

How to shape the humor of a robot - social behavior adaptation based on reinforcement learning

Klaus Weber, Hannes Ritschel, Ilhan Aslan, Florian Lingenfelser, Elisabeth André

Angaben zur Veröffentlichung / Publication details:

Weber, Klaus, Hannes Ritschel, Ilhan Aslan, Florian Lingenfelser, and Elisabeth André. 2018. "How to shape the humor of a robot - social behavior adaptation based on reinforcement learning." In *Proceedings of the 2018 on International Conference on Multimodal Interaction - ICMI '18, Boulder, CO, USA — October 16 - 20, 2018*, edited by Sidney K. D'Mello, Panayiotis (Panos) Georgiou, Stefan Scherer, Emily Mower Provost, Mohammad Soleymani, and Marcelo Worsley, 154–62. New York, NY: ACM Press. <https://doi.org/10.1145/3242969.3242976>.



How to Shape the Humor of a Robot - Social Behavior Adaptation Based on Reinforcement Learning

Klaus Weber, Hannes Ritschel, Ilhan Aslan, Florian Lingensfelser, Elisabeth André

Human Centered Multimedia

Augsburg University

Augsburg, Germany

lastname@hcm-lab.de

ABSTRACT

A shared sense of humor can result in positive feelings associated with amusement, laughter, and moments of bonding. If robotic companions could acquire their human counterparts' sense of humor in an unobtrusive manner, they could improve their skills of engagement. In order to explore this assumption, we have developed a dynamic user modeling approach based on Reinforcement Learning, which allows a robot to analyze a person's reaction while it tells jokes and continuously adapts its sense of humor. We evaluated our approach in a test scenario with a Reeti robot acting as an entertainer and telling different types of jokes. The exemplary adaptation process is accomplished only by using the audience's vocal laughs and visual smiles, but no other form of explicit feedback. We report on results of a user study with 24 participants, comparing our approach to a baseline condition (with a non-learning version of the robot) and conclude by providing limitations and implications of our approach in detail.

CCS CONCEPTS

• Computing methodologies → Learning from implicit feedback;

KEYWORDS

socially-aware agents, social adaptation, human-robot-interaction

ACM Reference Format:

Klaus Weber, Hannes Ritschel, Ilhan Aslan, Florian Lingensfelser, Elisabeth André. 2018. How to Shape the Humor of a Robot -, Social Behavior Adaptation Based on Reinforcement Learning. In *2018 International Conference on Multimodal Interaction (ICMI '18)*, October 16–20, 2018, Boulder, CO, USA. ACM, New York, NY, USA, 9 pages. <https://doi.org/10.1145/3242969.3242976>

1 INTRODUCTION

Embodied agents, such as robots who are capable of “mimicking” social interaction, are envisioned to enrich many social situations. In addition to addressing serious societal challenges, such as improving the quality of elderly care, social robots have the potential to satisfy people's hedonic needs of enjoying everyday life. To

demonstrate that robots can make an audience laugh, fellow researchers went as far as sending robots into theaters and stand-up comedy shows (e.g., [9, 11]). Laughing together with a robot may not only be fun and possibly healthy, but it may also strengthen a user's relationship with the social robot and turn interactions more enjoyable over a longer period of time.

Sjöbergh and Araki [21] found that jokes told by a robot are rated significantly funnier than jokes presented by text only. Thus, not only the content of a joke is important but also how it is delivered. In fact, several factors need to be addressed, including “intonation, posture, gaze, gesture, expression, and timing”, but also the communication between speaker and audience. [9]. Receiving feedback, such as smiles or raised eyebrows, can be very useful for the speaker, and thus, it seems that robots who aim to become humorous also need to learn how to elicit social signals, such as human laughter.

Developing humorously acting robots has been a research interest for Japanese Manzai [8] and as aforementioned in the standup comedy domain [9, 11] with a focus on entertaining larger audiences. Thus, research has investigated how to adapt spoken contents and presentation in a manner that optimizes a show for a larger audience, based on the audiences' explicit and implicit feedback. However, adaptation is also important when a robot has to tell jokes to a small group of listeners or even a single listener, as it would be the case for a domestic companion with joke-telling abilities.

As an aggravating factor, there are individual and possibly contextual differences in joke preferences: people may enjoy different types, such as gross-out, slapstick, or academic jokes. In order to entertain a human partner, a social robot would need to quickly acquire their human counterparts' sense of humor and preferences in non-verbal stimuli (e.g., mimics) to enrich the presentation of the joke. Thus, human feedback from the user is essential to learn these preferences. One way to gather user feedback is by directly asking the user, but as it requires additional interactions for both the robot and the user the flow of the conversation may suffer. Moreover, people are critical towards providing explicit feedback and serving as an “oracle” for a machine [1].

It seems unnatural to ask whether a joke was perceived as funny or not because during interpersonal communication, humans can easily interpret their conversation partner's reactions, such as if they react with a smile or a laugh. Laughter is used in different contexts within social interaction and may be the result of “diverse meanings and connotations” [3]. When making jokes, acoustic laughs and facial smiles are obvious, contemporary human reactions serving as an indication whether a joke is good or bad from the perspective of the human listener. Consequently, joke-telling robots should also be able to make use of their listeners' reactions and

(ideally) learn preferences implicitly by interpreting (non-verbal) social signals.

In order to explore if a robot can effectively learn their listener’s humor preferences solely by utilizing implicit feedback (i.e., facial smiles and vocal laughs), we have implemented a machine learning approach based on trial and error (i.e., Reinforcement Learning) and evaluated with users the approach’s performance.

Before we present our approach in detail, we outline in the next section research covering humor in human-robot interaction in combination with adaptation, as well as Reinforcement Learning based on social signals. An overview of the scenario, details of the adaptation approach and the process of sensing social signals are explained in Section 3. Finally, Section 4 presents the results of a study with 24 participants, which we conducted to evaluate the performance of our humor adaptation approach. The results of the user study demonstrate that learning preferences by human laughs and smiles help learn humor preferences in a joke telling scenario, and will significantly improve a robot’s joke telling performance. Furthermore, results also indicate the importance of “first impressions” and that a joke-telling robot needs to quickly acquire their listener’s sense of humor since it seems that users won’t change their opinion (easily) once it is established.

2 RELATED WORK

Related research can be separated in two streams; i.e., (i) research exploring scenarios in which robots aim to (help) entertain users and (ii) research on the application of Reinforcement Learning for adaptation purposes (e.g., to personalize robots based on trial and error). We summarize both streams, highlighting technical details in terms of what kinds of interactions and signals are (typically) measured and applied to adapt a robot’s behavior based on feedback.

2.1 Robots entertaining an audience

Japanese “Manzai” is a comedy dialog conversation, which is traditionally performed by two people. Hayashi et al. [8] have realized Japanese “Manzai” with two robots. Their system utilizes two Robovie robots capable of coordinating their communication with each other and reacting to external stimuli from the audience, including social signals in terms of laughter. Moreover, the audience’s clapping and laughing are estimated using a sound level meter. Sensor data are discretized in three levels to distinguish “burst out”, “laugh” and “cool down” reactions to adjust speech and motion timing of the robots’ performance depending on the audience.

Research by Knight et al. [11] uses a NAO robot which performs standup comedy while adapting the contents and performance to the audience. Both auditive and visual information are used to estimate the audience’s current enjoyment level, including noise caused by laughter or applause, as well as green or red cards held up by the spectators. Furthermore, the NAO robot gets explicit information on the audience’s valence based on direct prompts (i.e., likes or dislikes). In combination with several attributes associated with each joke and online learning (convex programming), the robot is able to select the best next joke to maximize audience enjoyment.

Katevas et al. [9] use the RoboThespian™ platform in the standup comedy domain with focus on joke delivery in terms of non-verbal

behavior (gaze and gestures). The robot reacts in real-time to audience responses by looking at peoples’ faces or responding to them, based on social signals sensed via audio (laughter/applause via directional microphone) and video (infrared HD camera, facial expression with SHORE™). Every comedy text includes a list of positive or negative responses, which can be used by the robot when laughter has been observed or not.

Humor research states that the presence of another person positively influences the cognitive and emotional response to humorous stimuli and inspired by this theory, Niewiadomski et al. [18] present a laughter-aware virtual agent that is meant to enjoy amusing videos together with the user. They present experiments, in which human social signals are sensed by a Microsoft Kinect sensor to detect the user’s voiced laughs. Based on the occurrence of laughter, an automated intensity estimation and context information from the given scenario, the agent adapts to the user’s amusement level by selecting an appropriate audiovisual response.

While most of the aforementioned research in the context of humor and robots addresses scenarios with a larger audience and the need to adapt in a manner that fulfills the preferences of multiple humans, our work focuses on an adaptation process for a single person. Moreover, the goal of the adaptation process, which we propose in this paper, is not only to adapt but more specifically to learn preferences depending on the user’s current emotional state in terms of amusement. This is where Reinforcement Learning comes into play, which is a machine learning approach based on trial and error and enables to learn appropriate behavior depending on the human’s current social signals.

2.2 Reinforcement Learning Applications

In Human-Robot Interaction, the combination of Reinforcement Learning with rewards based on social signals is used in a variety of scenarios. For example, Leite et al. [12] use smile, gaze and task-related information to learn empathic supportive strategies for playing chess. Facial expression is also used by Gordon et al. [7] in a student tutoring scenario. In the assistive domain, Liu et al. [13] use physiological signals. To learn social behavior interaction and feedback modalities, such as tactile [2] or prosodic [10] as well as interaction distance, gaze meeting, motion speed and timing [17] are used. Others have used engagement based on gesture and posture a storytelling scenario [19].

The robot humor learning scenario, which we address in this paper, also requires real-time adaptation based on human social signals. However, table-based algorithms for discrete state space approaches, such as Q-Learning or SARSA, suffer from the “curse of dimensionality”, which means that the memory and time required for learning increase exponentially when the amount of states increases.

One potential way to address this issue is to discretize social signals and include them in the state space; that is, for example, to differentiate a finite set of levels of valence [7] or engagement [7, 19], which then allows using social signals with table-based algorithms. However, it causes an information loss, as these algorithms cannot generalize knowledge across similar states. A second potential way is to use algorithms based on function approximation, which do not use a discrete set of states but parameters representing a

target function to learn and represent the state space. This second approach was previously used in Human-Robot Interaction scenarios by, for example, Mitsunaga et al. [17] and Tapus et al. [23] applying a Policy Gradient Reinforcement Learning. Following this approach, social signals can be represented as continuous floating point parameters without losing information.

In summary, most previous Reinforcement Learning approaches based on social signals in Human-Robot Interaction utilize a combination of implicit and explicit feedback. However, for a robot to acquire their human counterparts' sense of humor in an unobtrusive manner, the robot would need to learn to be humorous solely by learning from implicit feedback [19]. In this paper, we aim to address this open research question by exploring whether implicit feedback alone suffices to enable a robot to learn humor preferences.

In the next section, we present a social adaptation approach based on Reinforcement Learning, which we believe can achieve good performance results in having a robot learn to tell the right jokes based only on implicit user feedback. We describe in detail technical details, providing arguments for our design choices in our approach, including how we address typical issues with Reinforcement Learning, such as how and if to discretize sensor data related with social signals for the learning approach.

3 SOCIAL ADAPTATION PROCESS

As proposed by Ritschel et al. [19], we set following requirements for the adaptation process: (1) The adaptation has to work in real-time, i.e. it runs while the robot is interacting with users, (2) users are not forced to provide any explicit feedback in order to enable adaptation, e.g. the process utilizes non-verbal social signals, and (3) continuous signals have to be used for learning, i.e. the signals should not be discretized before being employed by the robot.

Ultimately, the social adaptation process is needed to enable the robot to elicit laughter in users by performing the right kind of actions (e.g., make funny grimaces, produce sounds, or tell the right kind of jokes) considering a user's preferences, which are continuously learned during interactions with a user. Figure 2 shows some examples of grimaces our robot is able to make. Figure 1 illustrates the scenario: The robot selects an action depending on the learned preference and presents it to the user who reacts by smiling and laughing. These signals are sensed and processed to compute the probabilities of laughter and smile, which are then used as reward (see fig. 4) to update the user's preferences based on RL.

We employ the Social-Signal-Interpretation framework (SSI) [27] to capture and analyze the required audio-visual signals. Audio signals are used for estimating the probability of the user's vocal laughter, whereas video signals are used to compute the estimated probability of the user's smile.

3.1 Problem definition

Due to our requirements, there are several problems that emerge. Not relying on explicit feedback implies the robot's need of getting another form of feedback that is reliable enough to learn from it and that can be sensed at any time. Although social signals can be easily acquired, they are difficult to interpret. The robot's goal is to keep the amount of the user's vocal laughs and facial smiles as high

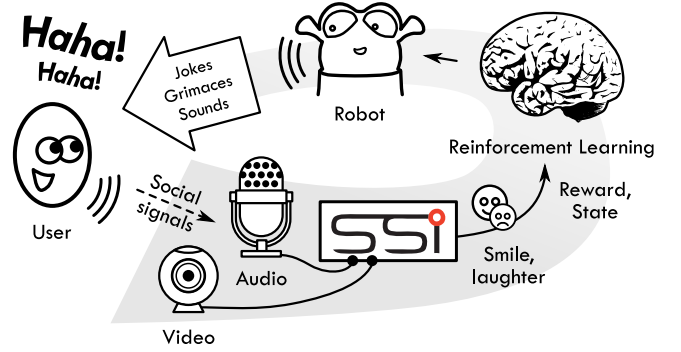


Figure 1: Interaction scenario involving a robot learning how to be funny from human social signals.



Figure 2: Some robot grimaces.

as possible. Therefore, both laughs and smiles are used to compute a reward $\mathcal{R}_t \in [0, 1]$ at time t (as defined in 3.1.4), which gives the robot a good indication about how well it is performing.

3.1.1 State space. Using the signals as state space \mathcal{S} enables the robot to adapt its behavior depending on the user's current enjoyment. E.g. the robot can learn how to react if the user is not amused, which may be different to the behavior when the user is already amused a lot. As we use two different kinds of social signals, we employ a two-dimensional non-discretized vector ϕ as state representation containing the probability of smiles and laughs as follows:

$$\phi = \begin{pmatrix} \mathcal{P}(\text{user's current facial smile}) \\ \mathcal{P}(\text{user's current vocal laughter}) \end{pmatrix} \quad (1)$$

3.1.2 Action space. We gave the robot different action categories as action space \mathcal{A} : sounds, grimaces and three types of jokes and the combination of them as illustrated in Figure 3.

Within each category, we defined different instances that the robot randomly selected from when having chosen an action. Thus, we enabled the robot to learn what category the user finds the funniest in order to not have the robot end up in telling the same joke over and over again. Overall, we defined 108 jokes split into the three joke categories, 19 grimaces and 23 sounds. Consequently, there exist 437 grimace-sound and 2052 grimace-joke combinations.

3.1.3 Algorithm. At the beginning, without any knowledge, the robot tries every of the aforementioned action categories once, measures how funny this category is and calculates an action value from it using RL before switching to an ϵ -greedy strategy with $\epsilon = 0.1$. As we use continuous signals for our state space representation,

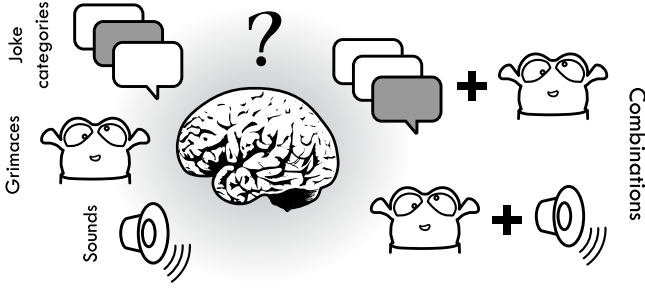


Figure 3: Robot learns the most efficient actions to make the user laugh.

we use a linear function approximator with normally distributed initialized weights based on Sutton et al. [22]

The main idea of linear function approximation is finding a weight vector ω for every action $a \in \mathcal{A}$ in order to compute a function $Q(s, a, \omega)$ that gives an approximation of the optimal action value function $Q^*(s, a)$ for every state $s \in \mathcal{S}$ contemporaneously, i.e. $Q(s, a, \omega) \approx Q^*(s, a)$. This allows for even learning the correct behavior for states not seen by the agent yet.

Every action value $Q(s, a, \omega)$ can be computed by using the following equation 2 [22]:

$$Q(s, a, \omega) := \phi(s) \circ \omega, \forall s \in \mathcal{S}, \forall a \in \mathcal{A} \quad (2)$$

To find the optimal weight vector, it is iteratively updated by using gradient descent. The easiest and conventional way uses the mean-squared-error (MSE) as objective function \mathcal{J} , the derivative of which is defined as follows (see Sutton et al. [22]):

$$\Delta\omega = \alpha \sum_s \mathcal{P}_s (Q^*(s, a) - Q(s, a, \omega)) \phi(s) \quad (3)$$

3.1.4 Reward function. It is important for every Reinforcement Learning problem that a suitable reward function is used. Yet, how can the robot finally use the social signals (see Section 3.2) to compute an appropriate reward that enables him to learn the expected behavior?

Wiewiora [28] indicated that most problems provide intuitive rewards. While executing an action, smile and laughter change continuously. These continuous signals occurring between the start state s_t and the new state s_{t+1} , which we reach after finishing the action, indicate whether or not the user finds the action funny. Thus, they could now be sampled and used for the reward function. Findings by Katevas et al. [9] indicate that people’s response to a joke usually peaks out just after the punchline. For actions not having a punchline, such as grimaces and sounds, Katevas et al. also showed that the best time for measurement is during and right after the action, but not before either.

Thus, we decided to compute an average reward based on all samples from the punchline to the end (Definition 3.1, for actions not having any punchline, the begin of the action is considered as punchline). In our work, we predefined the punchline for every joke in advance. Moreover, using an average based reward function also prevents the robot from learning a wrong behavior by occurring outliers. For the reward calculation, we used both smile and laughter

as visualized in Figure 4, which mathematically translates to the following definition.

Definition 3.1. Let $s_t \in \mathcal{S}$ be the state in which action $a \in \mathcal{A}$ has been chosen and $s_{t+1} \in \mathcal{S}$ the succeeding state. In addition let set \mathcal{F}_1 contain all relevant occurring smile events and set \mathcal{F}_2 all relevant laughter events between s_t and s_{t+1} by omitting all events before the punchline. The reward function \mathcal{R}_t is then defined as:

$$\mathcal{R}(s_t, a, s_{t+1}) = \frac{1}{2} \mathbb{E}[\mathcal{F}_1] + \frac{1}{2} \mathbb{E}[\mathcal{F}_2]$$

The set \mathcal{F}_2 is different to \mathcal{F}_1 because laughter events are sent as they occur, which means, as long as the user is not laughing, the events are 0. In order to give the occurring laughter events a higher weight, we decided to not use them as long as the laughter event has not been detected.

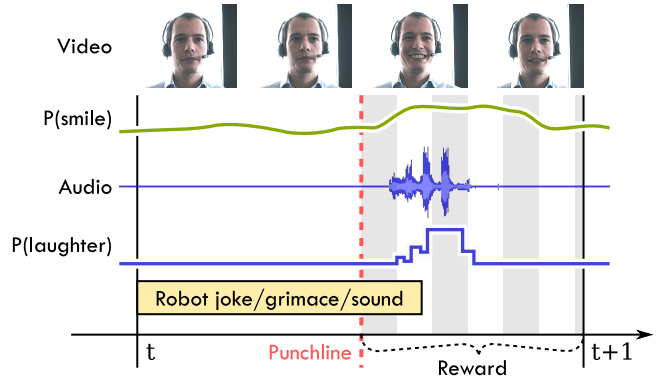


Figure 4: Reward calculation based on social signals.

3.2 Sensing social signals

Laughter has for years been identified as a crucial part of social interaction by traditional conversation analysis [6]. When modeling natural interactions between robots and humans, this key element should consequently not be neglected. The automated recognition of human laughs has been applied to a number of conversational interfaces in order to generate an engaging and pleasant user experience [18, 25]. Furthermore, efforts have been made to collect audio-visual data of human conversational laughter [16] that can be used to build automatic laughter recognition systems.

Given the presented joke telling scenario, a reliable recognition of voiced laughs and visible smiles of the human user is very important to provide social feedback to the robot. Therefore, we capture the human social signals during interactions with audio and video sensors and analyze them with the Social Signals Interpretation (SSI) framework [27] in real-time. In detail, we detect bursts of laugh in the audio signal with framewise recognition approach: The signal is analyzed within a one-second sliding window that is shifted every 400 milliseconds, resulting in a decision rate of 2.5 Hz. The overall activity is monitored by applying a voice activity transformation to the signals via hamming windowing and intensity calculation. We look for coherent signal parts (i.e. frames) in which the mean of squared input values - multiplied by a Hamming window - exceed predefined thresholds for intensity. If we identify

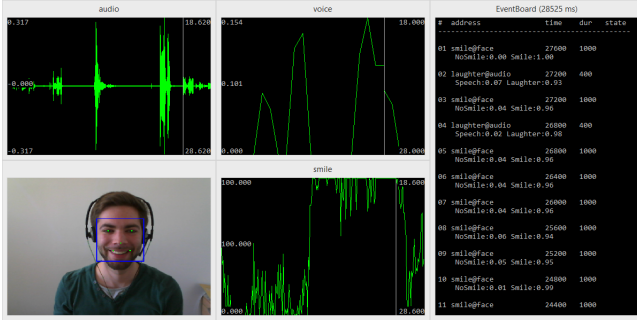


Figure 5: Sensing smile and laughter with SSI framework.

a voiced frame, Support Vector Machine (SVM) classification based on paralinguistic features is applied to decide if the contained vocal activity is more likely to be spoken words or laughter:

In our previous studies on the recognition of laughs [Anonymous reference], Mel Frequency Cepstral Coefficients (MFCC) spectral, pitch, energy, duration, voicing and voice quality features (extracted using the EmoVoice toolbox [26]) were successfully employed to describe the paralinguistic content of voice. These features were used within an SVM model trained on excerpts of the Belfast Storytelling Database [16], which contains spontaneous social interactions and dialogues with a laughter focused annotation. Person independent evaluation of the model on the training database showed an unweighted accuracy of 84% for the recognition of laughter frames.

In the same studies, we used a Microsoft Kinect™ to track 100 facial points, which were used as input for a visual smile classification model with an accuracy of 79%. To enhance performance further and forego the addition of a Kinect™ device in favour of a simple webcam, we chose to base smile recognition on commercial facial recognition software (Sophisticated High-speed Object Recognition Engine (SHORE) [20]) developed by the Fraunhofer Institute for Integrated Circuits IIS. The software has been integrated into the SSI framework. In addition to real-time face tracking, SHORE features a robust emotional analysis component based on facial expression. Using the probabilistic output of the SHORE analysis for each video frame and the same sliding window as used for the audio channel, we calculate a mean probability for user smiles at 2.5 Hz.

As we are interested in the intensity of recognized laughs and smiles as fundamental information for the social adaption process (state space + reward), we chose to provide continuous scores as final output of the recognition system. McKeown [15] states the strong connection of high intensity laughs with perceived humor, a continuous relation that has already been proposed by Darwin mentioning a range "from violent to moderate laughter, to a broad smile, to a gentle smile" [4]. Following these guidelines, it makes sense to incorporate an intensity assessment. To this end, we interpret the confidence scores given by our audio (audible laughs) and video (visual smiles) classifiers as intensity measurements. Though, there are of course other sophisticated approaches to quantify the intensity of vocal laughs [24] and facial smiles [14], we need to take into account the restrictions of real-time capability, and the

probabilistic output of the implemented classification systems turns out to be a good intensity measure that can be computed efficiently.

4 USER STUDY

In order to explore the performance of our social adaptation approach, we conducted a within-subjects user study, comparing our approach's effect (i.e., the performance of a learning robot, which chooses jokes by learning from users' reactions) against a baseline condition (i.e., the performance of a robot, which chooses a random function to choose the next joke). Consequently, the study explored (i) if users laugh more when being entertained by a learning robot and (ii) if users perceive a learning robot as more or less entertaining than a non-learning robot.



Figure 6: Overview of evaluation setup with Reeti robot.

4.1 Participants, apparatus, and procedure

We recruited 24 participants (12 male and 12 female, 18-29 years old) from a university campus. All participants were students. The study started with welcoming participants and providing an introduction to the user study. To this end, participants were handed out a short description, which informed them about the general study procedure, and a short questionnaire to report their self-perceived sense of humor. Moreover, participants were informed that there would be two sessions, one after the other, in which they would be asked to listen to a Reeti robot telling jokes and after each session, they would be asked to provide feedback about their general experience based on a questionnaire. Figure 6 depicts the general study setup, showing the robot, a participant with headphones, and a standard webcam, which was used in combination with the headphone's inbuilt microphone to analyze in real-time participants' visual and auditory reactions. The same setup was used in both sessions.

Participants were told that the two sessions were different in terms of the program version uploaded to the Reeti robot, but participants were not informed that they would experience the Reeti robot once telling random jokes and once telling jokes based on their reactions. At the end of all sessions, participants were asked which version of the Reeti robot they preferred overall.

In both sessions, 25 actions were executed, which took around 10-15 minutes, and the Reeti robot measured the reward of every

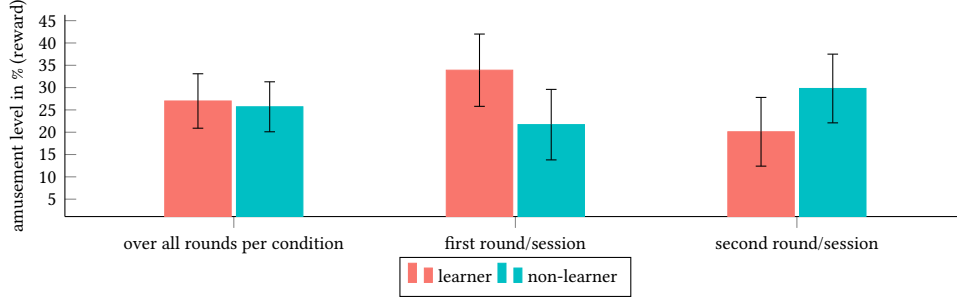


Figure 7: Average reward of both sessions/ amusement level of participants. Error bars denote 95% confidence interval.

performed action during interaction, although it was only used by the learning version (i.e., *learner* condition) to get a comparable value. The order of the sessions was counter-balanced with half of the participants starting with the *learner* version of the robot, and the other half starting with the *non-learner* version of the robot. The aforementioned questionnaire revealed an almost equal rated sense of humor as depicted in Figure 8.

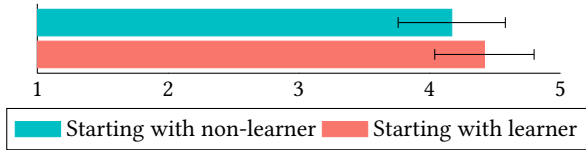


Figure 8: Participants' self-rated sense of humor from non-humorous (1) to humorous (5). Error bars denote 95% confidence interval.

4.2 Results

In the following, we first plot collected data to explore trends in the data and afterward present statistical tests, clarifying the significance of possible trends and systematic differences in user feedback.

4.2.1 General trends. Figure 7 presents a frequency plot of participants amusement levels separated by condition, which shows no obvious difference over all collected output measures (i.e., amusement level/reward). However, when frequencies are separated by round it seems that in round one, participants were more amused by the learning version of the robot and in round two, participants were more amused by the non-learning version of the robot.

There seems to exist a (somewhat unexpected) interaction between round/session (i.e., first and second) and condition (i.e., *learner* and *non-learner*). It seems that on the one hand that participants who experienced the learning version of the robot kept being amused even when in the second session the same robot kept telling randomly chosen jokes. On the other hand, participants who started to experience the robot telling randomly chosen jokes kept feeling less amused in the second session when the robot started to adapt its joke preferences to the user's reaction. Thus, the data seems to show a carryover effect, which we will address in the following statistical analysis and thoroughly discuss it Section 5.

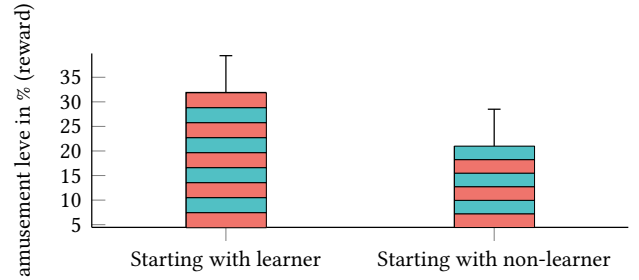


Figure 9: Participants' amusement level (reward) averaged over both rounds. Error bars denote 95% confidence interval.

4.2.2 Statistical analysis. Since repeated measurements were collected from each participant, we performed a paired *Student's t-test*, which shows (as expected after exploring the data plot in Figure 7) that condition (i.e., *learner* vs. *non-learner*) does not effect participants amusement level over all collected data ($M = 0.0123$, $SE = 0.0157$), $t(23) = 1.714$, $p = .022$ [5].

However, when we focus on the evaluation of round one only without including the results of round two to avoid the ordering effect, i.e. we divide our group in one control group (users having watched the non-learning version first) and an experimental group (users having watched the learning version first) with 12 participants each. Using the independent *Student's t-test* shows a significant effect that is also substantial ($p = .02$, $d = 0.87$, see fig. 7) with regard to the gained reward. Comparing the averaged amusement level over both rounds (see fig. 9) also shows an overall significant difference ($p = .03$, $d = 0.83$) between those two groups including the second round. We verified the variances to be homogeneous to satisfy the conditions for applying the independent *t-test*.

4.2.3 User preferences. At the end of the study, we had explicitly asked participants which version of the robot they preferred overall. (Please note, that participants didn't explicitly know how the two versions differed). Surprisingly, most participants preferred the version they experienced last (i.e., 54% preferred the robot they experienced in the second round, 25% preferred the robot they experienced in the first round, and the rest couldn't tell which

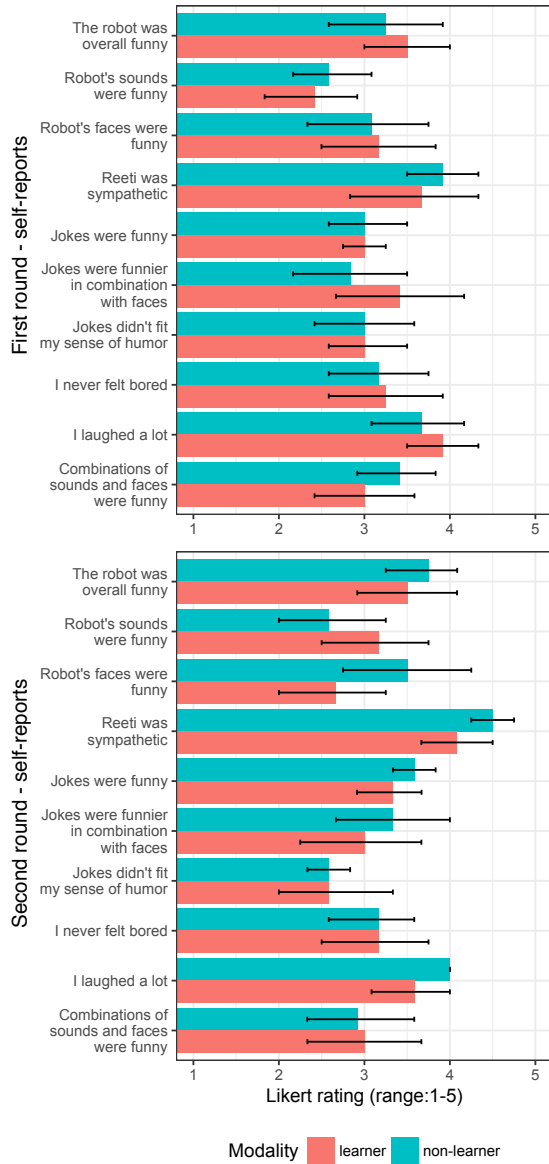


Figure 10: Overview of participants' self-reports, ranging from strongly disagree (1) to strongly agree (5). Error bars denote 95% confidence interval.

version they really preferred). Figure 10 provides an overview of self-reported data; i.e., results of the questionnaire, which we utilized after each of both sessions to collect subjective data about how participants perceived the robot's performance.

In the next section, we discuss implications of the results, including limitations our results may face.

5 DISCUSSION

In the beginning, we have argued that a shared sense of humor can result in positive feelings and moments of bonding, and thus, social robots would benefit from acquiring their human counterparts'

humor. We have also argued that robots should ideally acquire user preferences in an unobtrusive manner through observing and interpreting implicit user feedback, such as smiles. Despite being a challenging task, if social robots were able to shape their own humor and presentation style (e.g., how to include sounds and faces) according to their users' contextual requirements and moods, they would ultimately improve their skills of engagement and possibly also satisfy people's hedonic needs of enjoying everyday life.

To this end, we have presented a social adaptation process based on Reinforcement Learning and conducted a user study to explore how our social adaptation process performs in a joke telling scenario against a baseline condition in which a robot's output (selection of jokes) is based on a random function. We have measured amusement levels and asked for users' preferences in order to assess differences between both conditions. As a study setup, we have chosen a within-subjects study, which would allow us to collect repeated measurements from each participant.

Considering users' preferences, we found that many of the study participants preferred the robot they listened to in the last of two subsequent sessions. The result was somewhat unexpected since we counterbalanced the order of the two different robot conditions (i.e., *learner* and *non-learner*) to prevent ordering and carry-over effects. The data presenting amusement level/reward, which is a non-subjective measure, provides some explanation for why participants started to prefer the second version of the robot. Considering participants' amusement level in each round and condition, Figure 7 shows clearly that participants' amusement level did not change across rounds.

This observation suggests, in retrospective, that choosing a repeated measurement setup might have been naive since the interactions (i.e., telling jokes) and effects (i.e., positive emotions and amusement) that we are targeting may have been of slow nature and difficult to change in short periods of time.

Consequently, we addressed the carry-over effect, first by dropping the data collected in the second round and analyzing the remaining data as between-subjects data. Doing this, we were able to show that robot condition (i.e., *learner* vs. *non-learner*) had a main effect on amusement level and that the learning version of the (same) robot achieved significantly higher levels in amusement than one that doesn't learn from its user's implicit feedback. In summary, this procedure allows to clearly eliminate any carry-over effects and demonstrates that a robot can, based on our social adaptation approach, successfully acquire a users' humor (in a joke telling scenario) to such a degree that results in amusement level will significantly increase compared to a baseline condition.

5.1 Lessons learned

5.1.1 The importance of first impressions. While collecting repeated measurements have been problematic, there are still additional insights and lessons learned to take away. For example, Figure 7 also shows that amusement level was not only significantly influenced by the starting condition (*learner* vs. *non-learner*) (see Figure 9) but that the level stayed stable even when the robot changed its behavior/strategy of selecting and presenting jokes, indicating that in joke telling scenario the "first impression" of a robot is more important than having a robot adapt later on. It seems, that the

good impression in the first session left by the learning version of the robot made the users like the robot and the following non-learning version was found almost equally good because of the first perceived impression. On the other hand, when participants experienced the non-learning version in the first session, it seems to have left a worse first impression and, although in the second session the robot changed its behavior and started to adapt to the users' preferences, the first impression seems to have dominated participants' opinion about the robot. This hypothesis is in line with our study data, which show that in two out of three cases the learning version made users laugh less when being the second session (compared to only one out of six cases when being the first session).

Consequently, applying the ability to adapt to the human preferences can be vain if users have already had a bad first impression of the robot as mentioned above. This also shows that when trying to adapt to humans' preferences, the robot should not make too many mistakes and find out their preferences as quickly as possible.

5.1.2 Learning and Reward Function. Our (average-based) reward function, which uses both users' smile and vocal laughter, was successful, which we verified based on the data collected during the study, showing that the learning robot developed a tendency towards a specific action in contrast to the non-learning version. For example, the percentage of action categories that were frequently chosen (more than 7 times) was much higher for the learning robot (see fig. 11).

During the study we perceived some additional findings that are worth mentioning. The recognition of laughs and smiles was very reliable (see Section 3.2, $\sim 80\%$). A crucial point is, however, to determine when to take measurements as an indicator of the user's level of amusement. As seen in Figure 4, the interval that was used for collecting the frames required for assessing the user's amusement, was statically predefined by the punchline as suggested by Katevas et al. [9]. However, this interval sometimes turned out not to optimally fit the user's reactions to a particular joke, i.e. some users also reacted before the interval, or right after the interval, which resulted in a lower reward than it could actually have been. In the worst case, they only reacted after the interval, where the new action was already in progress because e.g. it took them a while to understand the joke. Nevertheless, such shifts in expected reactions did not happen often and this noise is likely overwritten in subsequent learning steps.

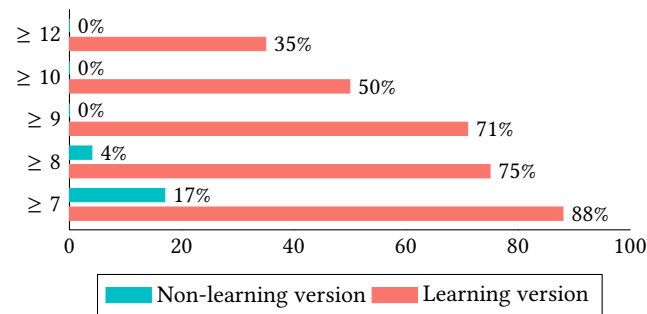


Figure 11: Percentage of frequently chosen action categories

5.1.3 Benefits of variation. As also indicated by participants' preferences and self-reports (see Figure 10; e.g., items *The Reeti was sympathetic* and *Robot's faces were funny*) the non-learning version of the robot was sometimes perceived as funnier. We believe it is because it varies its actions more drastically, which of course was not the goal of the learning version of the robot. In contrast, the learning version aims to continuously figure out, which action the user prefers most, which may result in less variation in shorter periods of time.

In a scenario that aims to entertain listeners, it seems that there are benefits to plan additional moments of surprise and variety. The data we collected has shown us that, a robot should leave a good first impression by adapting to users' sense of humor, but it also indicated that later on there is no harm in integrating variations.

5.2 Limitations and future work

Our work faces some limitations and fellow researchers should interpret results with some care. We have explored our social adaptation process utilizing a simple and predefined test scenario inside a lab, and participants couldn't, for example, explicitly interact with the robot and some participants mentioned this lack of interaction possibilities with the robot as a negative. The monotonous voice of the robot was also mentioned as a reason for turning potentially funny jokes unfunny and participants mentioned that some words were not pronounced correctly, which made it harder for them to perceive the robot as entertaining since participants felt that they had to put more effort into understanding some jokes.

In our future work we aim to build on our results and explore how we need to adapt our social adaption approach to address requirements of more natural and interactive settings in which the interaction space is not completely predefined and the robot has to consider and react to interaction dynamics, including changes in interaction possibilities and user needs over time.

6 CONCLUSION

In this paper, we presented a fully autonomous, real-time adaptation approach solely based on social signals (smile and laughter) and Reinforcement Learning in order to have a robot adapt to a single user's humor preferences without requiring him or her to give explicit feedback.

We have demonstrated the feasibility of the proposed adaptation approach utilizing a user study with a Reeti robot and 24 participants, in which we explored in an exemplary joke telling scenario the adaptation approach's performance. To this end, we have collected (subjective) self-reported data and measured (objective) amusement level of the study participants based on an advanced social signal interpretation setup. Results of the user study have shown that the proposed real-time adaptation approach performs significantly better, considering achieved amusement level against a baseline condition in which jokes were presented in a random fashion. We have described in detail lessons learned, including the importance of a robot to leave a good first impression by learning a user's humor preferences in the beginning. We hope the insights we provided, considering multiple dimensions will help fellow researchers in addressing future challenges in building humor adaptive robots and human-robot interactions.

REFERENCES

- [1] Saleema Amershi, Maya Cakmak, W. Bradley Knox, and Todd Kulesza. 2014. Power to the People: The Role of Humans in Interactive Machine Learning. *AI Magazine* 35, 4 (December 2014), 105–120. <https://www.microsoft.com/en-us/research/publication/power-to-the-people-the-role-of-humans-in-interactive-machine-learning/>
- [2] Rémi Barraquand and James L. Crowley. 2008. Learning polite behavior with situation models. In *Proceedings of the 3rd ACM/IEEE international conference on Human robot interaction, HRI 2008, Amsterdam, The Netherlands, March 12-15, 2008*. 209–216. <https://doi.org/10.1145/1349822.1349850>
- [3] Christian Becker-Asano and Hiroshi Ishiguro. 2009. Laughter in social robotics-no laughing matter. In *Intl. Workshop on Social Intelligence Design*. 287–300.
- [4] Charles Darwin. 1965. *The expression of the emotions in man and animals*. Vol. 526. University of Chicago press.
- [5] A.P. Field and G. Hole. 2003. *How to Design and Report Experiments*. Sage. https://books.google.de/books?id=awtu_c7NofwC
- [6] Phillip J Glenn. 1989. Initiating shared laughter in multi-party conversations. *Western Journal of Communication (includes Communication Reports)* 53, 2 (1989), 127–149.
- [7] Goren Gordon, Samuel Spaulding, Jacqueline Kory Westlund, Jin Joo Lee, Luke Plummer, Marayna Martinez, Madhurima Das, and Cynthia 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, February 12-17, 2016, Phoenix, Arizona, USA*. 3951–3957. <http://www.aaai.org/ocs/index.php/AAAI/AAAI16/paper/view/11759>
- [8] Kotaro Hayashi, Takayuki Kanda, Takahiro Miyashita, Hiroshi Ishiguro, and Norihiro Hagita. 2008. Robot Manzaï: Robot Conversation as a Passive-Social Medium. *I. J. Humanoid Robotics* 5, 1 (2008), 67–86. <https://doi.org/10.1142/S0219843608001315>
- [9] Kleomenis Katevas, Patrick GT Healey, and Matthew Tobias Harris. 2015. Robot Comedy Lab: experimenting with the social dynamics of live performance. *Frontiers in psychology* 6 (2015).
- [10] Elizabeth S Kim and Brian Scassellati. 2007. Learning to refine behavior using prosodic feedback. In *Development and Learning, 2007. ICDL 2007. IEEE 6th International Conference on*. IEEE, 205–210.
- [11] Heather Knight. 2011. *Eight Lessons Learned about Non-verbal Interactions through Robot Theater*. Springer Berlin Heidelberg, Berlin, Heidelberg, 42–51. https://doi.org/10.1007/978-3-642-25504-5_5
- [12] Iolanda Leite, André Pereira, Ginevra Castellano, Samuel Mascarenhas, Carlos Martinho, and Ana Paiva. 2011. Modelling Empathy in Social Robotic Companions. In *Advances in User Modeling - UMAP 2011 Workshops, Girona, Spain, July 11-15, 2011, Revised Selected Papers*. 135–147. https://doi.org/10.1007/978-3-642-28509-7_14
- [13] Changchun Liu, Karla Conn, Nilanjan Sarkar, and Wendy Stone. 2008. Online Affect Detection and Robot Behavior Adaptation for Intervention of Children With Autism. *IEEE Trans. Robotics* 24, 4 (2008), 883–896. <https://doi.org/10.1109/TRO.2008.2001362>
- [14] Robert Lynch. 2010. It’s funny because we think it’s true: laughter is augmented by implicit preferences. *Evolution and Human Behavior* 31, 2 (2010), 141–148.
- [15] Gary McKeown and Will Curran. 2015. The Relationship between laughter intensity and perceived humour. In *The Fourth Interdisciplinary Workshop on Laughter and other Non-Verbal Vocalisations in Speech*. 27–29.
- [16] Gary McKeown, William Curran, Johannes Wagner, Florian Lingensfelser, and Elisabeth André. 2015. The Belfast storytelling database: A spontaneous social interaction database with laughter focused annotation. In *Affective Computing and Intelligent Interaction (ACII), 2015 International Conference on*. IEEE, 166–172.
- [17] Noriaki Mitsunaga, Christian Smith, Takayuki Kanda, Hiroshi Ishiguro, and Norihiro Hagita. 2008. Adapting Robot Behavior for Human–Robot Interaction. *IEEE Trans. Robotics* 24, 4 (2008), 911–916. <https://doi.org/10.1109/TRO.2008.926867>
- [18] Radosław Niewiadomski, Jennifer Hofmann, Jérôme Urbain, Tracey Platt, Johannes Wagner, Bilal Piot, Huseyin Cakmak, Sathish Pammi, Tobias Baur, Stephane Dupont, et al. 2013. Laugh-aware virtual agent and its impact on user amusement. In *Proceedings of the 2013 international conference on Autonomous agents and multi-agent systems*. International Foundation for Autonomous Agents and Multiagent Systems, 619–626.
- [19] Hannes Ritschel, Tobias Baur, and Elisabeth André. 2017. Adapting a Robot’s Linguistic Style Based on Socially-Aware Reinforcement Learning. In *2017 26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. IEEE, 378–384.
- [20] Tobias Ruf, Andreas Ernst, and Christian Küblbeck. 2011. Face detection with the sophisticated high-speed object recognition engine (SHORE). In *Microelectronic Systems*. Springer, 243–252.
- [21] Jonas Sjöbergh and Kenji Araki. 2009. *Robots Make Things Funnier*. Springer Berlin Heidelberg, Berlin, Heidelberg, 306–313. https://doi.org/10.1007/978-3-642-00609-8_27
- [22] Richard S. Sutton, Hamid Reza Maei, Doina Precup, Shalabh Bhatnagar, David Silver, Csaba Szepesvári, and Eric Wiewiora. 2009. Fast Gradient-descent Methods for Temporal-difference Learning with Linear Function Approximation. In *Proceedings of the 26th Annual International Conference on Machine Learning (ICML ’09)*. ACM, New York, NY, USA, 993–1000. <https://doi.org/10.1145/1553374.1553501>
- [23] Adriana Tapus, Cristian Tapus, and Maja J. Mataric. 2008. User - robot personality matching and assistive robot behavior adaptation for post-stroke rehabilitation therapy. *Intelligent Service Robotics* 1, 2 (2008), 169–183. <https://doi.org/10.1007/s11370-008-0017-4>
- [24] Jérôme Urbain, Hüseyin Çakmak, Aurélie Charlier, Maxime Denti, Thierry Dutoit, and Stéphane Dupont. 2014. Arousal-driven synthesis of laughter. *IEEE Journal of Selected Topics in Signal Processing* 8, 2 (2014), 273–284.
- [25] Jérôme Urbain, Radosław Niewiadomski, Elisabetta Bevacqua, Thierry Dutoit, Alexis Moinet, Catherine Pelachaud, Benjamin Picart, Joëlle Tilmann, and Johannes Wagner. 2010. Avlaughtercycle. *Journal on Multimodal User Interfaces* 4, 1 (2010), 47–58.
- [26] Thirid Vogt, Elisabeth André, and Nikolaus Bee. 2008. EmoVoice – A Framework for Online Recognition of Emotions from Voice. In *Tutorial and Research Workshop on Perception and Interactive Technologies for Speech-Based Systems: Perception in Multimodal Dialogue Systems (PIT ’08)*. Springer-Verlag, Berlin, Heidelberg, 188–199. https://doi.org/10.1007/978-3-540-69369-7_21
- [27] Johannes Wagner, Florian Lingensfelser, Tobias Baur, Ionut Damian, Felix Kistler, and Elisabeth André. 2013. The Social Signal Interpretation (SSI) Framework: Multimodal Signal Processing and Recognition in Real-time. In *Proceedings of the 21st ACM International Conference on Multimedia (MM ’13)*. ACM, New York, NY, USA, 831–834. <https://doi.org/10.1145/2502081.2502223>
- [28] Eric Wiewiora. 2010. *Reward Shaping*. Springer US, Boston, MA, 863–865. https://doi.org/10.1007/978-0-387-30164-8_731