to be replaced by a new expression. Typically a happy expression is shown by raising the lip corners and low-eyelids; and an angry expression is represented by eyebrow frowns and tight lips.

In this example, the elicited emotional impulses could be due to the appraisal of six declarations from the employer of the character: "You are our best employee (first happiness stimulus), you have helped us tremendously (second stimulus) and I should increase your salary (third). However, I can't do this (first anger stimulus)—the company is not doing well (second anger), so we will all need to work harder (third)." This sequence of events produces various reactions (such as a smile) depending on the agent's mood, but they also change the emotional state of the character. This in turn influences the character's next action.

In row 1, the first screen shot shows some response to the happiness impulse. However this activation is reduced due to the influence of negative mood. The next two impulses show no response because the happiness impulse intensities are too low to achieve behaviour activation in this mood. In the neutral mood context, row 2, the first happiness impulse produces a stronger activation than in the context of negative mood. Here the second impulse also produces an activation but the intensity of the third impulse is still too low to generate a response. The positive mood, row

3, amplifies the happiness impulses, therefore they all produce facial expressions, and these expressions have stronger intensities than in the previous contexts.

The effects of anger impulses are influenced by mood states but in addition, they are also influenced by the level of the emotion *happy* generated by the previous happiness impulses. In row 1, all three anger impulses produce behaviour activations, showing an expression of anger represented by an eyebrow frown and tight lips. The different impulse intensities, 100%, 50% and 50%, produce different strengths of expression, but the differences are not marked since the effect of negative mood and the building anger amplifies the lower intensity impulses.

In row 2, the neutral mood, the response of the first anger impulse is stronger than the responses of the two other anger impulses. In fact, Figure 1 shows that no behaviour activation has been triggered by the second anger activation, therefore no new facial expression is displayed. The expression shown at the time of the second anger impulse is due to the visual momentum of the expression produced by the first anger activation which is still fading. The third anger impulse does produce an emotional expression, since happiness has decreased. The difference between these is the emotional momentum produced by the previous happiness impulses. Happiness impulses increase the level of happiness, which takes time to disappear and inhibits the effects of anger impulses. In contrast, the negative agent (row 1) never became very happy in the first place. While for the positive agent, (row 3) no anger impulses produce any behavioural response. This is due to two reasons. First, the positive mood reduces the effects of anger impulses. Second, the happiness momentum produced by the previous happiness impulses also reduces the effects of anger impulses. The largest impact of the bad news is the reduction of happiness which had previously soared.

The DER Basic Representation

The basic unit of the DER model is a modular representation based on the (Picard, 1997) description of emotion intensity and emotion filters. A DER network consists of a system of modules connected by filters.

We assume an emotion appraisal mechanism, such as those based on the OCC model (Ortony et al., 1990), classifies events, actions and objects, outputting emotion types and intensities. We call this output an *emotional impulse*. Emotional impulses are defined by:

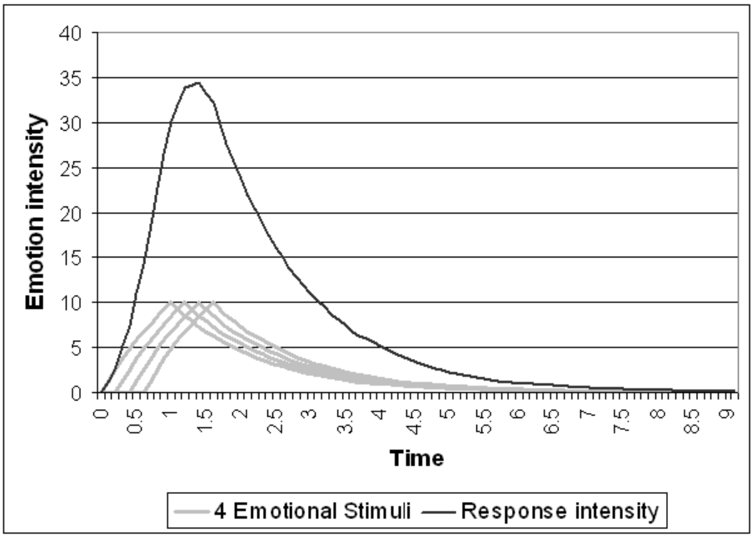
- the name of an emotion,
- an intensity value and
- a valence, which can be 1, 0 or -1.

The valence specifies whether a behaviour activation is positive, neutral or negative. Fast and slow mechanisms eliciting emotions can produce the same type of behaviour activations with different intensity or produce different types of behaviour activations affecting the same or different emotions.

Emotional impulses are transformed by the DER into emotion stimuli. The stimuli include timing information typical of an emotion type and represent intensity changes over time. The activation curve representing emotional stimuli, shown at the bottom of Figure 3, was chosen to represent the slow decay of emotion intensity (Picard, 1997). The effects of small emotional events are cumulative. Therefore emotional stimuli are summed to compute the intensity of an emotion, as shown by the top curve in Figure 3.

As shown in the example above (Figure 1), the effect of an emotional event depends on an agent's current emotional state. This is implemented in the DER by connecting modules through a system of *dynamic filters*. Sigmoid functions have been chosen for this role because they describe "a large variety of natural phenomena" (Picard, 1997). The sigmoid function parameters can be modified depending on the DER model's state, resulting, for example, in

Figure 3. Emotional stimuli are summed to compute the intensity of an emotion

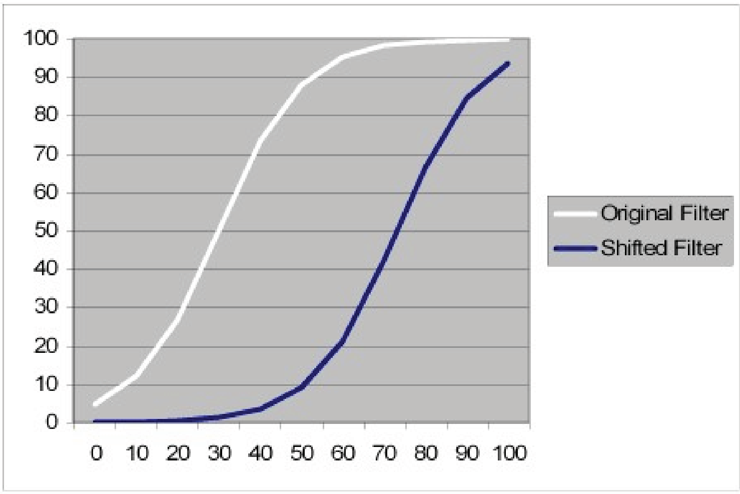


a shifted sigmoid curve as show in Figure 4. This effect simulates the change of sensitivity, for example, emotion threshold activation, to particular emotional stimulus in relation to the current emotional state of a person. In the DER model, we refer to each module as a *dimension*. A dimension may represent for instance a particular emotion or a particular behaviour activation. Dimensions consist of a list of dynamic filters. One dynamic filter modifies one parameter of a particular type of emotional stimulus, for example its peak intensity. Some characteristics, such as the decay duration, can be modified. In Figure 4, the horizontal axis represents the input value, such as the peak intensity of an emotional stimulus, and the vertical axis represents the output value, such as a new peak intensity value.

Modules influence each other by passing their output to a bus which modifies the parameters of other module's dynamic filters. This bus is the same one conducting the emotional stimuli; the influence on other modules occurs through their "input" filter. Figure 5 shows states affected by two types of emotional stimulus, *Anger* and *Happiness*, and whether the influence is positive or negative for each type of emotional stimuli. For instance, happy stimuli affect positively the emotional state of *Happiness* and negatively the state of *Anger*. This figure also shows that the emotional state Happiness amplifies the effects of happy stimuli and reduces the effects of angry stimuli. In practise, the higher the level of the emotional state *Happiness*, the more the sigmoid functions controlling the effects of emotional stimuli anger are shifted to the right. This mechanism decreases the positive effect of anger stimuli on the emotional state anger. This simulates the effects of good and bad news on different moods as demonstrated above.

The system can be configured such that any type of behaviour activation can affect any emotional state and any emotional state can influence the effects of emotional impulses on any emotional state. For the instance of the DER model integrated in the EE-FAS (described below), the tuning has been carried out using heuristic values and visualisation software plotting the resulting sigmoid functions as well as the modified emotional stimuli.

Figure 4. Sigmoid functions are used as dynamic filters by changing their parameters



An Embedded Instance of a DER

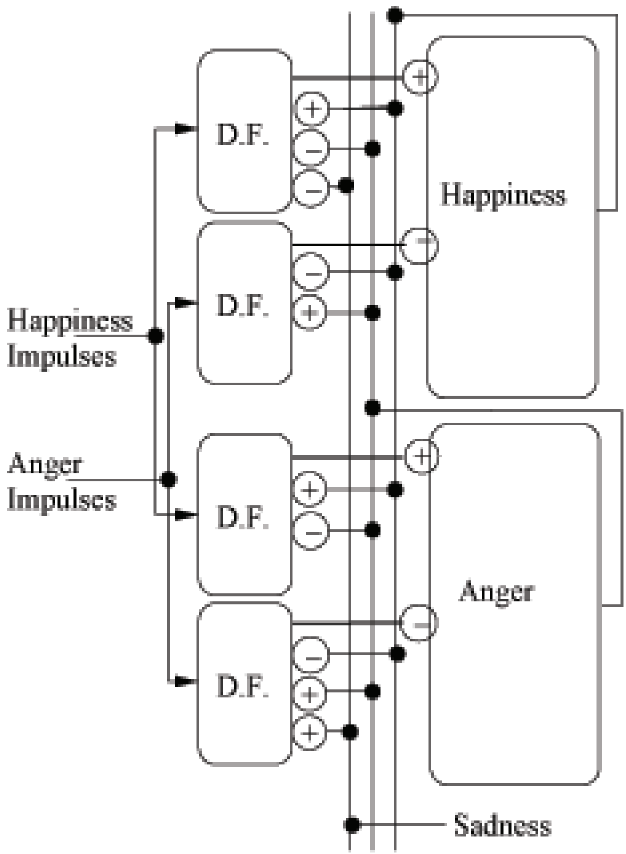
Using the representations just described, we created an instance of a DER and integrated it into the Emotionally Expressive Facial Animation System (EE-FAS), which is a custom-built virtual reality system for producing 3D facial avatars. As mentioned earlier, this instance of the DER is composed of three types of modules: *behaviour activations*, *emotions* and *moods*. Each represents persisting states changing on different timescales. Figure 6 shows a graphical representation of this DER.

The behaviour activations are generated by the DER due to emotional impulses. Impulses can come from any intelligent system capable of appraisal. We have so-far used impulses from three sources:

- As part of a script the EE-FAS avatar essentially acts. The scripts include XML mark up indicating where impulses should be inserted. The exact nature of the agent's acting is dependent on the state of the DER (the agent's mood) and can thus be extremely varied. Example videos of this can be found on our Web page.
- By the real-time appraisal of textual interactions with users. The appraisal of text for emotional content was provided for us by the Bristol company Elzware as a Web service.
- Through button presses on a GUI interface. This interface allows for the real-time influencing of the DER which can be used in concert with the script-reading for demos.

Any behaviour activation is displayed by the EE-FAS as one of the Ekman's emotional facial expressions. However, the duration of an emotional facial expression is different from the duration of the corresponding behaviour activation, as expressions have their own innate time courses representing underlying durative (probably chemical or hormonal) state. The graphs a, b and c in Figure 1 show the behaviour activations *Happy* and *Angry*. Emotions, such as *Happiness* and *Anger*, also produced by emotional impulses, last longer than behaviour activations. In the EE-FAS, emotions are also used to select facial signals corresponding to communicative functions when these are requested, for example, by a

Figure 5. Example of a network of influences between emotional states and emotional impulses in the DER model. D.F. stands for Dynamic Filters

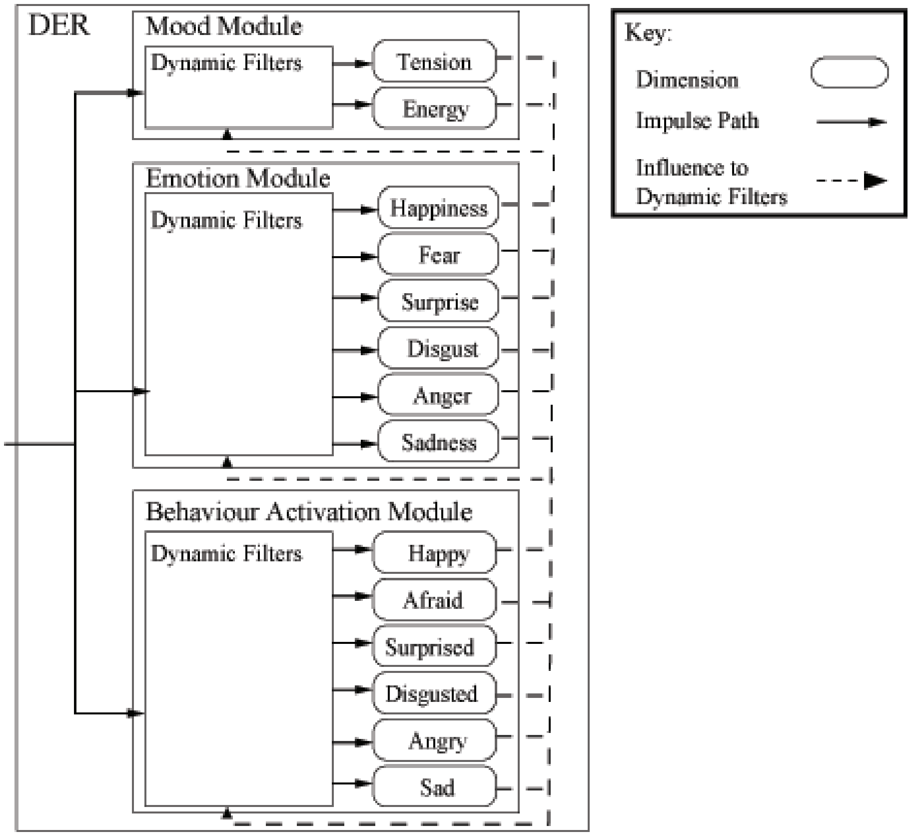


script. For example, a semi-deliberate facial expressions can be synchronised with speech, such as emphasis or deliberate smiles. Mood changes on a slower timescale than emotions and it influences and is influenced by the effects of emotional impulses on the DER state. More detail on the design of this DER can be found elsewhere (Tanguy, 2006).

Experiments and Applications of the DER

We are evaluating both the believability and the usability of the DER through a series of applications. So far, most of these have incorporated the DER instance described above and the EE-FAS. Ellis and Bryson (2005) describe an experiment using the DER which demonstrates the importance of graphical textures to human recognition of the emotions communicated by intelligent virtual agents. There were extremely significant differences in user assessment of the emotions of the EE-FAS between three different textures: a simple, cartoon-like texture, a photo-realistic male texture, and a photo-realistic female texture. In all three cases, the textures overlay the same standard (male) facial structure. The textures themselves determined

Figure 6. A DER composed of three types of state changing on different timescales: behaviour activations, emotions and moods



systematically what emotions viewers were likely to detect.

We have also used the system for live demonstrations both of the DER itself and of real-time chatbot technology where the natural language capabilities were provided by Elzware. We have worked on integrating the EE-FAS into an assistive home environment, making it give suggestions on typing breaks based on camera input as a prototype system. Below we describe briefly some experiments using the DER and EE-FAS to further assess the human perception of emotions.

There is an open question in psychology as to whether humans recognise emotions from facial components or from full-face configurations (Izard, 1997; Smith & Scott, 1997; Wehrle, Kaiser, Schmidt, & Scherer, 2000). Using the DER in the EE-FAS, we generated different animations from an animation script to test whether humans can recognise emotions from component-only changes—a test impossible with real human faces. By changing the state of the DER different facial signals corresponding to communicative functions are selected. We designed an experiment using videos displaying different facial component actions, such as *raise lip corners* or *eyebrow frown*. The hypothesis of this experiment was that individual facial component actions would influence subjects' perception of a virtual actor and that they communicate their own meanings.

The experiment involved 60 subjects. Each subject watched all 17 videos in randomised order and after each video the subject filled a short questionnaire about the virtual actor present in the video. We showed that emotions could be recognised with minimal cues: eyebrow frown (AU4) is interpreted as a sign of anger and displeased state; lip corner raise (AU12) is interpreted as a sign of happiness and pleased state; eyebrow oblique (AU1) is interpreted as a sign of sadness and displeased state. We thus find that facial component actions are sufficient for interpreting emotions without the presence of full-face configuration changes. The experiment also shows that eyebrow oblique is interpreted as a sign of sincerity where the combination of lip corner raise and eyebrow frown is interpreted as a sign of insincerity. The combination of these two facial component actions was in our system as an emergent property due to the presence of two emotions. All these results are highly significant with a p value under 0.001. Further details of the experimental procedure and results are reported by (Tanguy, 2006).

Comparison to Related Systems

A fundamental concept of the DER model is the division of emotion models into mechanisms eliciting emotions and emotion representations. This greatly simplifies scripting or direction for characters, since once the DER is set and an initial mood described, the script or direction can describe communicative actions abstractly rather than describing precise facial expressions. For agents situated in long-term real-time domains such as Web pages or interactive environments, this also provides a mechanism for greater variety of behaviour per scripted response, since the mood can depend on the recent interaction history.

Emotion models are heavily researched now; of these a number of researchers have made systems to which the DER is more or less similar. The main advantage of the DER is its modularity, flexibility, ease of configuration and relative autonomy (lack of required direction) once completed. Its main disadvantage is that it takes some time "out of the box" to configure, although we also distribute through our Web page a pre-configured version with the EE-FAS instance described earlier. The DER can:

- represent any number of emotions,
- represent emotions or other forms of durative state with different time scale,
- define interactions between emotions, and
- customise the influences of emotional impulses on each emotion.

Paiva et al. (2004) present an emotion model which assigns a decay function to each emotion elicited with a value higher than a personality threshold. In contrast to the DER, Paiva et al.'s work does not implement any interaction between emotions in its emotion representation. Egges, Kshirsagar, and Magnenat-Thalmann (2004) describe a generic emotion and personality representation composed of two types of affective states, moods and emotions. Any number of moods and emotions can be represented. In the implementation of their model the only influences between states is the influence of mood on emotions. Egges et al. compute the intensity of affective states by linear functions through the use of matrix operations. In the DER, sigmoid functions are used to control and change the influences of emotional stimuli on emotions and the influences between emotions. This mechanism produce non-linear behaviours closer to natural phenomena.

Bui (2004) does use a decay function to represent the durations of emotions, but the effects of new emotional impulses on an emotion are influenced by the intensity of the other emotions. Their decay functions are also influenced by personality parameters. Velásquez and Maes (1997) present another representation. The computation of the intensity changes of an emotion takes into consideration the intensity of other emotions, the decay in intensity and the previous intensity of the emotion itself. The influences of emotions on others are of the types inhibitory or excitatory. The DER model is rather like these two systems. However it differs in being highly customisable. Any durative state, such as mood, and any number of emotions can be represented, and influences of emotional stimuli on emotions can also be defined by the researcher. The main advantage of the DER model is that its representation can be adapted to different emotion theories and to different mechanisms eliciting emotions from the agent's environment. It is a tool that can help the community to model different emotion theories.

Compared to some systems though we have simplified the system slightly to make it more generic. Our definition of emotional impulses carries less information than the emotional structures described by Reilly (1996), which also contain the cause or referent of the emotions. Reilly and colleagues' interest was primarily in creating *believable* agents, like classic animated cartoon characters which exist to entertain and communicate. We have focussed instead on *realistic* models, which are more useful for experiments and long-term plotless semi-autonomous applications. The DER model focuses on the duration and interaction of emotions, cognitive referents can be tracked in other parts of an agent's intelligence, or the definition of impulse can be expanded. In our system, every emotions decays over time. An emotion such as hope might be thought to persist as long as the situation is the same. We take the position that hope will decay even if the situation stays the same. However, new appraisal of the same situation produces new emotional stimuli increasing the level of hope. Similarly, the DER could be used to represent drives like hunger. The only requirement is inverting the levels and decay function. A drive slowly increases over time, but then consummatory actions (rather than emotional impulses) such as eating can abruptly reduce its level.

Durative State for Goal Arbitration: Flexible Latching

Drives, like emotions, represent essential goals an animal needs to pursue. The hormone and endocrine systems underlying drives and emotions are an evolved system for providing smooth regulation of behaviour (Carlson, 2000). The problem of behaviour regulation includes allocating appropriate amounts of time to a variety of sometimes conflicting goals (Dunbar, 1993; Korstjens, Verhoeckx, & Dunbar, 2006). Through our experience building the DER for the EE-FAS, we realized that realistic, evolved emotions can have extremely complex interactions. If we are trying to build more expediently reliable real-time systems that self-regulate, we need a simpler and clearer system for describing the interaction of drives and emotions.

Drives as Simple Latches

The simplest way to represent drives is as a prioritised set of goals. The highest priority goal that is currently active or *released* directs the actions of the agent. For example, the goal of eating is only active when the agent is hungry, the goal of evading predators is only active when the agent is being chased. If the agent is both hungry *and* being chased, the fleeing should take priority.

The problem with this simple approach is in satiating multiple goals that gradually increase and decrease, for example the desire for food, water and sleep. When is the agent hungry enough to act on eating? If a single arbitrary threshold is chosen, then the agent will oscillate back and forth fairly rapidly between eating and not eating as eating takes its hunger just under the threshold, but then a little bit of time raises it over the threshold again.

A solution to this dithering is common in basic control theory, it is called the *latch*. Essentially, a goal is activated when a relatively high threshold value (e.g., of hunger) is achieved, but not *de*activated until a second, significantly lower threshold value occurs. Building one of these requires a simple module containing the following:

- **Level:** This is the only thing that the DER refers to as *dimensions* present in our simplified representation. For drives, the level increases gradually with time, while consummatory actions can reduce it. Emotions are the inverse --- the passage of time gradually reduces the level, while perceptual events can increase it dramatically.

- **Latch:** A latch consists of two thresholds and a single bit (true-false) of information. The first threshold determines when the level of a drive has reached a level such that, if the agent had not been currently influenced by the drive, it now will be; while the second threshold represents the point at which the drive will stop influencing behaviour if it previously had been. The bit simply records whether the drive's behaviour is active when the level is between the two thresholds. In biology (as in the DER) the thresholds are influenced by external circumstances such as the cost of changing behaviours or the attractiveness of an associated stimulus (e.g., rare food or an attractive mate), but we have not yet attempted to build into the simple drive system an appraisal system for such assessment. Rather, we fix the thresholds for the agents at design time.

Flexible Latching

Although the above simple latching is adequate to begin exploring complex interacting goals such as we see in primate social behaviour (Bryson, 2003), we discovered a problem as we began making more detailed models of primate social interactions.

As mentioned earlier, once a drive has passed the threshold required for it to begin influencing behaviour, it continues to do so until either a higher-priority activity interrupts, or another threshold indicating satiation has been reached. However, sometimes a dynamic external environment intervenes in a way that causes an *external* interruption. For example, if you were eating boiled eggs until you were very nearly satiate but then ran out of eggs, would you boil some more? In our primate simulations, a monkey that had not quite satisfied itself with grooming might pursue another grooming partner for five minutes, only to then satiate after just five seconds of further grooming.

It seems that a more plausible model would be the following. When an agent is reliably engaged in an activity consummatory to its goal (or, indeed, is failing to engage in a consummatory action for some length of time) it may essentially reset its lock status and "reconsider" its current top-level drive. Such a system still escapes dithering, but also provides a more efficient pursuit of goals—a heuristic mechanism for sensing when expectations have been violated and it is sufficiently likely that a goal may no longer be worth pursuing to justify reassessment. An illustration of the efficiency brought on by flexible latching can be seen in Figure 7. The cost for this efficiency is low: really, only adding mechanisms for recognising external interruption. Reassessing priorities after interruption is the same as initial assessment which must already be implemented.

In the system evaluated in Rohlfshagen and Bryson (2008) (see Figure 7) the agents had four goals. Two high-priority goals have to do with the immediate health of the agents: eating and drinking; while two have to do with the adaptive well-being of the agents: grooming (which correlates to safety from predation, see Lehmann & Bryson, 2007) and exploring. For the grooming goal, for example, consummation is the act of grooming, while preparatory actions include selecting and approaching a partner. Eating, drinking and grooming all rely on the stability of a dynamic environment—food sources can be depleted and grooming partners can become bored and leave. Also, grooming can be interrupted by an increased urgency of the higher-priority feeding goals. The evaluation of efficiency is done in terms of how much time the agent has to devote to its lower-priority goals in a fixed-duration "lifetime." This would recognise a small class of exogenous events that might interrupt the behaviour's expression. Full details of this system, how thresholds can be set, and experimental results can be found in Rohlfshagen and Bryson (2008).

An alternative implementation of Flexible Latching would be to include a more general-purpose mechanism for recognising interruptions, rather than requiring specific modification of the consummatory actions. The system might recognise if a preparatory, non-consummatory action involved in achieving a goal has been invoked after a period of consummatory behaviour. This period could be indicated by the drive threshold being below the triggering threshold for the latch, or by a timestamp indicating a recent expression of the consummatory action. For example, one might expect a short interval between eating multiple eggs during a meal which would be required to break into the next one. However, a longer break might indicate an unexpected difficulty in finding that next egg, which indicates an interruption. The advantage of such a strategy would be that it would be robust to detecting unanticipated forms of failure. The disadvantage is that it would complicate and slightly slow the basic action-selection mechanism, since it adds an additional check to executing every action associated with a complex goal, except for the consummatory ones.

CONCLUSION AND IMPLICATIONS

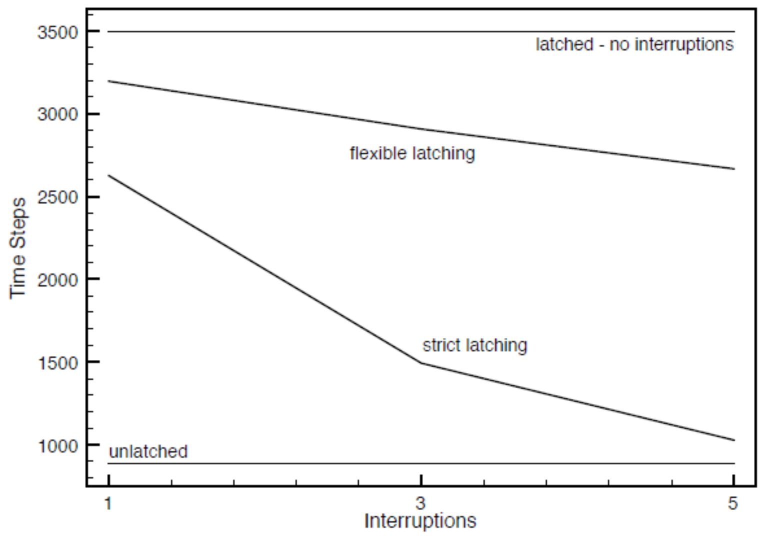
This article has presented representational systems for the relatively persistent action selection state often associated with moods, emotions and drives. First, the Dynamic Emotion Representation enables a programmer to represent any number of persisting states and the interactions between them, whether excitatory or inhibitory, linear or not. The system integrates these with ordinary temporal decay. We also described working systems using these representations. Second, we described a simplification of the DER called Flexible Latching that allows an agent to arbitrate efficiently between multiple conflicting goals. Both of these systems require design to set appropriate levels for behaviour, but in fact most commercial AI systems rely more on design than on automated planning or learning. What AI requires is an iterative development approach that optimises between the intelligence the artifacts can find for themselves and what is best provided by the designer (Bryson & Stein, 2001). What we emphasise here is getting the representations and ideas right such that designers have a relatively easy job for building emotions and control. We hope our work helps roboticists and other real-time AI designers when they specify and build their agents.

Interestingly, at least two published models of consciousness are similar to our representations of durative state. Norman and Shallice (1986) describe consciousness as a higher cost attentional system which is brought on line whenever the more basic, reliable, low-cost action-sequencing mechanism is unable to proceed. This is similar to our proposals in Flexible Latching. Shanahan (2005) proposes a model of mutually-inhibiting actions in a global workspace. We do not agree that Shanahan's model can account for all of action selection, for example, see the Tyrrell (1994) critique of Maes (1991). However, his model is similar to what we propose in the DER and with Flexible Latching for arbitration between certain types of high-level tasks. It may be that spreading activation models, while not effective for modelling detailed action selection, are good models for arbitration between high-level goals. This may in turn imply that every emotion or drive an animal has a corresponding adaptive goal for the agent, and vise versa. This would indicate that in agent design, every goal set for the agent should be associated with an element of durative state. Further, every autonomous agent should have as an essential core of its action selection a mechanism for arbitrating between these goals. For life-like agents, this arbitration should be done through a system of mutual excitation and inhibition, as demonstrated in the DER.

ACKNOWLEDGMENT

This work was supported by a PhD studentship for Tanguy from the Department of

Figure 7. Performance of three action-selection systems measured in terms of the number of time steps allocated to low-level priorities. This measure indicates the overall efficiency with which higher level goals are met while the agents' basic needs are satisfied sufficiently for it to stay "alive." Latching is always more efficient than dithering, but in the case of external interruptions strict latching is significantly less so. Flexibility in reassigning latched drives after an interruption ameliorates this problem.



Computer Science, University of Bath, and by the UK EPSRC AIBACS initiative, grant GR/S/79299/01. This article extends a previous conference article by TanguyIJCAI07. We would like to thank Philipp Rohlfshagen and Philip J. Willis for their contributions to the research described here. All code described here is available open-source on line and can be found from the AmonI Software page. Bryson's time on this article was supported in part by a sabbatical fellowship provided by the Konrad Lorenz Institute for Evolution and Cognition Research, in Altenberg, Austria.

REFERENCES

André, E., Klesen, M., Gebhard, P., Allen, S., & Rist, T. (1999, October). Integrating models of personality and emotions into lifelike character. In *Proceedings of the Workshop on Affect in Interaction — Towards a new Generation of Interfaces,* Siena, Italy (pp. 136-149). New York: Springer.

Badler, N., Allbeck, J., Zhao, L., & Byun, M. (2002, June). Representing and parameterizing agent behaviors. In *Proceedings of Computer Animation*, Geneva, Switzerland (pp. 133-143). Washington, DC: IEEE Computer Society.

Bizzi, E., Giszter, S. F., Loeb, E., Mussa-Ivaldi, F. A., & Saltiel, P. (1995). Modular organization of motor behavior in the frog's spinal cord. *Trends in Neurosciences*, *18*, 442–446. doi:10.1016/0166-2236(95)94494-P

Blumberg, B. M. (1996). *Old tricks, new dogs: Ethology and interactive creatures*. Unpublished PhD thesis, MIT Media Laboratory, Learning and Common Sense Section.

Brahnam, S., & De Angeli, A. (2008). Special issue on the abuse and misuse of social agents. *Interacting with Computers*, *20*(3), 287–291. doi:10.1016/j.intcom.2008.02.001

Brand, M. (2001, December). Morphable 3D models from video. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR'01)*, Kauai, HI (Vol. 2, pp. II456-II463). Washington, DC: IEEE Computer Society.

Breazeal, C. (2003). Emotion and sociable humanoid robots. *International Journal of Human-Computer Studies*, *59*(1–2), 119–155. doi:10.1016/S1071-5819(03)00018-1

Breazeal, C., & Scassellati, B. (1999, October). How to build robots that make friends and influence people. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS-99)*, Kyongju, Korea (pp. 858-863). Washington, DC: IEEE Computer Society.

Broekens, J., Kosters, W. A., & Verbeek, F. J. (2007). Affect, anticipation, and adaptation: Affect-controlled selection of anticipatory simulation in artificial adaptive agents. *Adaptive Behavior*, *15*(4), 397–422. doi:10.1177/1059712307084686

Bryson, J. J. (2000). Cross-paradigm analysis of autonomous agent architecture. *Journal of Experimental & Theoretical Artificial Intelligence*, *12*(2), 165–190. doi:10.1080/095281300409829

Bryson, J. J. (2003). Where should complexity go? Cooperation in complex agents with minimal communication. In W. Truszkowski, C. Rouff, & M. Hinchey (Eds.), *Innovative concepts for agent-based systems* (pp. 298-313). New York: Springer.

Bryson, J. J. (2008, March). The impact of durative state on action selection. In I. Horswill, E. Hudlicka, C. Lisetti, & J. Velasquez (Eds.), *Proceedings of the AAAI Spring Symposium on Emotion, Personality, and Social Behavior,* Palo Alto, CA (pp. 2-9). AAAI Press.

Bryson, J. J., & Stein, L. A. (2001, August). Modularity and design in reactive intelligence. In *Proceedings of the 17th International Joint Conference on Artificial Intelligence*, Seattle, WA (pp. 1115-1120). Morgan Kaufmann.

Bryson, J. J., & Thórisson, K. R. (2000). Dragons, bats & evil knights: A three-layer design approach to character based creative play. *Virtual Reality (Waltham Cross)*, *5*(2), 57–71. doi:10.1007/BF01424337

Bui, T. D. (2004). *Creating emotions and facial expressions for embodied agents*. Unpublished PhD thesis, University of Twente.

Cañamero, D. (2003). Designing emotions for activity selection in autonomous agents. In R. Trappl, P. Petta, & S. Payr (Eds.), *Emotions in humans and artifacts* (pp. 115-148). Cambridge, MA: MIT Press.

Carlson, N. R. (2000). *Physiology of behavior* (7th ed.). Boston: Allyn and Bacon.

De Rosis, F., Pelachaud, C., Poggi, I., Carofiglio, V., & De Carolis, B. (2003). From greta's mind to her face: Modelling the dynamics of affective states in a conversational agent. *International Journal of Human-Computer Studies*, *59*, 8–118.

Delgado-Mata, C., & Aylett, R. S. (2004, July). Emotion and action selection: Regulating the collective behaviour of agents in virtual environments. In *AAMAS ' 04: Proceedings of the 3rd International Joint Conference on Autonomous Agents and Multiagent Systems,* New York (Vol. 3, pp. 1304-1305). Washington, DC: IEEE Computer Society.

Dunbar, R. I. M. (1993). Coevolution of neocortical size, group size and language in humans. *The Behavioral and Brain Sciences*, *16*(4), 681–735.

Egges, A., Kshirsagar, S., & Magnenat-Thalmann, N. (2004). Generic personality and emotion simulation for conversational agents. *Computer Animation and Virtual Worlds*, *15*, 1–13. doi:10.1002/cav.3

Ellis, P. M., & Bryson, J. J. (2005, September). The significance of textures for affective interfaces. In T. Panayiotopoulos, J. Gratch, R. Aylett, D. Ballin, P. Olivier, & T. Rist (Eds.), *Intelligent Virtual Agents: Proceedings of the Fifth International Working Conference on Intelligent Virtual Agents*, Kos, Greece (LNCS 3661, pp. 394-404).

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

Gadanho, S. C. (1999). *Reinforcement learning in autonomous robots: An empirical investigation of the role of emotions*. Unpublished PhD thesis, University of Edinburgh.

Gratch, J., & Marsella, S. (2004, August). Evaluating the modeling and use of emotion in virtual humans. In *Proceedings of the 3rd International Joint Conference on Autonomous Agents and Multiagent Systems,* New York (Vol. 1, pp. 320-327). Washington, DC: IEEE Computer Society.

Graziano, M. S. A., Taylor, C. S. R., Moore, T., & Cooke, D. F. (2002). The cortical control of movement revisited. *Neuron*, *36*, 349–362. doi:10.1016/S0896-6273(02)01003-6

Hiller, M. J. (1995). *The role of chemical mechanisms in neural computation and learning*. (Tech. Rep. AITR-1455). Cambridge, MA: MIT AI Laboratory.

Izard, C. E. (1993). Four systems for emotion activation: Cognitive and noncognitive processes. *Psychological Review*, *100*(1), 68–90. doi:10.1037/0033-295X.100.1.68

Izard, C. E. (1997). Emotions and facial expression: A perspective from differential emotions theory. In J. A. Russel & J. M. Fernández-Dols (Eds.), *The psychology of facial expression* (pp. 57-77). Cambridge, UK: Cambridge University Press.

Korstjens, A. H., Verhoeckx, I. L., & Dunbar, R. I. M. (2006). Time as a constraint on group size in spider monkeys. *Behavioral Ecology and Sociobiology*, *60*(5), 683–694. doi:10.1007/s00265-006-0212-2

Lazarus, R. S. (1991). *Emotion and adaptation*. New York: Oxford University Press.

LeDoux, J. (1996). *The Emotional Brain: The mysterious underpinnings of emotional life*. New York: Simon and Schuster.

Lehmann, H., & Bryson, J. J. (2007, September). Modelling primate social order: Ultimate causation of social evolution. In F. Amblard (Ed.), *Proceedings of the 4th Conference of the European Social Simuation Society (ESSA '07),* Toulouse, France (p. 765). IRIT Publications.

Maes, P. (1991). The agent network architecture (ANA). *SIGART Bulletin*, *2*(4), 115–120. doi:10.1145/122344.122367

Marcella, S., & Gratch, J. (2002, July). A step toward irrationality: Using emotion to change belief. In *Proceedings of the 1st International Joint Conference on Autonomous Agents and Multiagent Systems*, Bologna, Italy (pp. 334–341). ACM Publishing.

Minsky, M., Singh, P., & Sloman, A. (2004). The St. Thomas common sense symposium: Designing architectures for human-level intelligence. *AI Magazine*, *25*(2), 113–124.

Morgado, L., & Gaspar, G. (2005, July). Emotion based adaptive reasoning for resource bounded agents. In *Proceedings of the 4th International Joint Conference on Autonomous Agents and Multi Agent Systems (AAMAS '05),* Utrecht, The Netherlands (pp. 921-928). ACM Publishing.

Norman, D. A., Ortony, A., & Russell, D. M. (2003). Affect and machine design: Lessons for the development of autonomous machines. *IBM Systems Journal*, *42*, 38–44.

Norman, D. A., & Shallice, T. (1986). Attention to action: Willed and automatic control of behavior. In R. Davidson, G. Schwartz, & D. Shapiro (Eds.), *Consciousness and self regulation: Advances in research and theory* (Vol. 4, pp. 1-18). New York: Plenum.

Ortony, A., Clore, G. L., & Collins, A. (1990). *The cognitive structure of emotions*. Cambridge, UK: Cambridge University Press.

Paiva, A., Dias, J., Sobral, D., Aylett, R., Sobreperez, P., Woods, S., et al. (2004, July). Caring for agents and agents that care: Building empathic relations with synthetic agents. In *AAMAS '04: Proceedings of the 3rd International Joint Conference on Autonomous Agents and Multiagent Systems,* New York (pp. 194-201). Washington, DC. IEEE Computer Society.

Picard, R. W. (1997). *Affective computing*. Cambridge, MA: MIT Press.

Plutchik, R. (1980). A general psychoevolutionary theory of emotion. In R. Plutchik & H. Kellerman (Eds.), *Emotion: Theory, research, and experience* (pp. 3-33). New York: Academic Press.

Prescott, T. J., Bryson, J. J., & Seth, A. K. (2007). Modelling natural action selection: An introduction to the theme issue. *Philosophical Transactions of the Royal Society, B— Biology, 362*(1485), 1521-1529.

Reilly, W. S. N. (1996). *Believable social and emotional agents*. Unpublished PhD thesis, School of Computer Science, Carnegie Mellon University, Pittsburgh.

Rohlfshagen, P., & Bryson, J. J. (2008, November). Improved animal-like maintenance of homeostatic goals via flexible latching. In A. V. Samsonovich (Ed.), *Proceedings of the AAAI Fall Symposium on Biologically Inspired Cognitive Architectures,* Arlington, VA (pp. 153-160). AAAI Press.

Russel, J. A., & Fernández-Dols, J. M. (1997). *The psychology of facial expression*. Cambridge, UK: Cambridge University Press.

Schaal, S., Ijspeert, A., & Billard, A. (2004). Computational approaches to motor learning by imitation. In C. D. Frith (Ed.), *The Neuroscience of Social Interaction: Decoding, Imitating, and Influencing the Actions of Others*, (pp. 199-218). New York: Oxford University Press.

Shanahan, M. P. (2005). Global access, embodiment, and the conscious subject. *Journal of Consciousness Studies*, *12*(12), 46–66.

Sloman, A. (2003). How many separately evolved emotional beasties live within us? In R. Trappl, P. Petta, & S. Payr (Eds.), *Emotions in humans and artifacts* (pp. 35-114). Cambridge, MA: MIT Press.

Sloman, A., & Croucher, M. (1981, August). Why robots will have emotions. In *Proceedings of the* 7*th* *International Joint Conference on Artificial Intelligence (IJCAI '81)*, Vancouver, British Columbia, Canada (pp. 1537-1542). William Kaufmann.

Smith, C. A., & Scott, H. S. (1997). A componential approach to the meaning of facial expressions. In J. A. Russel & J. M. Fernández-Dols (Eds.), *The psychology of facial expression* (pp. 295-320). Cambridge, UK: Cambridge University Press.

Tanguy, E. A. R. (2006). *Emotions: The art of communication applied to virtual actors* (Bath CS Tech. Rep. CSBU-2006-06). Bath, UK: University of Bath.

Tanguy, E. A. R., Willis, P. J., & Bryson, J. J. (2003). A layered dynamic emotion representation for the creation of complex facial animation. In T. Rist, R. Aylett, D. Ballin, & J. Rickel (Eds.), *Intelligent virtual agents* (pp. 101-105). New York: Springer.

Tanguy, E. A. R., Willis, P. J., & Bryson, J. J. (2006). A dynamic emotion representation model within a facial animation system. *International Journal of Humanoid Robotics*, *3*(3), 293–300. doi:10.1142/S0219843606000758

Tanguy, E. A. R., Willis, P. J., & Bryson, J. J. (2007, January). Emotions as durative dynamic state for action selection. In *Proceedings of the 20th International Joint Conference on Artificial Intelligence*, Hyderabad, India (pp. 1537-1542). Morgan Kaufmann.

Tomb, I., Hauser, M. D., Deldin, P., & Caramazza, A. (2002). Do somatic markers mediate decisions on the gambling task? *Nature Neuroscience*, *5*(11), 1103–1104. doi:10.1038/nn1102-1103

Tyrrell, T. (1994). An evaluation of Maes's bottom-up mechanism for behavior selection. *Adaptive Behavior*, *2*(4), 307–348. doi:10.1177/105971239400200401

Velásquez, J. D., & Maes, P. (1997, February). Cathexis: A computational model of emotions. In *Proceedings of the 1st International Conference on Autonomous Agents (Agents '97)*, Marina Del Ray, CA (pp. 518-519). ACM Publishing.

Wehrle, T., Kaiser, S., Schmidt, S., & Scherer, K. R. (2000). Studying the dynamics of emotional expression using synthesized facial muscle movements. *Journal of Personality and Social Psychology*, *78*(1), 105–119. doi:10.1037/0022-3514.78.1.105

Whiteson, S., Taylor, M. E., & Stone, P. (2007). Empirical studies in action selection for reinforcement learning. *Adaptive Behavior*, *15*(1), 33–50. doi:10.1177/1059712306076253

Wood, M. A., & Bryson, J. J. (2007). Skill acquisition through program-level imitation in a real-time domain. *IEEE Transactions on Systems, Man, and Cybernetics. Part B, Cybernetics*, *37*(2), 272–285. doi:10.1109/TSMCB.2006.886948

Zadeh, S. H., Shouraki, S. B., & Halavati, R. (2006). Emotional behaviour: A resource management approach. *Adaptive Behavior*, *14*(4), 357–380. doi:10.1177/1059712306072337

Joanna J. Bryson studies the structure of both natural and artificial cognitive systems. She holds degrees in behavioural science, psychology and artificial intelligence from Chicago (BA), Edinburgh (MSc and MPhil), and MIT (PhD). Since 2002 she has been a lecturer (assistant professor)

at the University of Bath where she founded Artificial Models of Natural Intelligence. She has over sixty peer-reviewed publications in artificial intelligence, biology, cognitive science and philosophy, and serves as en expert consultant on cognitive systems for the European commission. From 2007-2009 she served as pursued a sabbatical fellowship at the Konrad Lorenz Institute for Evolution & Cognition Research in Altenberg, Austria, studying the biological evolution of cultural evolution. Her main engineering interest is in broadening access to artificial intelligence by making it easier to design and develop.

Emmanuel Tanguy completed his PhD in June 2006 under the supervision of professor Phil Willis as his first supervisor and Joanna J. Bryson as his second supervisor, working in the field of facial animation and synthetic emotions with the developement of a dynamic emotion representation model. The title of his PhD thesis is Emotions: the Art of Communication Applied to Virtual Actors*. Currently, he is developing research on the influences of emotions on synthetic facial expressions and their effects on people's perception of embodied virtual agents. He developed an Emotionally Expressive Facial Animation System (EE-FAS: e-face) for embodied virtual agents, integrating a dynamic emotion representation inspired by natural models and psychological studies of human emotions.*