



**HAL**  
open science

# Adapting conversational strategies to co-optimize agent's task performance and user's engagement

Lucie Galland, Catherine Pelachaud, Florian Pecune

## ► To cite this version:

Lucie Galland, Catherine Pelachaud, Florian Pecune. Adapting conversational strategies to co-optimize agent's task performance and user's engagement. ACM International Conference on Intelligent Virtual Agents (IVA '22), ACM, Sep 2022, Faro, Portugal. pp.1-3, 10.1145/3514197.3549674 . hal-03818701

**HAL Id: hal-03818701**

**<https://hal.science/hal-03818701v1>**

Submitted on 18 Oct 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

# Adapting conversational strategies to co-optimize agent’s task performance and user’s engagement

Lucie Galland  
Computer Science departement, ENS,  
CNRS, PSL University  
Paris, France  
lucie.galland@ens.psl.eu

Catherine Pelachaud  
CNRS - ISIR, Sorbonne University  
Paris, France  
catherine.pelachaud@upmc.fr

Florian Pecune  
Waseda University  
Tokyo, Japan  
florian.pecune@gmail.com

## ABSTRACT

In this work, we present a socially interactive agent able to adapt its conversational strategies to maximize user’s engagement during the interaction. For this purpose, we train our agent with simulated users using deep reinforcement learning. First, the agent estimates the simulated user’s engagement depending on the latter’s non-verbal behaviors and turn-taking status. This measured engagement is then used as a reward to balance the task of the agent (giving information) and its social goal (maintaining the user highly engaged). Agent’s dialog acts may have different impact on the user’s engagement depending on the latter’s conversational preferences.

## KEYWORDS

Dialog management, Reinforcement Learning, Social Adaptation

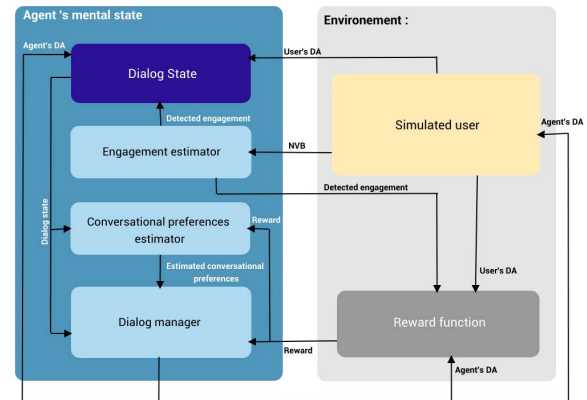
### ACM Reference Format:

Lucie Galland, Catherine Pelachaud, and Florian Pecune. 2022. Adapting conversational strategies to co-optimize agent’s task performance and user’s engagement. In *ACM International Conference on Intelligent Virtual Agents (IVA '22)*, September 6–9, 2022, Faro, Portugal. ACM, New York, NY, USA, 3 pages. <https://doi.org/10.1145/3514197.3549674>

## 1 INTRODUCTION

Engagement ensures high quality user experience during human-machine interactions [7]. In information-giving context, studies have demonstrated that addressees’ level of engagement has a significant impact on their motivation, effort, and memorizing gains [16]. Therefore, one of the main challenges for an information-giving agent is to dynamically manage the conversation by selecting the best dialog policy to fulfill its task and to maintain user’s engagement at the same time.

Manually authoring optimal dialog policies for deep and complex scenarios can be overwhelming [9], especially when the agent has to adapt its behavior according to its user’s conversational preferences. Indeed, while some people will appreciate a friendly interlocutor that liven up the interaction with jokes, personal anecdotes or opinions, others would prefer interacting with someone that solely focuses on the task [14]. Hence the agent should find the right



DA = Dialog act, NVB = non-verbal behavior

Figure 1: System architecture

balance and timing for choosing between using a task-oriented dialog act or a socially-oriented dialog act.

One solution to teach conversational agents the optimal sequence of actions to perform is Reinforcement Learning (RL). Some agents rely on RL to add a social layer on top of a predetermined scenario by learning appropriate social feedback to express during an interaction [1, 4, 6, 15]. Others rely on RL to learn which optimal sequence of actions would help users achieve their task and maximize their engagement at the same time [11, 17]. However, these works consider one single policy for their agent to optimize. None of these systems adapt their conversational behavior depending on whether their users care about the social aspect of the interaction or not. Hence, our aim is to build a Socially Intelligent Agent able to adapt its conversational strategies depending on user’s conversational preferences and deliver a certain quantity of information while maintaining user’s engagement during the interaction.

## 2 OUR APPROACH

Our architecture is composed of 4 modules interacting with each other as depicted in the Fig. 1: a simulated user (SU) (see Section 2.1), an engagement estimator, a conversational preferences estimator, and a dialog manager. At each speaking turn, the SU module approximates the behavior of real users and generates a dialog act [13] and a sequence of non-verbal behaviors [5]. The agent then relies on its engagement estimator to infer the SU’s perceived level of engagement from the latter’s non-verbal behavior and turn-taking information, as in [12]. This estimation updates the dialog state. Depending on how the SU reacted to the agent’s previous action, the agent updates the estimated user conversational preferences (CP) using the CP estimator module. Next, the agent’s dialog

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

IVA '22, September 6–9, 2022, Faro, Portugal

© 2022 Copyright held by the owner/author(s).

ACM ISBN 978-1-4503-9248-8/22/09.

<https://doi.org/10.1145/3514197.3549674>

manager produces either a task-oriented dialog act (ToDA) or a socially-oriented dialog act (SoDA) depending on the current dialog state and the SU’s estimated CP. Finally, the SU updates its own engagement according to the agent’s performed dialog act, and produces another dialog act and a sequence of non-verbal behaviors, updating the dialog state.

## 2.1 User simulator

To simulate users interacting with an information-giving conversational agent, we built the following user simulator.

**Conversational preferences** We rely on [8] to model SU’s conversational preferences. Each SU has one of the following preferences: (1) the SU prefers the agent to use SoDA and will disengage when faced with ToDA, (2) the SU prefers the agent to use ToDA and will disengage when faced with SoDA, or (3) the SU prefers the agent to use socially- as well as task-oriented dialog acts (SToDA).

**Engagement function** The engagement of the SU is updated at each turn according to the current simulated user’s state  $s \in S$ .

$$e_t = e_{t-1} * W(e_{t-1}) + i_{da} * l_{da}(e_{t-1}) - h_{t-1}(a_t) * l_h(e_{t-1}) + b$$

Where  $e_{t-1}$  is the previous engagement value, and  $W(e_{t-1})$  is a Gaussian function representing the weariness applied to the engagement value.  $i_{da}$  is the impact of the dialog act. It is positive if it corresponds to user’s conversational preferences and negative otherwise.  $h_{t-1}$  represents the history of the conversation. A growing penalty is added on the engagement each time the agent uses a strategy to model tiredness faced to repetitions.  $l_h$  and  $l_{da}$  represents the non linearity in engagement variation.  $b$  is a Gaussian random noise to model the variability of human interaction.

**Action selection** Each turn, the SU outputs a dialog act and a list of non-verbal behaviors. As in [10], the SU is a finite state machine. The SU also generates non-verbal behaviors depending on its engagement level by randomly selecting a speaking turn in the NoXi corpus [2] where the novice shows the same level of engagement as the SU. The SU displays the same non-verbal behaviors as the novice in the selected turn.

## 2.2 Adaptive Agent

We endow our agent with the following components:

**Engagement estimator** The engagement estimator is trained using the NoXi corpus [2]. Based on [12][3] and our Boruta analysis of NoXi, the input considered are non-verbal behaviors such as arm openness, arm closed, head nod, head shake, head touch, smile, look away, and whether the user is talking or not. The engagement estimator is composed of 5 linear layers separated by leaky relu activations and dropout layers. The optimal width and depth of the model were determined using a grid search. The model performs with a mean square error of 1.41.

**Conversational preferences estimator** To estimate the user’s CP, we build a 2 layers LSTM neural network. The estimator takes as input the previous state, the previous action, and the obtained reward. The estimator computes two values: the probability that the user prefers the agent to perform social behavior and the probability that the user prefers the agent to focus on the task. This approach of CP estimation detects the CP of the user during a conversation between our dialog manager and user simulator with a mean square error of 0.2.

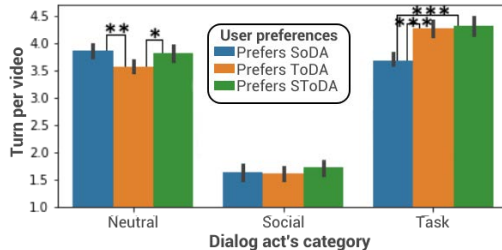


Figure 2: Dialog act distribution

(\* -  $p < 0.1$ , \*\* -  $p < 0.01$ , \*\*\* -  $p < 0.001$ )

**Dialog manager** The goal of the dialog manager is to find the optimal dialog policy and to adapt its conversational strategies to its user’s estimated engagement and estimated CP. The dialog manager is a Deep Q Neural network (DQN) composed of 5 linear layers separated by leaky relu activations and dropout layers. The DQN is trained on 500000 epochs with a buffer of 10000, a discount factor gamma of 0.8 and an exploration factor epsilon starting at 0.95 and progressively decreasing towards 0.05.

The state of our dialog manager is:  $S = \{ \text{user CP, agent’s DA, user DA, mean of the last 3 engagement values, number of turn, estimated topic engagement, historic of the number of time each strategy was used} \}$ . The action space is composed of 3 ToDA, 2 SoDA, and 2 neutral dialog acts. The reward function is crafted to balance the transmission of information and the engagement maximization.

## 3 OBJECTIVE EVALUATION

To evaluate our agent, we analyze the dialog act distribution generated after training (see Fig. 2). A one-way ANOVA is performed to compare the effect of the SU’s conversational preferences on the number of dialog acts generated by our agent. The one-way ANOVA reveals a statistically significant difference in the number of ToDA ( $F(2,970) = 25.85, p = 1.16e-11$ ). Tukey’s HSD test shows that an agent that interacts with a SU who prefers SoDA expresses significantly less ToDA than and an agent interacting with a SU who prefers ToDA ( $p < 2e-16, 95\% \text{ C.I.} = [0.4, 0.8]$ ) or than a SU who prefers socially- and task-oriented dialog acts (SToDA) ( $p < 2e-16, 95\% \text{ C.I.} = [0.4, 0.9]$ ).

## 4 CONCLUSION AND FUTURE WORK

In this work, we propose an agent able to adapt its conversational strategies to the conversational preferences of its users. The simulated users have conversational preferences determining whether they prefer an agent with socially- or task-oriented behaviors. These preferences influence the choice of users’ next dialog acts and engagement levels. The agent’s dialog manager is a DQN trained using RL to optimize both task performance and user’s engagement. The engagement is measured through an engagement estimator that takes as input the user’s non-verbal behavior and outputs the user’s engagement. A conversational preferences estimator is also developed to allow the dialog manager to adapt to the user. After training, we observe that our agent is able to adapt its behavior depending on the user’s conversational preferences. To extend this work, it would be interesting to train our model with a more complex user simulator and to evaluate our agent with real users.

## REFERENCES

- [1] Beatrice Biancardi, Maurizio Mancini, Paul Lerner, and Catherine Pelachaud. 2019. Managing an agent's self-presentational strategies during an interaction. *Frontiers in Robotics and AI* (2019), 93.
- [2] Angelo Cafaro, Johannes Wagner, Tobias Baur, Soumia Dermouche, Mercedes Torres Torres, Catherine Pelachaud, Elisabeth André, and Michel Valstar. 2017. The NoXi database: multimodal recordings of mediated novice-expert interactions. In *Proceedings of the 19th ACM International Conference on Multimodal Interaction*. 350–359.
- [3] Soumia Dermouche and Catherine Pelachaud. 2019. Engagement modeling in dyadic interaction. In *2019 international conference on multimodal interaction*. 440–445.
- [4] 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 AAAI conference on artificial intelligence*, Vol. 30.
- [5] Michele Grimaldi and Catherine Pelachaud. 2021. Generation of Multimodal Behaviors in the Greta platform. In *Proceedings of the 21st ACM International Conference on Intelligent Virtual Agents*. 98–100.
- [6] Iolanda Leite, André Pereira, Ginevra Castellano, Samuel Mascarenhas, Carlos Martinho, and Ana Paiva. 2011. Modelling empathy in social robotic companions. In *International conference on user modeling, adaptation, and personalization*. Springer, 135–147.
- [7] H. O'Brien and E. Toms. 2008. What is user engagement? A conceptual framework for defining user engagement with technology. *J. Assoc. Inf. Sci. Technol.* 59 (2008), 938–955.
- [8] Florian Pecune and Stacy Marsella. 2020. A framework to co-optimize task and social dialogue policies using Reinforcement Learning. In *Proceedings of the 20th ACM International Conference on Intelligent Virtual Agents*. 1–8.
- [9] Charles Rich and Candace L Sidner. 2012. Using collaborative discourse theory to partially automate dialogue tree authoring. In *International Conference on Intelligent Virtual Agents*. Springer, 327–340.
- [10] Jost Schatzmann, Blaise Thomson, Karl Weilhammer, Hui Ye, and Steve Young. 2007. Agenda-based user simulation for bootstrapping a POMDP dialogue system. In *Human Language Technologies 2007: The Conference of the North American Chapter of the Association for Computational Linguistics; Companion Volume, Short Papers*. 149–152.
- [11] Weiyan Shi and Zhou Yu. 2018. Sentiment Adaptive End-to-End Dialog Systems. *arXiv preprint arXiv:1804.10731* (2018).
- [12] Candace L Sidner and Myrosia Dzikovska. 2002. Human-robot interaction: Engagement between humans and robots for hosting activities. In *Proceedings. Fourth IEEE International Conference on Multimodal Interfaces*. IEEE, 123–128.
- [13] Andreas Stolcke, Elizabeth Shriberg, Rebecca Bates, Noah Coccaro, Daniel Jurafsky, Rachel Martin, Marie Meteer, Klaus Ries, Paul Taylor, Carol Van Ess-Dykema, et al. 1998. Dialog act modeling for conversational speech. In *AAAI spring symposium on applying machine learning to discourse processing*. 98–105.
- [14] Karen Tracy and Nikolas Coupland. 1990. Multiple goals in discourse: An overview of issues. *Journal of Language and Social Psychology* 9, 1–2 (1990).
- [15] Klaus Weber, 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 20th ACM international conference on multimodal interaction*. 154–162.
- [16] K. Wentzel. 1999. Social-motivational processes and interpersonal relationships: Implications for understanding motivation at school. *Journal of Educational Psychology* 91 (1999), 76–97.
- [17] Zhou Yu, Alan W Black, and Alexander I Rudnicky. 2017. Learning conversational systems that interleave task and non-task content. *arXiv preprint arXiv:1703.00099* (2017).