# Generating Norm-Related Emotions in Virtual Agents

Nuno Ferreira<sup>1</sup>, Samuel Mascarenhas<sup>1</sup>, Ana Paiva<sup>1</sup>, Frank Dignum<sup>2</sup>, John Mc Breen<sup>3</sup>, Nick Degens<sup>3</sup>, and Gert Jan Hofstede<sup>3</sup>

 <sup>1</sup> INESC-ID and Instituto Superior Técnico, Technical University of Lisbon, Av. Prof. Cavaco Silva, Taguspark 2744-016, Porto Salvo, Portugal
<sup>2</sup> Utrecht University 3508TB, Utrecht, The Netherlands
<sup>3</sup> Wageningen University Hollandseweg 1, 6706 KN Wageningen, The Netherlands

**Abstract.** The increased believability provided by emotions in virtual characters is a valuable feature in a multi agent environment. Despite much research on how to model emotions based on events that affect a character's goals, the current emotional models usually do not take into account other sources of emotions, such as norms and standards. Moreover, current normative systems usually do not consider the role of emotions. Systems that include emotions and norms are too domain-specific or lack flexibility. We propose a model for the generation of emotions based on the appraisal of actions associated with norm-related events, such as the fulfilment or violation of a norm.

Keywords: Emotions, Norms, Appraisal, Standards, Believability.

# 1 Introduction

Virtual environments often try to simulate social situations where agents should follow existing norms. We argue that for social agents to be believable, they should have emotional reactions related to the importance of the norms which are fulfilled or violated (by choice or necessity). However, despite research on how to model emotions and how to model norms in virtual agents, there is no link yet established between norm-related actions and the emotions that can arise by witnessing such actions. We propose a model to generate emotions in virtual agents that result from the evaluation (appraisal) of actions that are perceived to cause fulfilment or violation of norms. The normative emotional agents were then integrated into an architecture for virtual agents, and tested in a scenario. The results of a preliminary user study indicate that the emotional responses produced by the model were perceived by users and correlated to how strongly the user believed that the norm was important for the agents.

In section 2, we present related work on normative systems and emotions. In section 3, we present our appraisal model of norms. In section 4, we present a case study and the evaluation. Finally, we draw conclusions and discuss future work.

Y. Nakano et al. (Eds.): IVA 2012, LNAI 7502, pp. 97-104, 2012.

<sup>©</sup> Springer-Verlag Berlin Heidelberg 2012

#### 2 Related Work

There has been considerable research on how to use emotions to increase the believability of synthetic characters. Traditional animators suggest that to properly portray the emotional reactions of a character, the emotions must affect the reasoning process and consequences should be noticeable in the actions of the characters [18]. This led the computer science community to use emotion theories from psychology that model the generation of emotions in humans [3] [14] [2] [8] [9] [11].

However, emotional characters must balance their personal goals with their social environment to be believable. Normative models were developed to solve this problem without imposing hard constraints. A well-known normative system in virtual agents is *Thespian* [16]. In Thespian, obligations are created when an agent performs a certain action towards another agent. To satisfy the obligation the target agent must choose a proper action in response. Another normative system is *culturally affected behaviour* (CAB) [17], which focuses on so-called cultural norms. In this system, norms are represented using graphs named so-ciocultural networks, where actions are linked to cultural norms with a value that indicates whether the action conforms to the norm.

Some emotional models try to model norm-related emotions without a representation of norms, by casting norm violations as goal violations [10] or include norms in very specific domains [8]. Some normative systems, including Thespian and CAB, were further extended with emotional models [15] [4]. But those models typically are too domain-specific or lack flexibility. We argue that not only do emotions play a fundamental part in norm-related decision processes, but that the norms themselves influence the emotional state. Thus, virtual agents that connect emotions and norms will be far more believable.

#### 3 Linking Norms and Emotions

We aim to generate emotions in virtual agents by the appraisal of actions associated with the fulfilment or violation of a norm. Hence, our agents need to have a normative model so that they can recognize norms, and when they are fulfilled or violated. The agents must also have an emotional model that evaluates the actions of agents, and generates an emotional response based on their goals and standards.

Norms prescribe what behaviours are expected in a certain social context. In our model, a norm is specified by its activation conditions, which mark the norm as active, and its expiration conditions. The behaviour prescribed by a norm is represented by a set of conditions, called normative conditions. The agent should try to satisfy these conditions, if the norm is an obligation, and avoid them, if the norm is a prohibition. If the agent succeeds, the norm is fulfilled, when it fails, the norm is violated. Our norm model is based on the work presented in [19], [5], [7] and [12]. A norm contains the following components:

- ID: A unique identifier that is used to identify the norm.
- Name: A name that describes the norm.
- Type: A value that informs if the norm is an obligation or a prohibition.
- Targets: The agents that are expected to fulfil the norm (when active).
- Activation Conditions: Conditions that cause the activation of the norm.
- Expiration Conditions: Conditions that cause the expiration of the norm.
- Normative Conditions: Prescriptions for the behaviour of the targets of the norm.
- Salience: A value that "indicates to an individual how operative and relevant a norm is within a group and in a given context" [1]. The salience of a norm depends on several contextual, social and individual factors (cues), such as the level of compliance and the frequency of punishment.

In our model agents monitor their own norms that they (and others) should observe. Each agent has a Normative Environment to store information about norms, whether they are active, recently expired, or were fulfilled or violated. Obligations are fulfilled when the normative conditions become true and violated if they expire without being fulfilled. Prohibitions are fulfilled as long as their normative conditions remain true, and violated when they become false.

Our emotional model follows the OCC Appraisal theory of emotions (named after its creators Ortony, Clore and Collins) [13]. According to OCC, the appraisals focused on how actions conform or not with internalized standards will trigger "Attribution Emotions" (pride, shame, admiration and reproach). Pride and shame occur when the agent is appraising its own actions as praiseworthy or blameworthy, respectively, while admiration and reproach arises from appraising the actions of others as praiseworthy or blameworthy.

According to OCC, the praiseworthiness of an action is often assessed in terms of its (perceived) social value. So, in our model, actions that cause the fulfilment of a norm are considered praiseworthy while actions that violate norms are blameworthy. Four factors determine the value for the praiseworthiness or blameworthiness: the salience S of the norm  $(S \in [0, 1])$ , the estimated cost C of the action  $(C \in [0, 1])$ , if the action was intentional I or not  $(I \in [0, 1])$ , and if the agent is responsible R for the action  $(R \in [0, 1])$ .

The praiseworthiness P is given by  $RI(SW_s + CW_c)$ , with  $W_s$  and  $W_c$  being a weight for the salience and for the cost, respectively. An action is only praiseworthy when the agent is perceived as having the intention and the responsibility for it. If so, this value is proportional to the salience of the norm and the cost of the action. The blameworthiness B is given by R [(SWs + CWc) Wi (1 I)], where the factor Wi reduces the blameworthiness of less intentional actions. It is also related to the salience and the cost, and zero if the agent is not perceived as responsible.

Another appraisal variable that can influence the intensity of the attribution emotions is the expectation-deviation D. For instance, the admiration we would feel for a fire-fighter saving the life of a child is likely to be less intense than the admiration that we would feel if it was the child who saved the fire-fighter's life, because the latter deviates more from what is expected. In our model, the expectation-deviation is 1-S if the norm is fulfilled and S if the norm is violated.

The intensity of the attribution emotion is given by  $PW_p + DW_d$  if the norm is fulfilled and  $BW_b + DW_d$  if the norm is violated, where  $W_p$ ,  $W_b$  and  $W_d$  are weights for the praiseworthiness, blameworthiness and expectation-deviation, respectively.

# 4 Case Study and Evaluation

We implemented our model in an agent architecture called FAtiMA [6], a BDI architecture that endows agents with the ability to generate emotional reactions to events, based on the OCC model but in which there was no explicit notion of norms. With the addition of our model, agents constantly check if any norm becomes active or expires. Every time that an agent perceives a new event, it will check if it is an action of an agent that causes the fulfilment or violation of a norm. When a norm fulfilment is detected, the agent appraises that event and computes its praiseworthiness and expectation-deviation to determine the intensity of the resulting emotion.

Using the extended architecture, two versions of a simple social scenario were created. The scenario occurs in a bar where the user plays the role of a character that is sitting at a table with two friends (a smoker and a non-smoker) and there is a prohibition to smoke, as described in an introductory text. We made two versions of this scenario where the only difference was the salience (all weights were set to 0.5, intentionality and responsibility to 1 and cost to 0) of the nonsmoking norm (see Figure 1). In the low-salient version, the salience of the norm is set as 0.1. A friend starts smoking, the the non-smoker character perceives that as a norm violation and appraises the event as blameworthy. However, the blameworthiness is so low that it is not enough to exceed the threshold for triggering a reproach emotion, thus no emotional expression is made. In the high-salient version, the salience of the norm is set as 0.9. The smoker friend still smokes since the norm, while important, is still not as important as its goal to smoke. When the non-smoker friend perceives this norm violation, it appraises the action as very blameworthy, feels a strong reproach emotion, and reacts with a frown expression and the background character gestures his annovance.



**Fig. 1.** In the low-salience version (left image) the non-smoker does not react emotionally, while in the high-salient version (right image) the non-smoker reacts with a frown expression

The bar scenario that was previously described was used to conduct a small pilot study. The aim of the study was to investigate if users would perceive differences in the emotional response of the agents and if those differences would relate to the specified salience of the smoking ban in the virtual environment.

Participants were randomly assigned to interact with one of the two versions of the virtual bar that were previously described. Besides the different value assigned to the salience of the smoking ban norm, all of the other parameters of the agents in the two versions (goals, relations, properties) are exactly the same.

After they interacted with one of the versions, subjects were asked about which emotions did they agreed (using a 7-point Likert scale) that the non-smoking character felt after witnessing his friend lighting a cigarette. The rationale for these questions was to check if the frowning expression of the non-smoking friend was being correctly interpreted as an emotional response.

Participants were then asked if they agreed that from the perspective of the characters the smoking ban was important and if it was acceptable to smoke inside the bar. A 7-point Likert scale was used for both questions as well. Finally, we asked participants if they smoke and also their gender, age and nationality.

In total, we had 17 Portuguese subjects (82% male), aged between 22 and 40, with the average age being 27. A total of 8 participants interacted with the low-salience version of the virtual bar and the other 9 with the high-salience version. Figure 2 shows the results obtained.



**Fig. 2.** The left side shows the results for the perception of the non-smoker's emotional state. The right side depicts the results obtained for the perceived relevance of the norm. (1 - Strongy Disagree, 7 - Strongly Agree)

Regarding the perception of the emotional state of the non-smoker character after the norm is violated, we found the following significant differences. In the high-salience version, the one in which the non-smoker frowns, participants agreed significantly more that the character was feeling upset (U = 19, z =-1.7, p = .046, r = -.41), offended (U = 16, z = -1.98, p = .024, r = -.48) and angry (t(15) = -2.37, p = .016, r = .52). On the other hand, in the the lowsalience version subjects perceived the character as more amused (U = 12, z =-2.43, p = .008, r = -.59). There were no significant differences in the emotions of surprise, disgust, shame and embarrassment. Overall these results indicate that subjects, as expected, detected a significant change in the emotional state of the character after the norm is violated in the high salience version.

Concerning the questions about the perceived relevance of the smoking ban in the perspective of the characters, as shown in Figure 2 participants did in fact attribute a significantly higher importance (t(8.6) = -2.114, p = .032, r = 0.58) in the high-salience version. They also thought that it was more acceptable to smoke inside the bar in the low-salience version (U = 11, z = -2.47, p = .0065, r =-.60). To examine the link between these results and the non-smoker's emotional state, we run a Pearson's correlation test between the two. Concerning the user's perception of how important was the norm in the character's perspective it was significantly correlated with the perception of the non-smoker character being upset (r = .42, p = .046) and being angry (r = .56, p = .01). Similarly, the perception of how acceptable was for the characters to smoke inside the bar was significantly correlated with the non-smoker character being upset (r = -.68, p = .001), being angry(r = -.76, p < .0001) and also being offended (r = -.64, p = .003). Although preliminary, the results obtained suggest that users were able to perceive a relationship between the emotions generated by our model and the specified salience of the norm in the scenario. This is an important result because it indicates that generating these kind of emotions from the specified norms of a multi agent environment can help users to better understand the social context the agents are simulating.

# 5 Conclusion

We argued that the link between norms and emotions is important to consider when modelling virtual agents, as norm-related events can be appraised and trigger emotions that will increase the character's believability. We proposed a normative model for agents to be able to recognize when norms are fulfilled or violated by actions, and an emotional model capable of generating emotions when agents witness such events. The proposed model was then integrated in an architecture for virtual agents to create two versions of a scenario where the user interacted with characters with different needs and goals, that reacted emotionally to the violation of a norm. In one version this norm had a low salience and in the second version, the salience was high. A small pilot study was conducted in which a group of participants interacted with one of the two versions created. The aim was to see how users interpreted differences in the agents emotional behaviour, with those differences being generated by our model. The results suggest that users did relate the differences in the versions to the importance of the norm. As future work we plan to extend the model by introducing enforcing mechanisms and to conduct further tests.

Acknowledgments. This work was partially supported by the Portuguese Fundação para a Ciência e a Tecnologia (INESC-ID multiannual funding) through the PIDDAC Program funds, by the European Community (EC) and by a scholarship (BD/62174/2009) granted by FCT. Additionally, it was funded by the SEMIRA (ERA-Compl/0002/2009), eCute (ICT-5-4.2 257666) and GALA (ICT-2009.4.2) projects. The authors are solely responsible for the content of this publication. It does not represent the opinion of EU or the FCT, which are not responsible for any use that might be made of data appearing therein.

## References

- Andrighetto, G., Villatoro, D.: Beyond the Carrot and Stick Approach to Enforcement: An Agent-Based Model. In: European Perspectives on Cognitive Science (2011)
- Bates, J.: The nature of characters in interactive worlds and the Oz project. In: School of Computer Science, Carnegie Mellon University, Pittsburgh, PA Technical Report CMU-CS-92-200 (1992)
- Bates, J.: Virtual Reality, Art and Entertainment. In: Presence 1.1, pp. 133–138 (1992)
- 4. Bulitko, V., et al.: Modeling Culturally and Emotionally Affected Behavior. In: Artificial Intelligence and Interactive Digital Entertainment Conference (2008)
- Castelfranchi, C., Dignum, F., Jonker, C.M., Treur, J.: Deliberative Normative Agents: Principles and Architecture. In: Jennings, N.R. (ed.) ATAL 1999. LNCS, vol. 1757, pp. 364–378. Springer, Heidelberg (2000)
- Dias, J., Paiva, A.C.R.: Feeling and Reasoning: A Computational Model for Emotional Characters. In: Bento, C., Cardoso, A., Dias, G. (eds.) EPIA 2005. LNCS (LNAI), vol. 3808, pp. 127–140. Springer, Heidelberg (2005)
- Dignum, F.: Autonomous agents with norms. In: Artificial Intelligence and Law 7.1, pp. 69–79 (1999)
- 8. Elliott, C.D.: The Affective Reasoner: A process model of emotions in a multi-agent system. PhDThesis Northwestern University 1992 (1992)
- Gratch, J.: Émile: Marshalling passions in training and education. In: Proceedings of the Fourth International Conference on Autonomous Agents, pp. 325–332. ACM (2000)
- Gratch, J., Mao, W., Marsella, S.: Modeling social emotions and socialattributions. Cambridge University Press (2006)
- Marsella, S.C., Johnson, W.L., LaBore, C.: Interactive pedagogical drama. In: Proceedings of the Fourth International Conference on Autonomous Agents, pp. 301–308. ACM (2000)
- Oren, N., Panagiotidi, S., Vázquez-Salceda, J., Modgil, S., Luck, M., Miles, S.: Towards a Formalisation of Electronic Contracting Environments. In: Hübner, J.F., Matson, E., Boissier, O., Dignum, V. (eds.) COIN 2008. LNCS, vol. 5428, pp. 156– 171. Springer, Heidelberg (2009)
- Ortony, A., Clore, G.L., Collins, A.: The Cognitive Structure of Emotions. Cambridge University Press (1988)
- Scott Neal Reilly, W.: Believable Social and Emotional Agents. PhD thesis. Citeseer, p. 288 (1996)
- Si, M., Marsella, S.C., Pynadath, D.V.: Modeling appraisal in theory of mind reasoning. In: Autonomous Agents and Multi-Agent Systems 20.1, pp. 14–31 (2010)

- Si, M., Marsella, S.C., Pynadath, D.V.: Thespian: Modeling Socially Normative Behavior in a Decision-Theoretic Framework. In: Gratch, J., Young, M., Aylett, R.S., Ballin, D., Olivier, P. (eds.) IVA 2006. LNCS (LNAI), vol. 4133, pp. 369–382. Springer, Heidelberg (2006)
- 17. Solomon, S., et al.: A language for modeling cultural norms, biases and stereotypes for human behavior models. Tech. rep. DTIC Document (2008)
- Thomas, F., Johnston, O.: Disney Animation: The Illusion of Life. Abbeville Press (1981)
- Villatoro, D., et al.: Dynamic Sanctioning for Robust and Cost-Efficient Norm Compliance". In: 22nd International Joint Conference on Artificial Intelligence, Barcelona, Spain, July 16-22 (2011)