Socially-Conditioned Task Reasoning for a Virtual Tutoring Agent

Socially Interactive Agents Track

Zian Zhao
Shanghai Jiao Tong University
Shanghai, SH
zi.an.zhao@sjtu.edu.cn

Michael Madaio
Carnegie Mellon University
Pittsburgh, PA
mmadaio@andrew.cmu.edu

Florian Pecune
Carnegie Mellon University
Pittsburgh, PA
fpecune@andrew.cmu.edu

Yoichi Matsuyama
Carnegie Mellon University
Pittsburgh, PA
yoichim@andrew.cmu.edu

Justine Cassell
Carnegie Mellon University
Pittsburgh, PA
justine@cs.cmu.edu

ABSTRACT

Virtual agents have been shown to be more effective when incorporating social factors such as trust into task action selection. However, there has been less work on how virtual tutoring agents can incorporate social factors into pedagogical action selection. We propose and evaluate how a socially-conditioned task reasoner for a virtual pedagogical agent can incorporate both task and social factors into task reasoning. Our work contributes to the autonomous agent community by providing further evidence that incorporating information about dyadic social factors (e.g. rapport) can be beneficial for agents' task reasoning in the case of a tutoring agent.

KEYWORDS

Socially-aware; pedagogical agent; rapport; RL-LSTM

ACM Reference Format:

Zian Zhao, Michael Madaio, Florian Pecune, Yoichi Matsuyama, and Justine Cassell. 2018. Socially-Conditioned Task Reasoning for a Virtual Tutoring Agent. In *Proc. of the 17th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2018), Stockholm, Sweden, July 10–15, 2018,* IFAAMAS, 3 pages.

1 INTRODUCTION

With the increasing ubiquity of virtual agents in a wide variety of domains, it is becoming ever more critical for those agents to incorporate awareness of social dimensions of interactions [13]. While prior work has investigated the role that social factors like trust play in interactions between humans and agents [3, 7, 14], and modeled the human-agent relationship to inform the agent's selection of high-level tasks [4], there has been less work on using that social dimension to inform the selection of low-level task-related dialogue moves, particularly for tutoring agents.

Over the last several decades, intelligent tutoring systems have been developed to support student learning by providing adaptive problem-selection, step-level instructions, hints or feedback [16]. More recently, researchers have identified optimal pedagogical strategies through reinforcement learning [10, 11]. Prior research suggests that the best teachers build interpersonal closeness with their students, and change their teaching style accordingly [5]. However, while some work has studied virtual tutors' action selection based on students' individual affect [1, 12], those prior work didn't take into account dyadic social states, such as rapport.

In this paper, we investigate how an awareness of dyadic interpersonal social closeness (or, "rapport") and the social conversational strategies that may build that rapport can be incorporated into and improve an agent's task reasoning. Specifically, we focus on two research questions: (1) Will an agent's pedagogical task selection be more effective for learning when incorporating the rapport and social behaviors between tutor and tutee? (2) Will using the task and social interaction history instead of only the current observation improve the performance of a pedagogical reasoner?

2 MODEL CONSTRUCTION

The task of a *socially-conditioned* pedagogical reasoner is as follows: at each time step, the pedagogical reasoner will choose the optimal tutoring strategy from a set of available strategies, to maximize tutees' learning gains over time (normalized gain between pre and post-test [11]), given data of (a) tutees' previous problem-solving performance; (b) the tutees' and tutoring agent's prior labeled utterances; and (c) the rapport level between tutee and agent [8].

We use a peer tutoring dialogue corpus (described in [9]) of 22 pairs of 12-15 year old students tutoring each other in algebra, for a total of 30 hours of multimodal data. The transcripts were annotated for tutoring and learning strategies (TS) (e.g. propose step/answer, ask for help, etc) and two social conversational strategies (SS) (e.g. self-disclosure, praise, etc), described in [8, 9]). Each tutoring session was scored for the problems solved during the session and the normalized learning gain from pre-test to post-test (as in [11]). To find the rapport level for each 30-second "thin-slice" of our corpus, we used a crowd-sourcing approach, described in [8, 17].

To combine social factors (rapport and SS) and task factors (learning performance and TS) into pedagogical reasoning, we model the process into a reinforcement learning problem as follows: **States** are defined as the TS or SS the tutee has just used (e.g. propose step), concatenated with rapport level, if included. **Actions** are the set of TS available for the tutoring agent to choose from, used to tutor the tutees in algebraic problem-solving (e.g. provide hint, feedback,

Proc. of the 17th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2018), M. Dastani, G. Sukthankar, E. André, S. Koenig (eds.), July 10–15, 2018, Stockholm, Sweden.

prompt for step). The **Reward function** assigns rewards according to learners' performance (1 for problem solved, 10 for positive, and -10 for negative learning gains at the end of the session). The **Environment** is modeled by a Markov Decision Process (MDP) with **Discount coefficient** set to 0.98 to ensure the tutoring sessions' duration is realistic. We will refer to this model as an MDP reasoner.

To investigate RQ2, we replace the **Environment** with a long-short term memory (LSTM) to model possible long-term dependencies in pedagogical reasoning. We refer to this as an RL-LSTM reasoner. In order to find the optimal policy, we train the network to approximate the Q value, following the method in [2]. Our architecture contains a layer of 200 LSTM units, followed by a single hidden layer of 100 hidden units, with the input layer corresponding to the dimensions of the concatenated input vector. The input of the model is a concatenated vector:

$$< S_{t-1}, A_{t-1}, S_t, Rp_t >$$

which is the tutee's action (S_{t-1}) and agent's action (A_{t-1}) at the previous time step; the current tutee's action (S_t); and the current rapport level (e.g. low, medium, high) (Rp_t); each one-hot encoded. The output layer is the same length as the set of available task strategy actions for the virtual agent. The output layer is activated by *tanh* function and other layers are activated by *ReLU* function. At each time step, the network selects the action with the largest expected reward to be the agent's next pedagogical action. To compare performance across multiple models (i.e. MDP and RL-LSTM), we evaluate the models by comparing them to the ground truth of the human-human tutoring data in our corpus. Following [15], we compute the similarity of the learned policy to the cases in the corpus that result in larger learning gains, expressed as follows:

$$Sim(\pi_{ij}^*, \pi_i) = 1 - \frac{Lev(\pi_{ij}^*, \pi_i)}{m}$$

where m is number of tutoring moves and $Lev(\pi_{ij}^*, \pi_i)$ Levenshtein distance between learned policy π_{ij}^* and observed policy π_i.

3 EXPERIMENT

To investigate our research questions, we study: (1) the effect of incorporating social factors into the task reasoning models on the expected reward for student learning; and (2) whether using a longer memory of the utterances (task and social) and rapport is beneficial to task reasoning. All p-values are adjusted using the post-hoc Bonferroni correction. For **RQ1**, we investigate two types of social factors involved in the corpus: the thin-slice dyadic rapport and two social conversational strategies previously shown to contribute to rapport-building: self-disclosure and praise. To study the effect of these factors, we construct four MDP reasoners to capture each combination of the different features.

We can see in **Fig.1** that the model with TS, SS, and Rapport receives the highest expected reward (m=6.61, std=6.76). Specifically, incorporating rapport as a single additional input factor to task strategies significantly improves (t=56.04, p<.001) the overall performance of the MDP (m=6.34, std=6.81) over an MDP with TS only (m=3.59, std=8.06), while the model that includes both TS, Rapport, *and* social strategies (SS) performs the best (t=5.43, p<.001). This result suggests that including social factors improves virtual agents' pedagogical task reasoning performance.

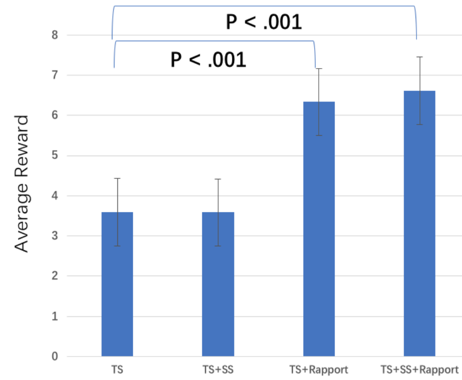

Figure 1: Effect of Social Factors on Task Reasoning in MDP

For **RQ2**, we compared RL-LSTM to MDP to evaluate the impact of memory on model performance. Based on the results from **RQ1**, we use TS, SS, and Rapport for both models. We set 20% of data as the test set and the remaining 80% as training. For test set a_i, we calculate the similarity score $Sim(\pi_{ij}^*, \pi_i)$ and the learning gain R_i of test case a_i, as in Section 2. For a successful pedagogical policy, we expect R_i to be positively correlated to $Sim(\pi_{ij}^*, \pi_i)$. Overall, policies generated by the RL-LSTM have a significant, greater similarity with effective tutoring sessions (ρ =0.591) than those from the MDP (ρ=0.136). This result follows prior work [6] that using information from the tutoring and social interaction history (rather than current information alone) will improve task success.

4 DISCUSSION

In this work, we study the role of social factors in task reasoning by incorporating rapport and social conversational strategies into the input of a pedagogical reasoner, using reinforcement learning. Our results show that social factors have a strong positive effect on a pedagogical task reasoner's performance. More precisely, using as input the learners' social conversational strategies along with the rapport between tutee and agent helps the reasoner make better tutorial decisions, increasing students' learning gain in our models. We also found a memory-based method (RL-LSTM) outperforms a memoryless method (MDP), showing that a tutoring agent that makes its decisions on both the current and previous observations is more effective than one that uses the current information alone.

However, this work has some limitations, namely lacking an evaluation in a deployed system. Further, tutors often teach differently according to their individual tutees, while our model presents a general approach to action selection, given the rapport level. A personalized model may thus be more effective. For future studies, given the flexibility of RL-LSTM models, we also intend to incorporate more features, such as nonverbal and acoustic features, and study their effect on pedagogical reasoning or broader applications for socially-conditioned task reasoning in virtual agents.

We intend for this article to contribute to the autonomous agent community for its study on incorporating social factors into task reasoning, and intend for it to contribute to the design of virtual agents that can interact with people on more human terms.

ACKNOWLEDGMENTS

The work is supported by NSF Award No.1523162, IES Grant R305B150008 to CMU, and the IT R&D program of MSIP/IITP [2017-0-00255,].

REFERENCES

[1] Mohamed Ben Ammar, Mahmoud Neji, Adel M Alimi, and Guy Gouardères. 2010. The affective tutoring system. *Expert Systems with Applications* 37, 4 (2010), 3013–3023.

[2] Bram Bakker. 2001. Reinforcement Learning with Long Short-term Memory. *Advances in Neural Information Processing Systems* (2001).

[3] Justine Cassell and Timothy Bickmore. 2003. Negotiated collusion: Modeling social language and its relationship effects in intelligent agents. *User modeling and user-adapted interaction* 13, 1-2 (2003), 89–132.

[4] William Coon, Charles Rich, and Candace L Sidner. 2013. Activity planning for long-term relationships. In *Intelligent Virtual Agents: 13th International Conference, IVA 2013, Edinburgh, UK, August 29-31, 2013, Proceedings*, Vol. 8108. Springer, 425.

[5] Brandi N. Frisby, Erin Berger, Molly Burchett, Emina Herovic, and Michael G. Strawser. 2014. Participation Apprehensive Students: The Influence of Face Support and Instructor–Student Rapport on Classroom Participation. *Communication Education* 63, 2 (2014).

[6] Helen Hastie, Mei Yii, Srini Janarthanam, Amol Deshmukh, Mary Ellen Foster, and Lynne Hall. 2016. I Remember You! Interaction with Memory for an Empathic Virtual Robotic Tutor. *Proceedings of the 15th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2016)* (2016), 931–939.

[7] Julia M Kim, Randall W Hill Jr, Paula J Durlach, H Chad Lane, Eric Forbell, Mark Core, Stacy Marsella, David Pynadath, and John Hart. 2009. BiLAT: A game-based environment for practicing negotiation in a cultural context. *International Journal of Artificial Intelligence in Education* 19, 3 (2009), 289–308.

[8] Michael Madaio, Rae Lasko, Justine Cassell, and Amy Ogan. 2017. Using Temporal Association Rule Mining to Predict Dyadic Rapport in Peer Tutoring. *In Proceedings of the 10th International Conference on Educational Data Mining* (2017).

[9] Michael Madaio, Kun Peng, Amy Ogan, and Justine Cassell. 2018. A climate of support: a process-oriented analysis of the impact of rapport on peer tutoring. *In Proceedings of the 12th International Conference of the Learning Sciences (ICLS)* (2018).

[10] Kimberly N Martin and Ivon Arroyo. 2004. AgentX: Using reinforcement learning to improve the effectiveness of intelligent tutoring systems. In *International Conference on Intelligent Tutoring Systems*. Springer, 564–572.

[11] Christopher M. Mitchell, Kristy Elizabeth Boyer, and James C. Lester. 2013. When to intervene: Toward a Markov Decision Process dialogue policy for computer science tutoring. *CEUR Workshop Proceedings* (2013).

[12] Aydée Liza Mondragon, Roger Nkambou, and Pierre Poirier. 2016. Evaluating the Effectiveness of an Affective Tutoring Agent in Specialized Education. In *European Conference on Technology Enhanced Learning*. Springer, 446–452.

[13] Florian Pecune, Magalie Ochs, Stacy Marsella, and Catherine Pelachaud. 2016. SOCRATES: from SOCial Relation to ATtitude ExpressionS. *In Proceedings of the 2016 International Conference on Autonomous Agents and Multiagent Systems* (2016), 921–930.

[14] Antonio Roque and David Traum. 2007. A model of compliance and emotion for potentially adversarial dialogue agents. In *Proceedings of the 8th SIGdial Workshop on Discourse and Dialogue*. 35–38.

[15] Jost Schatzmann, Matthew N. Stuttle, KarlWeilhammer, and Steve Young. 2005. Effects of the user model on simulation-based learning of dialogue strategies. *Proc. Autom. Speech Recognit. Understand. Workshop (ASRU'05)* (2005).

[16] Kurt Vanlehn. 2006. The behavior of tutoring systems. *International journal of artificial intelligence in education* 16, 3 (2006), 227–265.

[17] Hongliang Yu, Liangke Gui, Michael Madaio, Amy Ogan, Justine Cassell, and Louis-Philippe Morency. 2017. Temporally Selective Attention Model for Social and Affective State Recognition in Multimedia Content. *In Proceedings of the 25th ACM SIGMM Multimedia Conference* (2017).