Personality Simulation in Interactive Agents Through Emotional Biases

José Serra  
Instituto de Telecomunicações  
– Porto  
Rua Dr. Roberto Frias  
4200-465 Porto, Portugal  
jserra@dcc.fc.up.pt

Pedro Nogueira  
Artificial Intelligence and Computer  
Science Laboratory – Porto  
Rua Dr. Roberto Frias  
4200-465 Porto, Portugal  
pedro.alves.nogueira@fe.up.pt

KEYWORDS  
Emotions, personality simulation, BDI, cognitive simulation, framework.

ABSTRACT

Personality simulation has been for the past decades a recurrent topic, as it is a non-trivial problem where a multitude of equally valid approaches can (and have) been taken. However, no general consensus has emerged from this discussion, which further difficulties novel contributions. Being a multidisciplinary research topic between computer science and psychology further complicates the aforementioned issue, as a large number of models for emotion definition currently exist. Thus, methods capable of simulating the variations that occur at an emotional level due to personality and at the same time provide an easy way to define behavior are still lacking. In this paper, we propose the usage of emotions as the prime factor in the personality definition process, by using a height-map represented gravity field to define the emotional biases that each individual’s personality encompasses. The use of this model eases the definition of how the emotional variations occur. We propose a novel architecture, which is capable of receiving and produce any kind of stimuli and reaction. The method was implemented as a proof-of-concept and initial tests show it is effective, behaving as expected.

INTRODUCTION

Emotions play a crucial role on peoples lives, influencing their social interactions reactions to external stimuli both consciously and unconsciously. This phenomenon occurs due to emotions’ nature as complex psychophysiological experiences of an individual, which influenced by a state of mind, arise as the result of interacting biochemical reactions and environmental interactions. As they occur at a deep and sometimes instinctual or subconscious level, emotions influence humans in a very meaningful and critical way, often overriding even rational thought. In fact, emotions influence many aspects of the mental process, such as decision-making, planning, memory and attention (Berkowitz 2000), thus being crucial for human intelligence (Bower and Cohen 1982) and largely guiding how people adapt their behaviour to the world around them. Damasio has shown that people which lack emotional responses to stimuli, make poor decisions that can affect their integration in society (Damasio 1994). As a result, when simulating human-like interactions with an environment it is necessary to take into account the influence of emotions when choosing adequate courses of action.

An agent is an entity that receives inputs, usually from a sensor or set of sensors, analyses them and chooses an action according to some model or ruleset. In the case of humanoid agents, if the chosen action is always determined by a non-stochastic mechanism i.e. reacting in a nil entropic manner, the simulation will not be realistic. This is extremely important in areas such as crowd-simulation or video-games, where the chosen behaviours should be as realistic, coherent and believable as possible. If the behaviour selection mechanism does not take this into account, the simulation will, at some point, inevitably fail. This breakes the illusion of a realistic system and ultimately leads to an awkward and unengaging interaction with the player. Thus, it is crucial to model the emotional component when simulating human behaviours.

In this paper we propose a method for simulating a personality based on emotional biases as a means for simulating a believable emotional component in an agent’s behavior. In our model, each agent’s state is defined by its current mood, i.e. the emotion it has at a specific step in the simulation. This mood is modified by the occurrence/perception of external stimulus. Each stimulus has an associated emotional weight that determines how each one influences the current mood. However, the main feature of our model is that the aforementioned stimuli do not act in a linear fashion, as their effect on each agent depends on the agent’s personality. This personality represents the agent’s emotional biases, i.e. if an agent has a tendency towards positive emotional states and a negative emotional stimulus is given, the stimulus’s impact on the mood will be reduced. Conversely, an agent with a negative bias would interpret the stimuli in the opposite manner. Thus the agents’ mood converges to a specific range of emotional states depending on his personality. The suggested approach is based on a novel methodology to define personalities through the use of height-maps, where the height value of each mood is directly correlated to its influence on the provided stimuli. Based on the similarity to the current mood and its own predefined mood, the agent reacts to the stimuli by choosing an action. The created method is completely focused on the emotion selection and calculus based on...
the influence of the personality and the emotional weight of the perceived stimuli. In sum, the method is capable of automatically - and without the common necessity to define all possible emotion-stimulus combinations - model the emotions and behaviours of an agent by taking into account the influence of the current emotion and given stimulus.

This paper’s remaining sections are organized as follows: the state of the art describes the current research in emotional-based behaviour and some of the currently accepted emotional models, on which we base our system’s internal emotion representation. The method description specifies how the agent is configured and the calculus of the emotional state after an external stimulus occurs. The proof of concept section describes the model’s prototypical implementation and discusses the results obtained so far. Finally, the last section presents our conclusions and future work plans.

STATE OF THE ART

Emotion modelling significantly influences the believability of an agent. As result this is an active research topic in the field of artificial intelligence (Guojiang et al. 2010, Hu et al. 2010). Most emotional models are based on the psychological understanding of emotions, being important to note that, so far, there is no universally accepted theory. Hu et al. states that, from a psychological perspective, emotions are evaluations of the relation between oneself and his surrounding environment. The emotional reaction is generated by external stimuli and affected by the demands and requirements of each individual. From an agents’ perspective, emotions are the evaluations and reactions that are based on its internal state and the relation of its desires and plans towards an exterior environment (Hu et al. 2007). Thus, the model sees emotions as the result of the cognitive process. An example, of such a theory, is the OCC model (Ortony 1988, Steunebrink et al. 2009), which admits the existence of 22 separate emotions and categorises them according to the agent’s circumstances and situations. This division is performed based on the positive and negative factor (valence) of emotions that result from specific events and the interactions with other agents and objects. In our model, we concentrate on the agent’s emotional state as a means of altering its analysis and judgment of each possible action, the process through which it selects the most appropriate one to suit its desires or goals.

For a broader understanding of our model, consider the classic Belief-Desire-Intention (Rao1 and Georgeff 1995) model extended with the OCC model of emotion, presented in Figure 1 from (Parunak and Bisson 2006). The agent’s Analysis of its current state is fed by its Beliefs, which are inferred from the environment by its sensors (Perception), and its Desires, which may or may not change over time. In turn, this analysis leads the agent to an Intention, through which it interacts with the environment. In the extended BDI model, Beliefs also feed an Appraisal process that alters the agent’s current emotional state. This emotional state may then influence the agent’s Perception of the observed environmental stimuli, his analysis of it, or both. In our work, we deviate from the OCC model, using Russel’s model instead to continuously model emotion and allow the agent’s emotional state to affect its Perception of the environment. In turn, this re-interpretation of the environment affects the agent’s Beliefs and ultimately, its Analysis of the current situation. Parallel to this, we also allow the agent’s Emotion to directly affect its Analysis. In our experiments this is done by imposing an emotional stress tolerance that the agents may not exceed when choosing how to interact with the environment. For a more comprehensive description, we refer the reader to the proof of concept section, where we detail the created environment and the agent’s perception and interaction with it.

![Fig. 1. BDI enhanced with OCC, adapted from (Parunak and Bisson 2006).](Image)

**Computational Models of Emotion**

It is also important to note that emotions are classified according to two (Guojiang et al. 2010) categories, which are associated with different theories: the basic emotions theory, such as described by OCC and Ekman (Ekman and Friesen 1971) and the dimensional theory, where emotions are classified according to a coordinate system. An example of the latter is presented by Russel (Russel 1991). In this model, emotions are represented in a two dimensional referential, where the abscissa represents arousal and the ordinate is associated to valence (the hedonic component) (Russel 1991). An illustrative diagram, extracted from (Guojiang et al. 2010) is shown in Figure 2. Another example of the dimensional theory is Plutchik’s emotion wheel (Plutchik 1980), also shown in Figure 2. In this model, the vertical axis represents each emotions’ intensity and the circular component represents the similitude between them. Due to the familiarity and ease of computation presented by Russell’s model cartesian plane, it has been widely accepted by computer scientists, being one of the most popular computational models of emotion.
but that also takes into account the agents plans and goals. Guojiang et al. proposed a method that defines the emotions with the Russel model and combines them with motivations and behaviour decisions based on the Markov Decision Process (Guojiang et al. 2010). Finally, Bryson and Tanguy defined a framework for creating behaviour (Bryson and Tanguy 2010) specifically thought for the easy definition of virtual intelligent actors. In this work the emotions have onsets, sustains and decays, which is extremely important when simulating lasting emotions. They also added moods that reflect a general state of mind over a longer period of time than that exhibited by emotions.

The main problem with the current models is the difficulty associated with controlling the way the emotions are influenced by external stimuli, usually requiring the user to define a set of parameters that directly control the weights of each one. The proposed work is based on a bi-dimensional emotion model used to define agents’ personalities, which control the emotional variations that an external stimulus induces in the current emotional state. Besides presenting a simple method for defining emotional behaviour, our method also offers a seamless integration with existing architectures (e.g. BDI), such that its integration in novel work and subsequent comparison is straightforward.

METHOD

The presented method has the main goal of simulating the changes that occur in the emotional component of the human behaviour when an external stimuli occurs. The method relies on Russel’s bi-dimensional representation of emotions (Russel 1991) to define and control the agent’s personality and current emotion, which are associated to the given stimuli and current mood. The feature space is circular to maintain all points comparable based on their distance to the origin. Thus, each emotion is defined as a two coordinate set in this Cartesian plane. The architecture is divided in three layers: the input abstraction layer, responsible for the mapping between the stimuli and their emotional weights, the stimuli evaluation layer (SEL), that keeps track of the agent’s emotion and alters it according to the observed stimuli and personality, and the output abstraction layer, in charge of the action decision process. The described architecture can be seen in the Figure 3.

The input abstraction layer defines the mapping between the external stimuli and their emotional weight. Like emotions, the weight is defined by two coordinates in the emotional feature space. A stimulus is an abstract entity and can thus represent anything from socialinteractions to abstract sensory stimuli, such as an image or sound, as long as it has the two coordinates associated. It is also necessary to take into account that in the real-world stimuli do not usually occur in an isolated manner. Therefore, in the case of multiple

**Emotion Simulation Mechanisms**

Traditionally, the creation of artificial agents relies mostly on the definition of a rule-set that guides the agent’s behaviour and its action decision mechanism(s). In fact, this is still the most common and accepted approach in various applied/industrial research applications and multimedia industry, where the virtual characters are controlled using scripting, finite-state machines and/or visual logic representations, such as graphs. However, this approach usually leads to stiff behaviours, i.e. always following the same patterns, which inflicts a negative impact on the player experience. As a result, suspension of disbelief is broken by inadvertently gaining insight to the agent’s logic mechanism(s). This leads to lessened enjoyment and believability factors and ultimately shallower immersion levels (Nacke and Lindley 2010), all of which are essential aspects of the gaming/multimedia experience, thus presenting a critical issue. Consequently, work has been done with models capable of realistic behaviour in view.

A recently popular method of solving the behaviour repetitiveness is to add a fuzzy layer, thus introducing a certain degree of entropy to the behaviour selection process. However this may lead to behavioural incoherences if the entropy levels are not kept in check. In order to avoid this issue, formal rule verification mechanisms must be employed, resulting in an overly-complex and costly development cycle. An also popular alternative is to add emotion-simulation to the action-selection mechanism. Hu, et al attempted this by adding emotions to the BDI model (Hu et al. 2007). Their goal was achieving time-coherent action choices, where the choices were based on previous ones.

Chown et al. proposed a cognitive architecture (Chown et al. 2002) that used emotions to influence the decision making process. However the emotion representation was not quantified, thus making it difficult to create a real-world applicable application. In (Hu et al. 2010), the authors also used emotions to influence the action selection process. In their case, emotions were divided into: cognitive (based on the OCC model), expected and reflex. Gratch addressed the problem of simulating the emotional appraisal (Gratch 2000), which refers to the emotion that results not only from an external stimuli

**Fig. 2.** Two dimensional emotional classification models. On the left is the Russel model (Russel 1991), only with two dimensions. On the right is the Plutchik three dimensional model (Plutchik 1980).
concurrent stimuli there are (at least) two possible approaches. The first one is to simply consider the event as a sequence and sequentially feed them to the emotional response layer. The second alternative is to combine them according to their weights and create a new emotional weight. However, these two approaches are not equivalent, as in the former each stimuli would be analysed differently taking into account the result of the previous ones. On the other hand, the latter approach equates to applying only one stimuli. More formally, the emotional weight of a linear combination of stimuli is not necessarily equal to the individual sum of their weights. As such, caution is advised when selecting an alternative. Since the proposed method only takes an emotional weight it is up to the user to define how they are obtained and pre-processed.

Figure 3: Two dimensional emotional classification models. On the left is the Russel model only with two dimension. On the right is the Plutchik three dimensional model.

The agent’s emotional response is the Stimuli Evaluation Layer’s (SEL) responsibility. This layer contains the Emotional State Internal Representation (ESIR), describing the agent’s current mood, i.e. its current emotion and personality model (PM). In fact, the way the personality is defined and how it affects the mood is the core component of the layer, also being the main contribution of the present article. An agent’s personality is defined through a height-map, such as shown in Figure 4. The concept of height-map is usually associated with computer graphics, where it is used to represent the height of a terrain or texture, based on the brightness of each pixel. In the presented method, these brightness values represent the force that each valley in the height-map exerts on the perceived stimuli. A simple metaphor would equate this force to a gravitational pull towards a gravity well (the aforementioned valley). This gravitational pull influences the perception the agent has of the stimuli by simulating his emotional bias towards a specific emotion or group of emotions. In reference to Figure 4, the darker an area is for a corresponding are of the emotional space, the more it pulls the stimulus emotional vector towards it. As a result, if a stimulus tries to move the agent’s current mood away from a gravity well, the stimulus’s strength will be reduced, representing the countering gravitational vector. On the other hand, if the stimulus is in the same direction as the gravity well, the stimulus strength will be increased. Several personalities can be defined in the Personality Archive (PA) and each agent has only one associated with it, at each given time. An example can be seen in Figure 4.

A personality valley (or gravitational well, if using the gravitation analogy) is defined by its center coordinates, height, radius and by the definition of how the height at each point varies. There are multiple ways to define such height variation, it is possible to simply obtain the height based on a linear variation obtained from the distance to the center, or based on a gaussian distribution. It is up to the user to define how the height is obtained based on the previous parameters. In Figure 4 the variation from the valley’s center follows a Gaussian distribution.

Figure 4: An example of a personality height-maps with the outlook of a typical depressive person that after becoming sad it is hard to change to another emotional state.

The change of the current mood is done within the Stimuli Response Logic, SRL and is based on formula (1), where the $m_c$ represents the current mood’s value, defined by its two coordinates. This value is summed to the influence of the nearest personality valley, obtained from the multiplication of the $v_d$, i.e. the vector from the current mood to the center of the valley, and the height $h_c$ of the current mood, given by the distribution function. In the case of the $m_c$ being exactly in the middle of two or more personality valleys, the $v_d$ is the cross product of all the vectors that originate from the $m_c$ to each valley. It is important to note that the closer the current mood is from the valley center, the smaller
the \( vd \) is and the higher is \( hc \), thus influencing it more. Finally, it is also necessary to add the influence of the stimulus emotional weight \( s_w \), which is multiplied by the inverse of the current \( hc \). With this formula, the current mood is changed based on the stimulus and the personality.

\[
m_n = m_c + v_d \cdot h_c + s_w \cdot (1 - h_c)
\]  

(1)

SEL outputs the new current mood each time a stimulus is received, which is then passed to the Output Abstraction Layer (OAL). This layer is responsible for the choice of the agent’s behaviour, as a reaction to the provided stimulus. There are multiple ways to choose an action based on the current emotional state; however they usually have a common dictionary that translates a space coordinate to the desired action. Usually, action decision mechanisms rely on the creation of an emotion-to-action dictionary. Some examples in this domain are: choose the action based on the distance from the current mood to the emotional coordinates of the actions defined in the dictionary since, like the stimuli, actions have to be defined with 2 dimensional coordinates. It is then just a matter of calculating the euclidean distance to all the actions and choosing the one that is nearest. Another option is to define areas for each action in the feature space and choosing the one that has the current mood inside it. A third option is the obtain the action directly from the current emotion through a parameter interpolation technique, which is useful e.g when generating a 3D character expression from the current mood.

**PROOF OF CONCEPT**

The proposed architecture was implemented in C++ and validated by analysing the variations produced in the agent’s emotional state by a stimuli set. The tests were performed on 2 personality models and for a series of 500 and 1000 random stimuli chosen from the game event dictionary. The emotional variations associated with the depressive personality are presented in the form of a scatter graphic shown in Figure 5.

**Discussion**

The graphics show that the current emotion tends to converge to the personality valleys previously defined. However, this does not prevent the current emotion from moving outside each valley’s influence radius, thus converging to the personality valleys. It is then possible to conclude that the personality model successfully influences the agent’s emotional states in the intended way. However, as seen in Figure 4, these initial tests revealed an interesting issue with the proposed model: the high concentration on the limits of the circumference. This phenomena is easily attributed to two causes: one of human error and a second central to the model itself. The first of these, was that the manually created stimuli were too high (in the range of 0.3 and 0.6 Arousal or Valence units), which means that if a current emotion is already near the limit, it will attempt to exceed it and thus clipping occurs. The other problem is a consequence of not simulating the emotional decay that occurs over time, i.e. as time passes after even a strong stimuli, its effect will degrade and thus becomes less effective in influencing the current emotion, resulting in the current emotion decaying either to a neutral emotional state or to the personality’s main tendency. As this is not simulated the current emotion always stays at the limit until an opposite stimuli is perceived. It is also important to note that in our tests generated stimuli’s distribution is completely random. However, this may not be the case in reality, as most experiences have a limited range of emotional states which they elicit (their emotional spectrum, so to speak). For example, when watching a horror movie, it is unlikely that positively valenced stimuli arise. Therefore, this must also be taken into account when interpreting these results or trying out new personalities for specific case studies.

![Figure 5: Personality model for a depressible person and scatter graphic after 1000 randomly chosen stimuli from the dictionary. The scatter graphic shows the distribution of the emotional variations, based on the personality model, that occur after all the stimuli were applied.](image)

**CONCLUSION & FUTURE WORK**

Throughout this paper three main contributions were presented. The first one was PERSONA, which proposes a novel architecture for personality-based emotional response simulation. Secondly, the input and output abstraction layers provide a transparent and powerful abstraction of stimuli and their elicited reactions. Finally, and at the core of these three contributions, a simple personality-definition method through the use of height-maps, was invented.

The variations induced by this approach were tested at section 4, validating it as a proof of concept, however further tests are needed. The mood distribution tests showed that the given a random distribution of stimuli, the agent did tend to interpret them according to his personality model, while the trajectory tests proved that our influence model adequately influenced the given
stimuli. Thus, it was proven that the proposed method is capable of describing the psychological tendencies that compose a personality, therefore being able to adequately simulate it’s responses. Some examples of uses of the abstraction layer were given as a way to point into practical uses of the proposed approach, thus we believe that in synchrony with an agent they are a versatile and powerful tool that has real-time in further applications.

The validation tests did however point-out that since the agent has no notion of time, he remains on the elicited emotional state until elicited otherwise, which means he will easily reach the maximum emotional charge allowed. Performing a further analogy to the real world, this is as if all the stimuli were given almost instantaneously to someone, all the while not letting them process the information and emotionally straining them. This issue will be addressed in the future work through the implementation of a stimuli timestamp and an emotional decay function over a time axis. Other aspects that need consideration include: a GUI for increased interaction ease, as defining a personality solely on parameters is time-consuming and error-prone, and the increased versatility of an height-map definition through a mesh instead of Gaussian mixtures. As a final statement, the authors would like to point out that although the proposed approach does not consider all the possibly available information in a real-world scenario (nor does it strive to), both the proposed architecture and imbeded agent logic have shown themselves capable of modelling individuals’ personalities through emotional responses.

ACKNOWLEDGMENTS

The authors would like to thank Instituto de Telecomunicações – Porto and Artificial Intelligence and Computer Science Laboratory for hosting and supporting their respective PhDs research. They also acknowledge the Portuguese Foundation for Science and Technology (FCT) for fully funding the aforementioned PhD research under the scholarships: SFRH/BD/79905/2011 and SFRH/BD/77688/2011.

REFERENCES


AUTHOR BIOGRAPHIES

José Serra, born in 1988, is a Ph.D. student in MAP-i Doctoral programme. He is currently being advised by Verónica Orvalho from University of Porto and Miguel Sales Dias from Microsoft (MLDC). His work deals with combining procedural facial animation with artificial intelligence to create a fully automatic facial animation system based on external stimuli. He has previously worked on the NECO (NEwork COding) simulator, which was integrated in the N-Crave project and is available online for download. He has also worked on the issue of automatically animating a speech based on an audio input as part of his MSc thesis, within the scope of the LIFEisGAME project. Other interests include human computer interaction, artificial intelligence, physics and security. His e-mail address is: jserra@dcc.fc.up.pt and his web-page can at http://www.dcc.fc.up.pt/~jserra.
PEDRO NOGUEIRA is a Ph.D. student at the Artificial Intelligence and Computer Science Lab in Portugal and the HCI and Game Studies Group in the University of Ontario’s Institute of Technology. He is co-supervised by Prof. Eugénio Oliveira and Prof. Rui Rodrigues at the University of Porto, Prof. Lennart Ncake at the University of Ontario. Before starting his Ph.D. studies, Pedro has worked in various European research projects. His work has ranged from conceptualizing and prototyping usable biometric monitoring systems for high-risk emergency scenarios, to developing intelligent biomedical image annotation algorithms for drug research. His current work focuses on the study of how user’s gameplay experience is affected by emotional state variations and how these can be enforced to produce more immersive and enjoyable experiences. To this end he is working on methods to determine players’ reactions to specific game events, which can be collated into time-evolving, self-adapting affective reaction profiles. His final goal is to use these user-independent profiles to provide a biofeedback loop with the necessary information on how it should plan future interactions, effectively implementing an automatic, data-driven regulation of the user’s affective experience. His current institutional e-mail address is: pedro.alves.nogueira@fe.up.pt.