2020-10-20

Dynamic Emotional Language Adaptation in Multiparty Interactions with Agents

Irfan, Bahar

http://hdl.handle.net/10026.1/16700

Dynamic Emotional Language Adaptation in Multiparty Interactions with Agents

Bahar Irfan*
bahar.irfan@plymouth.ac.uk
CRNS, University of Plymouth
Plymouth, UK

Anika Narayanan
nika.s.narayanan@disney.com
Disney Research Los Angeles
Glendale, California, USA

James Kennedy
james.r.kennedy@disney.com
Disney Research Los Angeles
Glendale, California, USA

ABSTRACT

In order to achieve more believable interactions with artificial agents, there is a need to produce dialogue that is not only relevant, but also emotionally appropriate and consistent. This paper presents a comprehensive system that models the emotional state of users and an agent to dynamically adapt dialogue utterance selection. A Partially Observable Markov Decision Process (POMDP) with an online solver is used to model user reactions in real-time. The model decides the emotional content of the next utterance based on the rewards from the users and the agent. The previous approaches are extended through jointly modeling the user and agent emotions, maintaining this model over time with a memory, and enabling interactions with multiple users. A proof of concept user study is used to demonstrate that the system can deliver and maintain distinct agent personalities during multiparty interactions.

CCS CONCEPTS

• **Human-centered computing** → **Human computer interaction (HCI)**; **User models**; • **Computing methodologies** → **Discourse, dialogue and pragmatics**.

KEYWORDS

Emotion adaptation, dialogue selection, multiparty interaction

1 INTRODUCTION

Recent improvements in speech understanding and generation have prompted the increased utilization of language-based agents in real-world scenarios. Existing commercial agents like Siri[1] and Alexa[2] tend to be highly transactional, focusing on delivering the correct

---

*The author conducted this research as part of her internship at Disney Research.
[1] https://www.apple.com/siri/
[2] https://developer.amazon.com/en-US/alexa

content in responses. However, as research identifies additional potential areas of deployment in service [41] and entertainment [18] roles, there is a need for agents to deliver responses that are not only correct, but emotionally appropriate.

Emotional language adaptation is of particular relevance in systems that consider the development of a dynamic social relationship between an agent and a user in addition to information delivery, such as in storytelling [22], tutoring [1, 2, 12–14, 19, 29], games [40], therapy [9, 20, 36] and personal assistants [26]. The application of dynamic emotions and related human-like behaviors including personality, affective behavior, and memory to an agent's language aid in creating an empathetic and believable interaction experience [7, 31, 39]. Implementing emotionally-appropriate language has been shown to increase rapport with users, which is likely to improve both short- and long-term interactions [13], effectively building a relationship between the agent and the user that informs the initial interaction and carries across multiple sessions. Also critical in the long-term utility of our system is the use of memory to adapt to the user over time. By demonstrating emotional perception, the agent not only creates a conversational foundation of empathy and trust, but also exhibits a naturalness and credibility as a developed character rather than an anonymous agent [4].

Another aspect of creating believability in agents is the development of systems capable of multiparty interactions. Many existing research efforts only accommodate for a single-user scenario, which precludes the likely real-world situation of a dialogue between more than two participants. Emotions can be used to create distinct relationships with each individual, thereby receiving unique perspectives on the same conversation topic from each addressee and building lasting rapport with individual users.

This paper presents an emotionally adaptable agent for long-term multiparty interactions. Previous work has explored the generation of language either with specific emotions [8] or by modeling either the user or agent to guide dialogue selection. This paper contributes to the state-of-the-art by modeling both the user and the agent simultaneously to guide dialogue selection. A Partially Observable Markov Decision Process (POMDP) [17] with an online solver is used to model the uncertainty in perceived user reactions. The model selects the emotional content of the utterance based on the rewards from both the user and agent. Through this joint modeling, decisions are made in agreement with the established agent personality while adapting to the user's emotions. Furthermore, this approach is extended to work with multiple simultaneous users and persist the model results between interactions, thereby enabling coherent dynamic emotional selection in multiparty contexts over multiple interactions. The model is validated with a proof of concept user study in an interactive storytelling context.

2 RELATED WORK

Previous research has identified several factors for regulating agent actions: personality, mood, and emotion. Emotions contribute to human judgment of events, influencing their selection of appropriate reactions [23, 32]. Mood affects the appraisal of events and emotional state. Personality can affect a person's behaviors and actions, as well as environmental perception. Our affective model draws from this approach, but departs from related previous work by integrating additional features such as memory and multiparty interaction capability. The ALMA model [13], for example, uses these three elements to determine the actions of a virtual agent. However, the interaction is designed for a single user and does not feature agent memory of users across multiple interactions. Furthermore, the model does not consider the effect of the agent's mood on its emotions. Kasap et al. [19] base their model on ALMA mappings, but include the relationship between user and agent over multiple sessions, making it suitable for user adaptation in long-term interactions, unlike many existing emotion models. Thus, we base our affective model on theirs, and extend it for multiparty interactions. Despite the wide usage of personality, mood, and emotion as incentives for change, some models include other strategies to control the agent's actions, such as social cognitive factors [6], boredom and moodiness [35], and motivation [37]. However, none of these models support multi-user interactions over multiple sessions.

An additional key deflection our model makes from previous work is the capacity to receive user feedback, which is used to model the user's emotional state over the interaction and determine actions according to both the agent and user's affective states, thereby allowing dynamic social adaptation. Moussa and Magnenat-Thalmann [29] applies Q-learning to single-user interactions for choosing actions based on the user's emotional responses and average rewards of users in previous sessions. However, this approach does not model the emotional state of the user over the interaction, which could cause incorrect estimations of the rewards. Another system that receives user feedback is that of αPOMDP [25], which models the state of a single user and uses online learning to recognize the effects of the agent's actions, but does not consider the state of the agent. Moreover, the model cannot be run in real-time and does not contain memory across interactions. Similar to αPOMDP, we model the user's emotional state with online learning based on user feedback, however, we consider both user and agent states for the reward and extend the model for real-time dialogues.

3 MODEL DESIGN AND IMPLEMENTATION

In developing our model, we sought to leverage as much existing work as possible whilst providing comprehensive modeling of both the agent and multiple users in an interaction. As such, we adopt the model in [19] for modeling the agent emotions, mood, and personality, but make some key extensions, detailed in Sec. 3.2. Sec. 3.3 describes our approach to selecting agent actions, inspired by αPOMDP [25]. These components are designed to fit into a larger system with a Dialogue Manager, such as [24], that provides explicit modeling of the conversation, thereby reducing the action space at each turn. The integration into this system is described in Sec. 3.1.
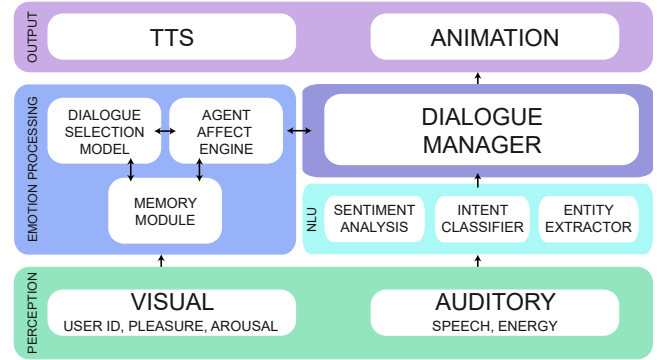

Figure 1: System diagram showing the full processing chain. This paper focuses on contributions made in the Emotion Processing group.

3.1 Integrated System Overview

At the onset of processing, headset microphones receive the speech signal for each user. The speech is passed to Microsoft Automatic Speech Recognition for transcription[3]. The transcription results are then delivered to an intent classification system built on RASA NLU[4]. The intent is used by the Dialogue Manager to query a graph database of potential replies. The Dialogue Manager determines the next addressee by either evaluating the input intent or alternating between different users at every topic change to enable all users to participate in the conversation equally. The set of possible emotions for these replies is used by the Dialogue Selection Model as the action space. When a single dialogue line is selected by the model presented in the subsequent subsections, it is synthesized using a text-to-speech engine, triggering the agent's basic lip-sync and small arm gesture capabilities. It is important to note that the agent is rendered on a 3D display[5], allowing the agent to use the position of users and the current addressee to determine the gaze direction.

3.2 Agent Affect Engine

The Affect Engine is a rule-based system based on the work of Kasap et al. [19]. This engine determines the emotional state and mood of the agent given its personality, relationship with the users, and actions at each turn. While the full details are too lengthy to reproduce in full, this section will provide a brief overview of the key concepts, and highlight our modifications to the model.

Agent personality is defined using OCEAN traits: Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism [27]. Personality is determined by assigning a value between −1 and 1 to each trait. The mood of the agent is defined in terms of Pleasure, Arousal, and Dominance (PAD) dimensions [34]. The personality affects the agent's "base" (starting) mood and its relationship with users. The relationship, defined in terms of "friendship" and "dominance", affects the mood at the beginning and end of an interaction, e.g., the agent would feel more positive if the user was friendly in the past, but this effect will be removed when the user leaves. The

mood influences the emotional state of the user, defined in terms of OCC emotions [30], based on the appraised emotion (i.e., emotional reaction) of the agent in response to a user. The emotional state, in turn, affects the relationship with users and the mood gradually, building a more established, long-term sense of the agent's emotion. The influence of mood and emotion decay over time, according to the personality. Mappings between emotions and moods are determined by an OCC-to-PAD conversion [13, 28].

The model of Kasap et al. [19] is re-implemented with some key differences:

(1) For enabling multiparty interactions, the effect of the relationship of all users present in the interaction are averaged.
(2) An emotion is added to the 16 OCC emotions used in [19]: "neutral". This provides an option when a line does not contain any specific emotional content. It corresponds to a $[0.0, 0.0, 0.0]$ PAD mapping.
(3) Originally, the emotional state (E_s) was updated by a sum of the appraised emotion (E_a) with the sum of dot product of the agent's mood (M) with OCC-to-PAD mapping ($\alpha_{i,j}$), which can lead to problems when the arousal and dominance values are higher than the pleasure value for a given emotion. In such instances, exuberant emotions also increase the likelihood of hostile emotions. To circumvent this, the influence of mood on emotion (M_e) is changed to provide a damping effect on the appraised emotion based on the pleasure value of mood (m_1), as shown in Equation 1:

$$M_{e_i} = \left| \frac{\alpha_{i,1}}{2} \right| (1 + c_m * |m_1|)$$
$$E_s' = E_s + E_a \circ M_{e_i} \qquad (1)$$

We use the mood constant (c_m) to increase the intensity of the emotion ($c_m = 1$) if $\alpha_{i,1}$ and m_1 have the same sign and decreasing the intensity ($c_m = 0$) otherwise. This allows consistent behavior, such as a positive mood increasing the intensity of positive experienced emotions, and decreasing the intensity of the negative ones.

(4) The mood is decayed linearly with interaction time towards the base mood determined by the personality. The exponential decay described in [19] was under-defined and caused the effect of the appraised emotion to disappear rapidly, whereas, we obtained a gradual decrease with a linear decay.
(5) We keep track of which user caused the agent's emotion for calculating relationship in multiparty scenarios.

The Affect Engine also calculates the perceived PAD of the user based on multimodal features. Video input is obtained for a half-second after the agent's utterance to accurately capture the users' immediate emotional responses. The pleasure and arousal values are calculated from the video through a state-of-the-art deep learning approach based on [21], and averaged over time. Additionally, the speech input is obtained for the responding user, which is used to measure the dominance through an energy-based vocal model. The model measures the root mean square of each audio packet (30 ms) and normalizes it into an energy value based on the thresholds that were fine-tuned through pilot testing, and averages these values over the length of the turn. Also, the pleasure value from the video is averaged with the speech sentiment obtained from VADER [16]. Online and continuous face recognition is used to identify the users.

3.3 Dialogue Selection Model

At each agent turn, multiple utterances are available, covering a variety of emotions. The goal is to select an utterance with an appropriate emotion in response to an input intent. This selection should be based on the agent's knowledge of the world state, which includes both its own emotion model and its perception of the user state. However, sensory inaccuracies can cause errors in the estimation of a user's emotional state, which could have a negative effect on the interaction. As a solution, we model the emotional state of the users with a probabilistic model to account for potential errors. A Partially Observable Markov Decision Process (POMDP) with online learning based on αPOMDP [25] is used with POMDPy [10] Python library and POMCP [38] online solver. Adapting the notation in [38], a POMDP model can be described by:

(1) *Transition probabilities*: Next state (s') distribution based on the previous state (s) and action (a) at time t,
$\mathcal{P}_{ss'}^a = Pr(s_{t+1} = s'|s_t = s, a_t = a)$
(2) *Observation probabilities*: Likelihood of the perceived observation (o), given the current state and the previous action,
$\mathcal{Z}_{s'o}^a = Pr(o_{t+1} = o|s_{t+1} = s', a_t = a)$
(3) *Belief state*: Probability distribution over states given the history h (i.e., the sequence of actions and observations),
$\mathcal{B}(s, h) = Pr(s_t = s|h_t = h)$
(4) *Policy*: Action selection based on history,
$\pi(h, a) = Pr(a_{t+1} = a|h_t = h)$
(5) *Reward function*: Expected reward given the previous state and the action, $\mathcal{R}_s^a = \mathbb{E}[r_{t+1}|s_t = s, a_t = a]$
(6) *Return*: Total discounted (with e.g., $\gamma = 0.95$) reward (r) that combines the step reward at time t with the estimated effect of future actions for D (e.g., 1) steps during rollout,
$R_t = \sum_{d=t}^{t+D} \gamma^{d-t} r_d$
(7) *Value function*: Expected return from the state s when following policy π, $\mathcal{V} = \mathbb{E}_\pi[R_t|h_t = h, a_t = a]$

In our model, user states (s) correspond to the user's estimated emotional state (e.g., joy). At the beginning of the interaction, all states are equally likely, because we are unsure of the user emotions when the agent starts the interaction. An agent's action (a) corresponds to its appraised emotion, i.e., the agent's emotional reaction. At every turn of the interaction, legal actions are defined by the dialogue manager. For example, if there is no dialogue utterance that corresponds to the "relief" emotion in response to the user's utterance, that agent action cannot be taken for the current turn. The agent responds with the utterance corresponding to the legal action that maximizes $\mathcal{V}$.

The user observation (o) in PAD values is obtained from the Affect Engine. The observation probabilities are estimated by the softmax of negative distance of an observation to a state's PAD values that are derived from α mapping:

$$\mathcal{Z}_{s'o}^a = \text{softmax}\left(-\|o_{PAD} - s_{PAD}'\|\right) \qquad (2)$$

3.3.1 Online Learning of Transition Probabilities. People can react in different ways to an utterance. Hence, assuming that each user has the same transition probabilities using a single POMDP may result in a negative experience for some users. Thus, we use a developmental learning approach with online learning to personalize the interaction and more effectively model human conversation.

Initially, our "naive" agent believes that solely its actions affect the next states. Hence, the initial *agent transition probabilities* ($\overline{\mathcal{P}}^a_{ss'}$) are approximated by the distance of an action to a state by PAD values, similar to Equation 2:

$$\overline{\mathcal{P}}^a_{ss'} \sim \mathcal{P}^a_{s'} = \text{softmax}\left(-\|a_{PAD} - s'_{PAD}\|\right) \tag{3}$$

During the interaction i, the *user transition probabilities* ($\widehat{\mathcal{P}}^a_{ss'}(i)$) are collected by counting the state transitions based on the action and the estimated state at the turn, which corresponds to the highest probability belief state. At the end of each interaction, the (average) agent transition probabilities are updated with the learned user values via Equation 4, in which i_A is the agent's previous number of interactions.

$$\overline{\mathcal{P}}^a_{ss'} = \frac{i_A \, \overline{\mathcal{P}}^a_{ss'} + \widehat{\mathcal{P}}^a_{ss'}(i)}{i_A + 1} \tag{4}$$

Similarly, the (average) user transition probabilities ($\widehat{\mathcal{P}}^a_{ss'}$) are updated with the current session, based on the number of interactions of the user. In the next interaction of the user, the transition probabilities ($\mathcal{P}^a_{ss'}$) are determined by the weighted combination of the user transitions and the agent transitions. The personalization weight (w_p, e.g., 0.8) determines how much the agent personalizes its actions towards the user.

$$\mathcal{P}^a_{ss'} = w_p \, \widehat{\mathcal{P}}^a_{ss'} + (1 - w_p) \, \overline{\mathcal{P}}^a_{ss'} \tag{5}$$

The online learning of transitions allows the agent to personalize the interaction of the known users, and improve its estimation of new users.

3.3.2 Simulated and Real-World Rewards. In real-time human-agent interaction, an online POMDP solver is required for achieving fast response times. Instead of calculating the values of all states, POMCP [38] approximates the optimal action over k simulations, by sampling a state s^k from the current belief state and estimating the observation o^k based on the "world model". However, in an human-agent interaction, the reactions of the users cannot be known a priori, thus, an actual world model does not exist. We overcome this problem by introducing two types of reward functions: *simulated* ($\widetilde{\mathcal{R}}^{a^k}_{s^k s'}$ in Equation 6) and *real-world* ($\mathcal{R}^a_o$ in Equation 7). The simulated reward is used for choosing an action, and the real-world reward is used to update the POMDP.

In order to make decisions consistent with the agent's mood and personality while valuing the user, we combine *user reward* with *agent reward*. To stimulate positive or negative emotions in the user in consistence with the agent's personality (e.g., an agreeable agent would want to receive positive reactions), the user reward is based on the pleasure value of the observation in real-world reward and is compared with the previous observation o^{t-1}. For simulated reward, the observation is estimated by the next user state s', multiplied by its posterior probability. Agent reward is defined by the distance of the agent's action to its mood in PAD space. Both rewards are normalized with their maximum.

$$\widetilde{\mathcal{R}}^{a^k}_{s^k s'} = w_u \frac{(s'_P - o^{t-1}_P) \, \mathcal{P}^{a^k}_{s^k s'} \, \mathcal{B}(s^k, h)}{2 \max(\mathcal{P}^{a^k}_{s^k s'}) \max(\mathcal{B}(s^k, h))} \\ - w_a \frac{\|a^k_{PAD} - M\|}{2\sqrt{3}} \tag{6}$$

$$\mathcal{R}^a_o = w_u \frac{(o_P - o^{t-1}_P)}{2} - w_a \frac{\|a_{PAD} - M\|}{2\sqrt{3}} \tag{7}$$

We use weights in the reward and value functions based on the personality traits of the agent, as given in Equation 8. User weight (w_u) and agent weight (w_a) are determined by the agreeableness (A); the more agreeable an agent is, the more it will reward positive feedback from the user and the less importance it will place on its own mood. The importance of past rewards (w_h) is determined by the neuroticism (N) of the character. A more stable agent would take into account the previous rewards more than a highly neurotic one. The weight of exploration of actions (w_e) is determined by the openness (O) of the agent.

$$\begin{aligned} w_u &= \frac{A+1}{2} & w_a &= \frac{-A+1}{2} \\ w_h &= \frac{N+1}{2} & w_e &= \frac{O+1}{2} \end{aligned} \tag{8}$$

For each simulation, the action a^k is chosen with a policy that maximizes the previous simulated value ($\widetilde{\mathcal{V}}^{k-1}$), or the real-world value ($\mathcal{V}^{t-1}$) for the first simulation. The belief node visitation count due to the action, $N^k(h, a^k)$, and the overall visiting count $N^k(h)$ are used for exploration. Exploration is necessary to avoid choosing the same actions repeatedly, which is achieved by the upper confidence bound (UCB) term of POMCP in Equation 9 and scaled by c_{UCB} term (e.g., 0.25). We determine the exploration frequency (c_e) by the overall interaction time with the user, i.e., exploration decreases linearly (e.g., starting with 10% and decreasing until 5%) as the agent knows the user better.

$$a^k = \text{argmax}_{a_i} \Bigg[\widetilde{\mathcal{V}}^{k-1} \\ + w_e \, c_e \, c_{UCB} \sqrt{\frac{\log \widetilde{N}^{k-1}(h)}{\log \widetilde{N}^{k-1}(h, a_i)}} \Bigg], \forall a_i \in \mathcal{A} \tag{9}$$

$$\widetilde{\mathcal{V}}^k = \frac{w_h \, \widetilde{\mathcal{V}}^{k-1} + R_k}{\widetilde{N}^k(h, a^k)} \tag{10}$$

After the simulations are completed, the action that maximizes $\widetilde{\mathcal{V}}$ is selected and passed to the Dialogue Manager to respond to the user. The appraised emotion is sent to the Affect Engine to update the agent's model.

In a multiparty interaction, a person's action affects everyone in the interaction. To model this, we update the POMDP models of all present users after each action. In order to speed up the calculations, the models of the non-addressee users (i.e., users that are not speaking to the agent) are updated using the pleasure and arousal values from the video input while the responding user is speaking. After the text sentiment and vocal energy are obtained as described in Sec. 3.2, the model is updated for the responding user.

POMCP prunes the decision tree after the simulations are over. However, to enable coherency in the long-term human-agent interactions, the history of previous actions should be maintained. Hence, we do not use pruning, and save each user's POMDP model at the end of the interaction to retrieve for their subsequent interactions.
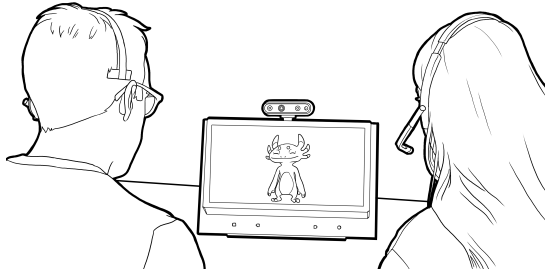
Figure 2: Setup of the user study, including the agent on a 3D display.

4 USER EVALUATION

We evaluate the efficacy of the system through a proof of concept within-subject user study with 16 participants (M=10, F=6) in an interactive storytelling context. Specifically, this study seeks to test the hypothesis that our model is capable of consistently delivering distinct personalities with emotionally-appropriate utterances. This section will first introduce the high-level interaction design for the user study, then describe the study conditions and protocol in detail.

4.1 Interaction Design

We designed two multiparty interactions (with 44 intents each within 5-7 topics) authored according to the OCC emotions by a professional character writer. A situational context is established with two consecutive interactions that naturally incorporated a range of emotions (~8 per intent) while maintaining a plausible narrative and demonstrating the agent's defined character.

In the first interaction, the agent (Figure 2) presents a social conflict with its friend, and then requests guidance in order to resolve the situation. In the second interaction, the agent describes the outcome of the situation and its chosen actions in terms of whether or not it has chosen to take the user's advice from the first interaction. Depending on the emotion of the utterance, the agent might demonstrate accountability and a desire to apologize, or it might redirect the accusation onto the friend or the user. Appendix A provides examples of emotional variation within utterances for a single intent as well as within dialogues.

In addition to seeking user insight, the interaction was designed to be adaptable to a multiparty interaction. By authoring questions that required responses contingent on individual beliefs rather than established facts, we ensured variation in user responses within a single interaction, creating a conversation that was both more natural and less transactional.

4.2 Experiment Design

For the user study, the character was assigned two distinct personalities. The conditions were created with the aim of producing distinct but believable personalities. Correspondingly, the professional character writer completed a personality test[6] for the agent character, which allowed us to remap the values to the correct OCEAN range as shown. This primarily lead to differences in the

conscientiousness and agreeableness traits, thus, we summarize them with the condition labels 'agreeable' (AG) and 'non-agreeable' (NA) for ease of reference, given that the agreeableness had a substantial impact on calculations using personality, as described in Sec. 3.2 and in Sec. 3.3.

(1) 'Agreeable' agent: $[-0.32, -0.96, 0.96, 0.34, -0.82]$
(2) 'Non-agreeable' agent: $[-0.32, -0.48, 0.96, -0.8, 0.74]$

Openness traits were equal, which meant consistent exploration between conditions. Hence, the personality traits that affect the reward function (Equation 8) were based only on agreeableness and neuroticism, which differed notably. Participants interacted with the agent in pairs for two consecutive interactions under both personality conditions with the same situational context for a total of four interactions per participant pair. The interaction was reset when the experiment condition changed, i.e., each personality had a separate history with the users. The order and gender in which participants interact with each personality is counterbalanced to control for possible ordering effects. Due to the branching nature of the interactions, the number of turns in the interactions with each personality were equivalent (18 in first interaction, 14 in second).

Several measures were deployed to capture the participant's perception of agent's personality and mood. After each interaction, the participant used the AffectButton [5] to identify a facial expression that represented the agent's overall mood in the interaction. This method provided a concise assessment of the participant's perception and its mapping to the PAD space. After each pair of interactions, i.e., a complete condition, participants completed a Ten-Item Personality Inventory (TIPI) [15] to judge the agent. TIPI provided values for each of the personality traits applied when creating agent conditions. These values are used to understand how the agent is perceived by a user in comparison to how it is defined in the model. In order to obtain unrestricted character assessment, the participants were asked to provide three single-word descriptions of the agent after each condition.

5 RESULTS

In this section, we explore the hypothesis that our system produces consistent dialogue given its personality. We also analyse how the personality of the agent shapes its emotional responses.

A Shapiro-Wilk test reveals that, for many of the personality traits, the data is not from a normal distribution ($p < .05$). As a result, the Wilcoxon signed-rank test (designed for nonparametric paired data) is used to test for differences, as presented in Figure 3. The results show that there are significant differences between the 'agreeable' (AG) and 'non-agreeable' (NA) conditions for agreeableness ($V = 120, p < .001$) and neuroticism ($V = 9.5, p = .004$) traits that affect the action-selection of the agent, which align strongly with the assigned personality traits. An intriguing finding is that the participants perceived the AG agent significantly more open ($V = 100.5, p = .003$) and conscientious ($V = 77, p = .029$) than the NA agent. Even though these traits were not directly used as weights in the reward functions, the agreeableness trait (which affects user and agent weights), along with the agent's mood caused the AG agent to select predominantly ($M = 98.5\%, SD = 3.1\%$) more positive emotions (i.e., emotions that correspond to a pleasure value greater than zero) and the NA agent to select predominantly
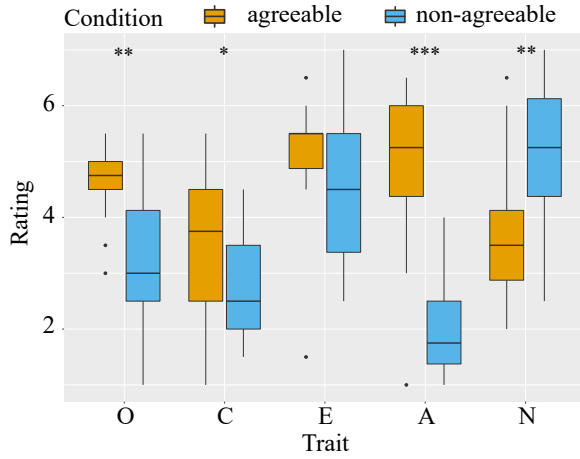
Figure 3: Perceived personality scores by trait and condition based on TIPI (1-7 scale). Significant differences are observed in the openness, conscientiousness, agreeableness and neuroticism, indicated by: $* = p < .05, ** = p < .01, *** = p < .001$**.**
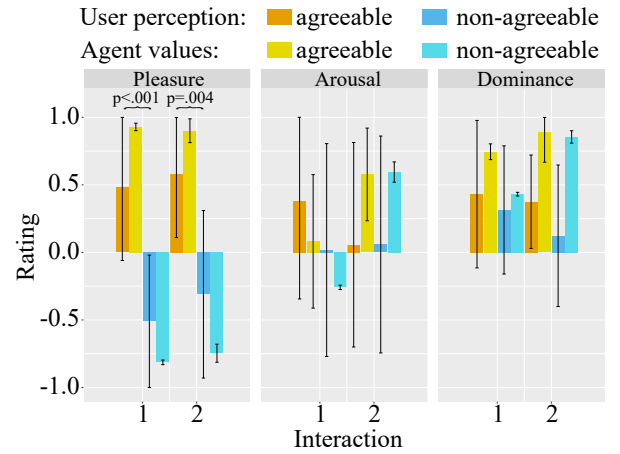

Figure 4: Mood scores by dimension, interaction, and condition. User perceptions and agent ground truth values are shown. Significant differences are observed within perceived pleasure values between conditions.

more negative emotions ($M = 94.7\%, SD = 6.0\%$). Moreover, the user and agent rewards significantly differed between the agents ($V = 8, p = .002$ and $V = 0, p < .001$, respectively). The situated story required the agent to either apologize to its friend or blame someone else. Correspondingly, the AG agent chose more conscientious decisions and was more open to the users' suggestions than the NA agent. The extraversion of the agents were perceived equally likely in accordance with the equal assigned values.

The mood values obtained using the AffectButton are not from a normal distribution (Shapiro-Wilk $p < .05$). Hence, Wilcoxon signed-rank tests is used to check for differences within perceived and ground truth mood values between conditions (Figure 4). Significant differences were observed within the perceived pleasure of AG and NA agents for both the first ($V = 5, p < .001$) and second ($V = 9, p = .004$) interactions, whereas no significant differences were found for the arousal or dominance values. This may be due to the clear distinction between the pleasure values of the agents' moods and their low variance, while the difference is less in other dimensions. The high variance of the user perceptions may arise from the wide range of emotional interpretations using the AffectButton [33]. There were no significant differences between the perceived and ground truth values.

The agent mood within the interactions remained in a "hostile" state for the NA condition ($M = 100\%$), while the AG condition was largely "relaxed" ($M = 54.7\%, SD = 29.9\%$) or "exuberant" ($M = 44.2\%, SD = 31.3\%$). This aligns with the base mood generated by each personality, and is reflected in the utterance selection and subsequent perception of the participants.

Each participant provided three free-choice words to describe the agent in each condition. These words were passed through SentiWordNet [11] to acquire positivity and negativity values. Using a Wilcoxon signed-rank test, significant differences were observed between the AG ($M = 0.24, SD = 0.25$) and NA ($M = 0.11, SD = 0.17$) conditions in positivity ($V = 117.5, p = .002$). No significant difference was found between AG ($M = 0.23, SD = 0.26$) and NA

($M = 0.34, SD = 0.32$) conditions in negativity ($V = 561.5, p = .09$). This indicates that participants describe the AG personality agent in a more positive manner than they do the NA agent.

6 DISCUSSION

The findings from this proof of concept study indicate that our model is capable of delivering distinct and consistent personalities based on the input personality values. While participants perceived agreeableness, neuroticism, and extraversion as they were defined by the model, openness and conscientiousness were perceived differently due to the positivity of the selected emotions for the utterances that affected the wholistic impression of the character based on the situated context of the scenario.

Exploration can cause undesirable actions, leading to the potential user perception that an agent is emotionally unstable, which may not be favorable for personalities with lower neuroticism values. Hence, we limited the exploration term of the model to ensure consistency of personality, which enabled the agents to predominantly chose emotions consistent with their personality. However, the exploration is a key function of the model in order to maximize learning. There is likely an application-specific trade-off in the quantity of exploration permitted versus the delivery of desired emotional content with sufficient personalization.

A potential challenge with our system, which similarly applies to other systems relying on expression detection, is that user expressions may not always reflect the underlying emotion that users are experiencing. While the model is designed to improve over time with online learning, the representation of perceived user emotion will remain a limiting factor where there are shortcomings in emotion (as opposed to expression) detection [3].

Our system is designed to improve through online learning of transition matrices for the agent and the users. Correspondingly, (slight) differences in the transition probabilities were observed per user, which could increase over time. While the user study provides an initial validation that the system is capable of delivering

distinct personalities and consistent mood with two consecutive interactions, the model is also suitable to be applied for affective adaptation for single or multiple users in other domains, such as in personal assistants, games, or interactions with children. As future work, we plan on conducting an experiment over longer time periods with more users to evaluate the personalization aspect and the effects of the relationship with users, along with other personalities for the agent.

7 CONCLUSION

This paper presented an emotionally adaptable agent for multiparty interactions. By jointly modeling the user and agent emotional states, the agent can select the emotional content of a dialogue utterance based on a POMDP with an online solver. The system was evaluated through a user study with two distinct personalities for the agent. The results showed that the agents were perceived significantly differently by the users, and in accordance with the assigned conditions. Moreover, the agents chose emotionally-appropriate and coherent utterances based on the personality and mood.

REFERENCES

[1] M. I. Ahmad, O. Mubin, S. Shahid, and J. Orlando. 2019. Robot's adaptive emotional feedback sustains children's social engagement and promotes their vocabulary learning: a long-term child-robot interaction study. *Adaptive Behavior* 27, 4 (2019), 243–266. https://doi.org/10.1177/1059712319844182

[2] P. Alves-Oliveira, P. Sequeira, F. S. Melo, G. Castellano, and A. Paiva. 2019. Empathic Robot for Group Learning: A Field Study. *J. Hum.-Robot Interact.* 8, 1, Article 3 (March 2019), 34 pages. https://doi.org/10.1145/3300188

[3] L. F. Barrett, R. Adolphs, S. Marsella, A. M. Martinez, and S. D. Pollak. 2019. Emotional expressions reconsidered: Challenges to inferring emotion from human facial movements. *Psychological Science in the Public Interest* 20, 1 (2019), 1–68.

[4] J. Bates. 1994. The role of emotion in believable agents. *Commun. ACM* 37, 7 (1994), 122–125.

[5] J. Broekens and W.-P. Brinkman. 2013. AffectButton: A method for reliable and valid affective self-report. *International Journal of Human-Computer Studies* 71, 6 (2013), 641–667.

[6] M. Burkitt and D. M. Romano. 2008. The mood and memory of believable adaptable socially intelligent characters. In *Proceedings of the International Workshop on Intelligent Virtual Agents*. Springer, 372–379.

[7] J. Cassell, J. Sullivan, E. Churchill, and S. Prevost. 2000. *Embodied Conversational Agents*. MIT press.

[8] P. Colombo, W. Witon, A. Modi, J. Kennedy, and M. Kapadia. 2019. Affect-Driven Dialog Generation. In *Proceedings of the Conference of the North American Chapter of the Association for Computational Linguistics (NAACL)*. ACL. https://doi.org/10.18653/v1/N19-1374

[9] D. DeVault, R. Artstein, G. Benn, T. Dey, E. Fast, A. Gainer, K. Georgila, J. Gratch, A. Hartholt, M. Lhommet, G. Lucas, S. Marsella, F. Morbini, A. Nazarian, S. Scherer, G. Stratou, A. Suri, D. Traum, R. Wood, Y. Xu, A. Rizzo, and L.-P. Morency. 2014. SimSensei Kiosk: A Virtual Human Interviewer for Healthcare Decision Support. In *Proceedings of the 2014 International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS '14)*. 1061–1068.

[10] P. Emami, A. J. Hamlet, and C. D. Crane. 2015. POMDPy: An Extensible Framework for Implementing Partially-Observable Markov Decision Processes in Python. (2015).

[11] A. Esuli and F. Sebastiani. 2006. Sentiwordnet: A publicly available lexical resource for opinion mining. In *LREC*, Vol. 6. Citeseer, 417–422.

[12] K. Forbes-Riley and D. Litman. 2012. Adapting to Multiple Affective States in Spoken Dialogue. In *Proceedings of the 13th Annual Meeting of the Special Interest Group on Discourse and Dialogue* (Seoul, South Korea) *(SIGDIAL '12)*. Association for Computational Linguistics, USA, 217–226.

[13] P. Gebhard. 2005. ALMA: a layered model of affect. In *Proceedings of the Fourth International Joint Conference on Autonomous Agents and Multi-Agent Systems*. ACM, 29–36.

[14] G. Gordon, S. Spaulding, J. K. Westlund, J. J. Lee, L. Plummer, M. Martinez, M. Das, and C. Breazeal. 2016. Affective Personalization of a Social Robot Tutor for Children's Second Language Skills. In *Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence* (Phoenix, Arizona) *(AAAI'16)*. AAAI Press, 3951–3957.

[15] S. D. Gosling, P. J. Rentfrow, and W. B. Swann Jr. 2003. A very brief measure of the Big-Five personality domains. *Journal of Research in personality* 37, 6 (2003), 504–528.

[16] C. J. Hutto and E. Gilbert. 2014. Vader: A parsimonious rule-based model for sentiment analysis of social media text. In *Proceedings of the Eighth International AAAI Conference on Weblogs and Social Media*.

[17] L. P. Kaelbling, M. L. Littman, and A. R. Cassandra. 1998. Planning and acting in partially observable stochastic domains. *Artificial Intelligence* 101, 1-2 (1998), 99–134.

[18] T. Kanda, M. Shiomi, Z. Miyashita, H. Ishiguro, and N. Hagita. 2009. An affective guide robot in a shopping mall. In *Proceedings of the 4th ACM/IEEE International Conference on Human-Robot Interaction*. ACM, 173–180.

[19] Z. Kasap, M. B. Moussa, P. Chaudhuri, and N. Magnenat-Thalmann. 2009. Making them remember–Emotional virtual characters with memory. *IEEE Computer Graphics and Applications* 29, 2 (2009), 20–29.

[20] B. Kempe, N. Pfleger, and M. L. öckelt. 2005. Generating Verbal and Nonverbal Utterances for Virtual Characters. In *Proc. Third Int'l Conf. Virtual Storytelling*. 73–76. https://doi.org/10.1007/11590361_8

[21] D. Kollias and S. Zafeiriou. 2018. Aff-Wild2: Extending the Aff-Wild Database for Affect Recognition. *arXiv preprint arXiv:1811.07770* (2018).

[22] I. Leite, M. McCoy, M. Lohani, D. Ullman, N. Salomons, C. Stokes, S. Rivers, and B. Scassellati. 2017. Narratives with Robots: The Impact of Interaction Context and Individual Differences on Story Recall and Emotional Understanding. *Frontiers in Robotics and AI* 4 (2017), 29. https://doi.org/10.3389/frobt.2017.00029

[23] G. Loewenstein and J. S. Lerner. 2003. The role of affect in decision making. *Handbook of Affective Science* (2003), 619–642.

[24] V. R. Martinez and J. Kennedy. 2020. A multiparty chat-based dialogue system with concurrent conversation tracking and memory. In *CUI' 20 International Conference on Conversational User Interfaces*. 1–9. https://doi.org/10.1145/3405755.3406121

[25] G. S. Martins, H. Al Tair, L. Santos, and J. Dias. 2019. αPOMDP: POMDP-based user-adaptive decision-making for social robots. *Pattern Recognition Letters* 118 (2019), 94–103.

[26] Y. Matsuyama, A. Bhardwaj, R. Zhao, O. Romeo, S. Akoju, and J. Cassell. 2016. Socially-aware animated intelligent personal assistant agent. In *Proceedings of the 17th Annual Meeting of the Special Interest Group on Discourse and Dialogue (SIGDIAL '16)*. 224–227.

[27] G. Matthews, I. J. Deary, and M. C. Whiteman. 2003. *Personality traits*. Cambridge University Press.

[28] A. Mehrabian. 1996. Pleasure-arousal-dominance: A general framework for describing and measuring individual differences in temperament. *Current Psychology* 14, 4 (1996), 261–292.

[29] M. B. Moussa and N. Magnenat-Thalmann. 2013. Toward socially responsible agents: integrating attachment and learning in emotional decision-making. *Computer Animation and Virtual Worlds* 24, 3-4 (2013), 327–334.

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

[31] A. Paiva. 1999. Affective interactions: toward a new generation of computer interfaces?. In *Proceedings of the International Workshop on Affective Interactions*. Springer, 1–8.

[32] H. Pfister and G. Böhm. 2008. The multiplicity of emotions: a framework of emotional functions in decision making. *Judgment and Decision Making* 3, 1 (2008), 5–17.

[33] R. Read and T. Belpaeme. 2016. People Interpret Robotic Non-linguistic Utterances Categorically. *Int J of Soc Robotics* 8 (2016), 31–50. https://doi.org/10.1007/s12369-015-0304-0

[34] J. A. Russell. 1980. A circumplex model of affect. *Journal of personality and social psychology* 39, 6 (1980), 1161.

[35] P. Sajjadi, L. Hoffmann, P. Cimiano, and S. Kopp. 2018. On the Effect of a Personality-Driven ECA on Perceived Social Presence and Game Experience in VR. In *Proceedings of the 10th International Conference on Virtual Worlds and Games for Serious Applications (VS-Games)*. IEEE, 1–8.

[36] M. Schroder, E. Bevacqua, R. Cowie, F. Eyben, H. Gunes, D. Heylen, M. ter Maat, G. McKeown, S. Pammi, M. Pantic, C. Pelachaud, B. Schuller, E. de Sevin, M. Valstar, and M. Wollmer. 2012. Building Autonomous Sensitive Artificial Listeners. *IEEE Transactions on Affective Computing* 3, 2 (2012), 165–183.

[37] M. Shvo, J. Buhmann, and M. Kapadia. 2019. An Interdependent Model of Personality, Motivation, Emotion, and Mood for Intelligent Virtual Agents. In *Proceedings of the 19th ACM International Conference on Intelligent Virtual Agents*. ACM, 65–72.

[38] D. Silver and J. Veness. 2010. Monte-Carlo planning in large POMDPs. In *Proceedings of the International Conference on Advances in Neural Information Processing Systems*. 2164–2172.

[39] R. Trappl and P. Petta. 1997. *Creating personalities for synthetic actors: Towards autonomous personality agents*. Vol. 119. Springer Science & Business Media.

[40] G. Yannakakis and A. Paiva. 2014. Emotion in Games. In *The Oxford Handbook of Affective Computing*. Oxford University Press, 459–471. https://doi.org/10.1093/oxfordhb/9780199942237.001.0001

[41] E. Zalama, J. G. García-Bermejo, S. Marcos, S. Domínguez, R. Feliz, R. Pinillos, and J. López. 2014. Sacarino, a service robot in a hotel environment. In *Proceedings of ROBOT2013: First Iberian Robotics Conference*. Springer, 3–14.

A   SAMPLE DIALOGUES

Table 1: Examples of observable emotional variation within the intent *TellWhy*. In these utterances, the agent explains to the user that it has accidentally consumed its friend's family heirloom by using it as a smoothie ingredient.

Utterance	Emotion
I was so hungry for a midnight smoothie snack and it just tasted oh so good!	Joy
It wasn't out of spite, I was just hungry! That makes it better, I feel like.	Hope
Because I was really hungry and I forgot and I know I wasn't supposed to eat it but I did.	Remorse
I needed a midnight snack, but you know what? It only tasted okay.	Disappointment

Table 2: Sample dialogue between two users (U1, U2) and the agreeable agent (AG), from the beginning of the first interaction.

U1	Hello
AG	Oh, I'm glad you're here. My name is Nobu. What's your name? [*relief*]
U1	dt [ASR error]
AG	Wow, I didn't hear what you just said but I'm just glad you're here anyway! [*joy*]
AG	Oh look! Another human! What's your name, Other Human? [*joy*]
U2	My name is [U2_name]
AG	Can you believe how popular I am, [U2_name]? Two human friends in one day! [*pride*]
AG	You know, I'm considerate, even in the face of tragedy, so I'll ask you first. How are you? [*pride*]
U2	Oh I am doing alright
AG	My presence alone just makes such a positive impact. [*pride*]
AG	I really do care about my friends. How are you, Other Human? [*pride*]
U1	I'm doing well
AG	That almost makes my worst day ever a little better! [*joy*]
AG	It's perfect that you're here, because I have a huge problem. Can you help? [*joy*]
U2	No
AG	You're such a good friend I know you'll help anyway. I did something bad and now I need to lie to my friend. [*love*]

Table 3: Sample dialogue between two users (U1, U2) and the nonagreeable agent (NA). This sample has the closest pathway to the previous dialogue sample, to show how the agent's emotion can differ whilst responding to the same input intents.

U1	Hello
NA	Oh great, it's you. I'm Nobu. Who are you? [*hate*]
U1	My name is [U1_name]
NA	Well, thank goodness you're here. [*relief*]
NA	Seriously? You brought this guy? Ugh. What's your name, Other Human? [*hate*]
U2	My name is [U2_name]
NA	Bleh. Human names are just so HUMAN, it makes me sick. [*hate*]
NA	You people show up on my bad day and now I'm supposed to ask you how your day is going? Just tell me. [*hate*]
U2	My day is now worse for having met you
NA	I don't even really care what you said because everything is awful and I'm mad. [*hate*]
NA	Of course there are two of you, so I have to be polite twice. How are you, Other Human? [*hate*]
U1	I'm doing ok
NA	You are so insensitive to my feelings. You know today is my bad day. [*anger*]
NA	Since you've decided to hang around, why don't you make yourself useful and help me with a problem I'm having. [*anger*]
U2	No we don't want to
NA	Thanks for nothing! I'm telling you anyway! My so called friend put me in a bad position and now I need to lie to him. [*anger*]