Check for updates

# Multimodal Affective Pedagogical Agents for Different Types of Learners

Nicoletta Adamo[1(✉)], Bedrich Benes[1], Richard E. Mayer[2], Xingyu Lei[1], Zhiquan Wang[1], Zachary Meyer[1], and Alyssa Lawson[2]

[1] Department of Computer Graphics Technology, Purdue University, West Lafayette, IN 47907, USA
{nadamovi, bbenes, lei64, wang4490, meyer204}@purdue.edu
[2] Department of Psychological and Brain Sciences, University of California Santa Barbara, Santa Barbara, CA 93106-9660, USA
{rich.mayer, alyssa.lawson}@psych.ucsb.edu

**Abstract.** The paper reports progress on an NSF-funded project whose goal is to research and develop multimodal affective animated pedagogical agents (APA) for different types of learners. Although the preponderance of research on APA tends to focus on the cognitive aspects of online learning, this project explores the less-studied role of affective features. More specifically, the objectives of the work are to: (1) research and develop novel algorithms for emotion recognition and for life-like emotion representation in embodied agents, which will be integrated in a new system for creating APA to be embedded in digital lessons; and (2) develop an empirically grounded research base that will guide the design of affective APA that are effective for different types of learners. This involves conducting a series of experiments to determine the effects of the agent's emotional style and emotional intelligence on a diverse population of students. The paper outlines the work conducted so far, e.g., development of a new system (and underlying algorithms) for producing affective APA. It also reports the findings from two preliminary studies.

**Keywords:** Animated pedagogical agents · Computer animation · Emotions · Affective computing · Intelligent tutoring systems

## 1 Introduction

Research has shown that animated pedagogical agents (APA) can be effective in promoting learning, but not equally for all learner populations, learning subjects, and contexts. Despite growing evidence in support of the positive value of APA, many questions still remain unanswered, particularly concerning the emotional design of APA. For instance, it is unclear which types of emotional expression, degree of embodiment and personalization benefit a particular leaner population and why. To advance knowledge in this field and maximize the agent's positive impact on learning, there is a need to further investigate the effects of the agent's affective features, and whether they are moderated by learner characteristics, learning topics, and contexts.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2021
D. Russo et al. (Eds.): IHSI 2021, AISC 1322, pp. 218–224, 2021.
https://doi.org/10.1007/978-3-030-68017-6_33

Specifically, with the growing understanding of the complex interplay between emotions and cognition, there is a need to develop life-like, convincing agents that not only provide effective expert guidance, but also believable emotional interactions with the learner. Although a few researchers have tried to equip intelligent agent systems with affective capabilities [2, 8, 9], much work still needs to be done in this area. In short, there is a need to explore the benefits of adding emotional intelligence to online agents, that is the ability to express emotions and to recognize and adapt to the emotional state of others [5]. The work reported in the paper is a first step to filling this need.

Further, the animation quality of existing APA is low. Existing agents either fall in the uncanny valley [10], or are too stylized to display clear and believable personalities and emotions. Learners may fail to connect with such agents and fail to engage in the learning activity. On the other hand, animated characters that captivate and engage by displaying convincing personality and emotions are now possible, as demonstrated by the enormous popularity of computer graphics applications in entertainment, such as movies and games. Nonetheless, these highly engaging characters have not yet had a substantial impact on the educational domain. The work reported in the paper aims to improve on the state-of-the-art in agent animation.

In particular, the paper describes the development of a new animation system (and underlying algorithms) for generating life-like, believable pedagogical agents that display clearly recognizable emotions and that will respond and adapt to the learner's detected emotional state (the emotion detection module of the system is still under development, results will be reported in a future publication). The paper also reports findings from two preliminary studies which examined students' perception of the affective agents generated with the system.

Once completed, the research work described in the paper will have important implications. First, it will contribute to our understanding of the effects of animated affective agents on people's learning and the interaction between agent's affective features, learner's characteristics and learning effects. Second, it will contribute to the development of new, effective algorithmic approaches to emotion representation in animated agents and to emotion recognition.

## 2 Affective Agents: State-of-the-Art

Pedagogical agents are animated characters embedded within a computer-based learning environment to facilitate student learning. Many studies confirm the positive learning effects of systems using these agents. A meta-analysis by Schroeder showed that lessons with APA led to statistically significant learning improvements compared to lessons without them [12].

Affective pedagogical agents are agents with emotional capabilities. They can support simulated social interactions between learner and computer. According to Kim, Baylor and Shen [6], it is "*the provision of such simulated social interactions that may distinguish pedagogical [affective] agents from traditional computer-based tutoring, seemingly offering a unique instructional impact*". A few affective agent systems have been developed to date. The system by Lisetti and Nasoz [8] combines facial expressions and physiological signals to recognize a limited set of user's emotions, like fear and anger. A multimodal anthropomorphic agent then adapts its interface by responding to the user's emotional states, and provides multi-modal feedback to the user. The IA3 system by Huang et al. [4] was an early attempt at developing intelligent affective agents that recognize human emotion, and based on their understanding of speech and emotional state, decide on how to respond.

D'Mello and Graesser [2] have shown that students display a variety of emotions during learning with the intelligent tutoring system, AutoTutor (such as boredom, flow/engagement, confusion, and frustration) and they have shown how AutoTutor can be designed to detect and respond to them. SimSensei [9] is a virtual agent that engages in interviews with the user in order to elicit behaviors that can be automatically measured and analyzed. SimSensei uses a multimodal sensing system that captures a variety of signals that are used to assess the user's affective state, as well as to inform the agent so she/he can provide appropriate feedback. Research indicates that the manipulation of the APAs' affective states can significantly influence learner beliefs and learning efficacy [13]. A study by Kim et al. [6] showed that an agent's empathetic responses to the student's emotional states had a positive influence on learner self-efficacy for the task. A meta-analytic review that examined findings from studies on the efficacy of affective APAs showed that the use of affect in APAs has a significant and moderate impact on students' motivation, knowledge retention and knowledge transfer [3].

The problem with existing systems is that the pedagogical agents display a limited set of emotions, which often appear robotic and unnatural. Equipping agents with a wide range of life-like emotions provides more believability in the agent's behavior, and might lead to increased student learning. However, such a result cannot be attained without a set of rules that determine which affective behavior the agent should exhibit. The goal of the project described in the paper is to improve on the state of the art by developing new algorithms that produce life-like agents that display a large variety of natural emotions, and integrating them with rules for emotional design and emotion regulation in the context of online learning.

## 3 The Affective APA System

The system uses the following as the building blocks of emotion animation: (1) whole body poses, (2) specific arm/hand gestures, (3) facial expressions, (4) speech, and (5) motion dynamics.

(1) Whole body poses include the 4 types of body postures commonly used in animation of acting: open, closed, forward, and back poses. We use a continuum of body poses ranging from open to closed to express Russell's valence dimension (from positive to negative), and a continuum of body poses ranging from back to forward to express Russell's arousal dimension (from deactivation to activation) [11].

(2) For specific arm/hand gestures we use McNeill's classification of communicative gestures (e.g. Iconics, Metaphorics, Deicticts, Beats, Emblems).

(3) Facial expressions: the algorithms make use of a face model comprised of 30 joints with 55 DOF which allows for representing 36 Action Units+tongue, teeth, chin, ears movements.
(4) Speech animation is generated using a standard set of 20 visemes (mouth shapes), 12 representing the consonants, 7 representing the monophthongs, and 1 viseme representing the neutral pose, i.e. silence.
(5) Motion parameters such as velocity, acceleration, amplitude, frequency are manipulated to better express the agent's emotional state and to enforce adherence to fundamental principles of animation.

The system takes as input (1) a rigged character, (2) a text script and recorded audio, which determine what the agent says and the specific gestures the agent makes in support of the speech, and (3) the agent emotional style and the detected learner emotion, which determine the types of body poses, facial expressions and movement dynamics.

During the interpretation phase the animation algorithms analyze the input elements and identify the animation components necessary to compute the agent's animation. During animation instantiation, the identified animation elements are transformed into animation data segments. In the compositing phase the algorithms compute realistic transitions between consecutive character poses, enforce adherence to 6 animation principles and perform movement modulation. The animation data is applied to the agent rigged model using animation retargeting and is rendered in real time.

A flowchart illustrating the system is included in Fig. 1.

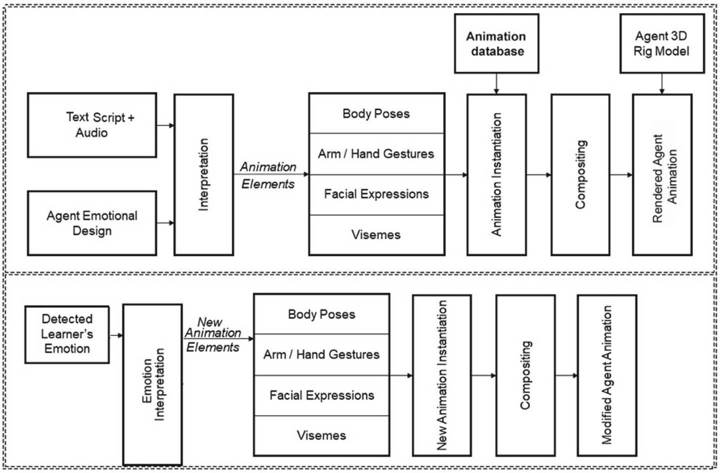


**Fig. 1.** Overview of APA system.

Videos of APA produced by the system can be accessed using the links below:
Happy: https://youtu.be/JjOjuYSJ8pk; Frustrated: https://youtu.be/oLYlFq6IRi8
Content: https://youtu.be/lEZy-LR0XGU; Bored: https://youtu.be/_8-Ooazd9bI

When a leaner's emotional state different from neutral is detected, the emotion is interpreted by the algorithms and new animation elements are selected, instantiated, composited and rendered in order to express the agent's response to the learner. As previously mentioned, the emotion recognition module of the system is still under development, a first version of the emotion detection module is described in [14].

## 4 Findings from Preliminary Studies

The objectives of the two studies conducted so far were to examine people perceptions of the affective agents generated with the system.

The first study [1] explored the effects of certain body gestures on viewer perception of the agent's emotion in terms of valence and arousal. More specifically, the study investigated the extent to which modifications to the range of movement of three beat gestures, e.g., both arms synchronous outward gesture, both arms synchronous forward gesture, and upper body lean, and the agent gender had significant effects on viewer perception of the agent's emotion along Russel's valence and arousal dimensions [11]. The study included two experiments and in each experiment the range of movement of the agent gestures was varied at two discrete levels. The stimuli for both experiments were sets of 12-s animation clips generated with the APA system. Each clip showed a different combination of body gestures and ranges of motion. In the first experiment subjects were asked to watch the stimuli clips and rank them from highest to lowest arousal and from positive to negative valence. In the second experiment subjects were divided into two groups: one group was assigned the clips featuring the female agent and one group the clips featuring the male agent. All subjects were asked to watch the stimuli clips and rate the valence and arousal using a 7-point Likert scale.

Findings showed that the more open and forward the gestures the agent makes, the more positive the perceived emotional valence and the higher the perceived emotional arousal. Findings also showed that female agents' emotions are perceived as having higher arousal and more positive valence that male agents' emotions. These results are important, as one main goal of our research is to create APAs that can convey clearly perceivable emotions not only through speech and facial expressions, but also through body gestures. Furthermore, there has been considerable debate as to whether posture and movement reliably convey emotions, or rather convey only the intensity of the emotion. Findings from the study demonstrate how body gestures are effective at expressing both the quality of the agent's emotion and its level of activation.

The second study [7] examined how well participants were able to perceive different emotions portrayed by a human instructor and by an animated pedagogical agent generated by our system in a video lecture on statistics. Participants were shown short video clips of either a human instructor or a pedagogical agent displaying four different emotions: happy, content, bored, and frustrated. The participants were asked to rate how well each video clip displayed each of those four emotions. Findings showed that subjects were able to recognize each of the emotions displayed by the instructor for both the human instructor and the animated agent. Subjects were much better at distinguishing between positive (happy and content) and negative (bored and frustrated) emotions than between active (happy and frustrated) and passive (content and bored) emotions. Emotions that involved higher activity (happy and frustrated) were more easily recognized in a human instructor than an animated agent. These results demonstrate that the emotions displayed by the agents generated with our system are recognizable, however they are not as clear as the ones conveyed by a real teacher. In future work, we will examine whether implementing exaggeration, one of the 12 principles of animation, will improve the clarity and hence the recognition of the emotions portrayed by our animated agents.

**Acknowledgments.** The work reported in the paper is supported in part by NSF – Cyberlearning Collaborative Research: Multimodal Affective Pedagogical Agents for Different Types of Learners, Award #1821894.

## References

1. Cheng, J., Zhou, W., Lei, X., Adamo, N., Benes, B.: The effects of body gestures and gender on viewer's perception of animated pedagogical agent's emotions. In: Kurosu, M. (ed.) Human-Computer Interaction. Multimodal and Natural Interaction. HCII 2020. Lecture Notes in Computer Science, vol. 12182. Springer, Cham (2020)
2. D'Mello, S., Graesser, A.: AutoTutor and affective AutoTutor: learning by talking with cognitively and emotionally intelligent computers that talk back. ACM Trans. Interact. Intell. Syst. **2**(4), Article 23 (2012)
3. Guo, Y.R., Goh, D.H.L.: Affect in embodied pedagogical agents: meta-analytic review. J. Educ. Comput. Res. **53**(1), 124–149 (2015)
4. Huang, T.S., Cohen, I., Hong, P., Li, Y.: IA3: intelligent affective animated agents. In: ICIG 2000 (2000)
5. Johnson, W.L., Lester, J.C.: Face-to-face interaction with pedagogical agents, twenty years later. Int. J. Artif. Intell. Educ. **26**(1), 25–36 (2016)
6. Kim, Y., Baylor, A.L., Shen, E: Pedagogical agents as learning companions: the impact of agent emotion and gender. J. Comput. Assist. Learn. **23**, 220–234 (2007)
7. Lawson, A., Mayer, R.E., Adamo-Villani, N., Benes, B., Cheng, J., Lei, X.: Recognizing the emotional state of human and virtual instructors. Comput. Hum. Behav. **114**, 106554 (2021)
8. Lisetti, C.L., Nasoz, F.: MAUI: a multimodal affective user interface. In: 10th ACM International Conference on Multimedia (Multimedia 2002), pp. 161–170 (2002)
9. Morency, L.P., Stratou, G., DeVault, D., Hartholt, A., Lhommet, M., Lucas, G., Morbini, F., Georgila, K., Scherer, S., Gratch, J., Marsella, S., Traum, D., Rizzo, A.: Demonstration: a perceptive virtual human interviewer for healthcare applications. In: 29th AAAI Conference on Artificial Intelligence (2015)
10. Mori, M.: The uncanny valley. IEEE Rob. Autom. **19**(2), 98–100 (2012). https://doi.org/10.1109/mra.2012.2192811
11. Russell, J.A.: Core affect and the psychological construct of emotion. Psychol. Rev. **110**, 145–172 (2003)
12. Schroeder, N., Adesope, O.O., Gilbert, R.B.: How effective are pedagogical agents for learning? A meta-analytic review. J. Educ. Comput. Res. **49**(1) (2013)

13. Zhou, L., Mohammed, A.S., Zhang, D.: Mobile personal information management agent: supporting natural language interface and application integration. Inf. Process. Manage. **48** (1), 23–31 (2012)
14. Zhou, W., Cheng, J., Lei, X., Benes, B., Adamo, N.: Deep learning-based emotion recognition from real-time videos. In: Kurosu, M. (ed.) Human-Computer Interaction. Multimodal and Natural Interaction. HCII 2020. Lecture Notes in Computer Science, vol. 12182. Springer, Cham (2020)