Role of Emotions in Perception of Humanness of Virtual Agents

Extended Abstract

Moojan Ghafurian
School of Computer Science,
University of Waterloo
moojan@uwaterloo.ca

Neil Budnarain
School of Computer Science,
University of Waterloo
bbudnarain@edu.uwaterloo.ca

Jesse Hoey
School of Computer Science,
University of Waterloo
jhoey@cs.uwaterloo.ca

ABSTRACT

We analyzed the performance of an agent based on an appraisal theory of human emotion with respect to how it modulates play in a social dilemma game. An experiment with 117 participants showed how the agent was rated on dimensions of Human-Uniqueness (HU), separating humans from animals, and Human-Nature (HN), separating humans from machines. We showed that our appraisal theoretic agent significantly improved on both HN and HU ratings, compared to the baselines. We also showed that perception of humanness positively affects cooperation and enjoyment.

KEYWORDS

Emotions; human likeness; OCC; prisoner's dilemma; virtual agents

ACM Reference Format:
Moojan Ghafurian, Neil Budnarain, and Jesse Hoey. 2019. Role of Emotions in Perception of Humanness of Virtual Agents. In *Proc. of the 18th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2019), Montreal, Canada, May 13–17, 2019,* IFAAMAS, 3 pages.

1 INTRODUCTION

The concept of "human-ness" has seen much debate in social psychology, particularly in relation to work on stereotypes and dehumanization [3]. Haslam *et al.* [8] examined how people judge others as human or non-human (dehumanized). In their model, *humanness* is broken down into two factors: (1) *Human Uniqueness* (HU) distinguishes humans from animals, and (2) *Human Nature* (HN) distinguishes humans from machines. HU traits are civility, refinement, moral sensibility, rationality and maturity. HN traits are emotionality, warmth, openness, agency (individuality), and depth. While much research in AI focuses on building machines with HU traits, there is much less work on improving HN traits.

Several studies have verified that humanness and emotions of virtual agents can affect people's behaviour and strategies (e.g.,[1] and [2], respectively). Here, we develop a virtual agent in the same spirit as EMA [7], where emotional displays are made using the Ortony, Clore and Collins (OCC) model [10], and a set of coping rules map the game history augmented with emotional appraisals to actions for the virtual agent. This "*OCC* agent" uses emotions to generate expectations about future actions [5, 14]. We evaluate how appraised emotions relate to two dimensions of humanness, and study the impact of emotion modeling on users' cooperation.

Proc. of the 18th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2019), N. Agmon, M. E. Taylor, E. Elkind, M. Veloso (eds.), May 13–17, 2019, Montreal, Canada.

2 EXPERIMENT

We present results from a study involving 117 participants who played a simple social dilemma game with a female virtual agent named "Aria". The agent was originally developed for speech and language therapy [12]. The game was a variation of Prisoner's Dilemma (PD), in which each player could either give two coins or take one coin from a common pile. Participants could also send an emotional signal by choosing one of 20 labeled emoji. The participants were awarded a bonus according to their total score in the game, and so had incentive to cooperate.

Following the OCC Model [11], momentary emotions were appraised on the *consequences of events* and the *actions of agents*.[1] Both were appraised for both self and other. The prospect-based *consequences of events* were evaluated on each subsequent turn. Two key examples are shown in Table 1.[2] In the first, Aria takes, but the player gives and shows a positive emotion, and Aria is pleased about the prospects of events, and approving of the player's action, leading to emotions of hope and gratitude, but disapproving of her own action, leading to shame. In the second, the player defects after a round of cooperation, and shows no regret, leading to negative emotions of fear, disappointment and reproach. Five coping strategies taken from [7] (*acceptance, seeking support, restraint, growth, and denial*) gave Aria's next action. In the first example in Table 1, Aria copes with a newly cooperative player with *growth* and cooperates. In the second, Aria copes with a sudden defection with *denial* and cooperates as well. At the start, Aria copes with an initial emotion of hope by seeking support, and thus cooperates.

Aria's facial expressions are generated with 3 controls corresponding to three dimensions of "Happy/Sad", "Surprise/Anger" and "Fear/Disgust" ('HSF' space) that mapped to sets of facial muscles. We mapped OCC emotion words and the endpoint words of HSF space to the Evaluation/Potency/ Activity (EPA) space using [4], and used a distance-based mapping to map to HSF space. Utterances were selected from 8 manually generated sets, one for each combination of agent/human action and utterance valence: (1) an embedding (vector) for each emotion label was computed using the pre-trained Word2Vec model [9], (2) the mean of the Google embeddings of all the words in a phrase were calculated, and (3) the closest phrase to an emotion label was found by cosine distance.

Experimental Conditions: Participants were recruited on Amazon Mechanical Turk (74 male, 48 female, 1 other, and 1 did not wish to share, age: [21,74]). Participants received an initial payment of $0.7 and a bonus according to their performance in the

[1]In our game, the *aspects of objects* will not change substantially over the course of the interaction. We therefore focused on actions and events only.
[2]For the full table and other details see [6]. We are grateful for funding from NSERC and SSHRC, and thank Nattawut Ngampatipatpong, Robert Bowen, and Sarel van Vuuren (University of Colorado, Boulder) for help with the virtual human.

Table 1: Emotions/actions based on OCC decision tree/coping. 🙂: “pleased”, ☺: approving. ♡: desirable, and ✓: confirmed.

GAME PLAY				VALENCED APPRAISALS								APPRAISED EMOTIONS			NEXT MOVE BASED ON THE COPING STRATEGY	
				Consequences						Actions of agents						
				other		self										
Previous	Most Recent					prospects relevant?										
	Move		Emotion		♡	yes			no			Momentary		Prospect-Based		
					for	🙂?				self	other				Coping	Action
Player	Aria	Player	Player	🙂?	other?	future	present	✓	🙂?	☺?	☺?	Single	Compound	Single		
take 1	take 1	give 2	positive	no	no	yes	yes	yes	yes	no	yes	pity joy,admiration shame	gratitude	hope relief	growth	give 2
give 2	give 2	take 1	no regret	no	yes	no	no	no	no	yes	no	resentment distress,reproach pride	anger	fear disappointment	denial	give 2

game ($0.05 per point). The data from 7 participants (5 male and 2 female) were removed as they failed to pass the attention checks or the interface did not work. [3] The participants were evenly split into three conditions: (1) **OCC:** Agent's emotional displays and next actions were selected from the set corresponding to the current game play (examples shown in Table 1); (2) **Emotionless:** Agent played tit-for-2-tats, showed no emotional expressions in the face, and did not talk; (3) **Random:** Agent played tit-for-2-tats. Emotions were randomly drawn from a set of 20 emotions. Facial expressions and utterances were selected from that emotion label. Participants played for 25 rounds, and then answered four questionnaires: demographic, HU/HN (from [8]), enjoyment, and tendency to anthropomorphise (IDAQ [13]).

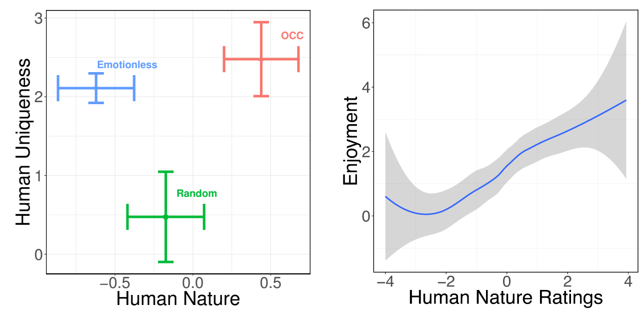
3 RESULTS

Humanness: Figure 1 (a) shows the results. As hypothesized, *OCC* agent was perceived to be more human-like on both HN and HU traits. Two linear mixed effect models were fit to predict HN and HU ratings based on experimental condition. General tendency to anthropomorphize (IDAQ), age, gender, and bonus were controlled for. The *OCC* agent's HN traits were perceived to be significantly higher than the Emotionless agent, and its HU traits were perceived to be significantly higher than the Random agent. Overall, the *OCC* agent was rated as significantly more human-like than baselines.

Cooperation/Trust: All agents played the same strategy (i.e., tit-for-two-tats); therefore, the difference in cooperation rates among conditions can reflect the effect of the different emotional displays on participants' tendency to cooperate (i.e., trusting the agent). OCC had the highest cooperation rate and encouraged cooperation. This difference was significant between OCC and Random agents ($se = 1.851, t = -2.006, p < 0.05$), but not however, between the cooperation rates of OCC and Emotionless agents.

Enjoyment: A linear model was fit to study how perception of HN and HU traits affected users' enjoyment. IDAQ and final bonus were controlled. Perception of HU traits did not seem to affect users' enjoyment, however, perception of HN traits significantly affected enjoyment in the game. In other words, playing against an agent that was perceived to be more human-like in HN traits (with the exact same strategy and actions) significantly increased enjoyment in the game. Figure 1 (b) shows the results.

[3] Participation limited to North Americans with more than 50 HITs and an MTurk approval rate of 96%. Approved by the University of Waterloo's Ethics Review Board.



(a) (b)

Figure 1: (a) Humanness ratings (b) Enjoyment based on perception of HN traits. 95% confidence intervals are visualized.

4 DISCUSSION AND CONCLUSION

This paper utilized Haslam *et al.*'s definition of humanness [8] to study how emotions affect perception of *Human Nature (HN)* and *Human Uniqueness (HU)*, distinguishing humans from machines and animals, respectively. We asked how emotions affect perception of HU and HN traits of an agent, therefore, people's opinion and behaviour towards the agent. We hypothesized that agents capable of showing emotions will be perceived more human-like, especially on the HN traits. We used a social dilemma to test this hypothesis in which the players cooperate if they trust the opponent. An agent based on the Ortony, Clore and Collins (OCC) model [10], capable of showing meaningful emotions, was perceived significantly more human-like on both HN and HU traits when compared to random or emotionless agents (baselines). Participants' cooperation rate and enjoyment were also improved. *Any* expression of emotion, even by the random agent, improved perception of HN traits. However, displaying random emotions negatively affected HU traits, as it can make the agent look irrational and immature. That is, while showing proper emotions enables agents with HN traits and fills the gap between humans and machines, showing random emotions that are not necessarily meaningful can be worse than showing no emotions, making the agent more animal-like. This work suggested that while perception of HU traits can be successfully improved by making agents smarter, emotions are critical for perception of HN traits. Emotions aligned with human expectations are important for the development of more human-like virtual agents.

REFERENCES

[1] Andry Chowanda, Martin Flintham, Peter Blanchfield, and Michel Valstar. 2016. Playing with social and emotional game companions. In *International Conference on Intelligent Virtual Agents*. Springer, 85–95.

[2] Celso M De Melo, Peter Carnevale, and Jonathan Gratch. 2010. The influence of emotions in embodied agents on human decision-making. In *International Conference on Intelligent Virtual Agents*. Springer, 357–370.

[3] Stéphanie Demoulin, Jacques-Philippe Leyens, Maria-Paola Paladino, Ramón Rodriguez-Torres, Armando Rodriguez-Perez, and John Dovidio. 2004. Dimensions of "uniquely" and "non-uniquely" human emotions. *Cognition and emotion* 18, 1 (2004), 71–96.

[4] Clare Francis and David R. Heise. 2006. Mean Affective Ratings of 1,500 Concepts by Indiana University Undergraduates in 2002-3. [Computer file] Distributed at Affect Control Theory Website. (2006).

[5] Nico H. Fridja. 2010. The Psychologists' Point of View. In *Handbook of Emotions* (third edition ed.). The Guildford Press, 68–77.

[6] Moojan Ghafurian, Neil Budnarain, and Jesse Hoey. 2019. Improving Humanness of Virtual Agents and Users' Cooperation through Emotions. *ArXiv e-prints* (March 2019). arXiv:1903.03980

[7] Jonathan Gratch and Stacy Marsella. 2004. A domain-independent framework for modeling emotion. *Cognitive Systems Research* 5, 4 (2004), 269 – 306. https://doi.org/10.1016/j.cogsys.2004.02.002

[8] N. Haslam, S. Loughnan, Y. Kashima, and P. Bain. 2008. Attributing and denying humanness to others. *European Review of Social Psychology* 19, 1 (2008), 55–85.

[9] T. Mikolov, I. Sutskever, K. Chen, G. Corrado, and J. Dean. 2013. Distributed Representations of Words and Phrases and their Compositionality. *ArXiv e-prints* (Oct. 2013). arXiv:cs.CL/1310.4546

[10] A. Ortony, G.L. Clore, and A. Collins. 1988. *The Cognitive Structure of Emotions*. Cambridge University Press.

[11] A. Ortony, D. Norman, and W. Revelle. 2005. Affect and proto-affect in effective functioning. In *Who needs emotions: The brain meets the machine*, J. Fellous and M. Arbib (Eds.). Oxford University Press, 173–202.

[12] Sarel Van Vuuren and L.R. Cherney. 2014. A Virtual Therapist for Speech and Language Therapy. In *Proc. Intelligent Virtual Agents, LNAI 8637*. Springer, 438–448.

[13] Adam Waytz, John Cacioppo, and Nicholas Epley. 2010. Who sees human? The stability and importance of individual differences in anthropomorphism. *Perspectives on Psychological Science* 5, 3 (2010), 219–232.

[14] R.B. Zajonc. 2000. Feeling and Thinking: Closing the Debate over the Independence of Affect. In *Feeling and Thinking: The Role of Affect in Social Cognition. Studies in Emotion and Social Interaction*, J. P. Forgas (Ed.). Vol. 2. Cambridge University Press, New York, 31–58.