Toward Affective Dialogue Modeling using Partially Observable Markov Decision Processes

Trung H. Bui, Job Zwiers, Mannes Poel, and Anton Nijholt

Human Media Interaction, Department of Computer Science, University of Twente
Postbox 217, 7500 AE Enschede, The Netherlands
{buith, zwiers, mpoel, anijholt}@cs.utwente.nl

Abstract. We propose a novel approach to developing a dialogue model which is able to take into account some aspects of the user's emotional state and acts appropriately. The dialogue model uses a Partially Observable Markov Decision Process approach with observations composed of the observed user's emotional state and action. A simple example of route navigation is explained to clarify our approach and preliminary results & future plans are briefly discussed.

1 Introduction

We aim to develop dialogue management models which are able to act appropriately by taking into account some aspects of the user's emotional state. These models are called *affective dialogue models*. Concretely, our affective dialogue manager processes two main inputs, namely the user's action (e.g., dialogue act) and the user's emotional state, and selects the most appropriate system's action based on these inputs and the context. In human-computer dialogue, this work is difficult because the recognition results of the user's action and emotional state are ambiguous and uncertain. Furthermore, the user's emotional state can change quickly. Therefore, an affective dialogue model should take into account both the basic dialogue principles (such as turn-taking and grounding) and the aspects of the user's emotional state. We found that Partially Observable Markov Decision Processes (POMDPs) are suitable for use in designing these affective dialogue models.

In this paper, we first introduce a short overview of POMDP and its application to the dialogue management problem. Second, a general affective dialogue model using POMDP is described. Then, we present a simple example to illustrate our ideas and discuss future work.

2 POMDP and dialogue management

A POMDP is defined by the tuple $\langle S, A, Z, T, O, R \rangle$, where S is the set of states (of the environment), A is the set of the agent's actions , Z is the set of observations the agent can experience of its environment, T is the transition model, O is the observation model, and R is the reward model (Fig. 1a).

In a dialogue management context, the agent is the system (i.e., the dialogue manager) and a part of the POMDP environment represents the user's state. The system uses a state estimator (SE) to compute its internal belief about the user's current state and a policy π to select actions. SE takes as its input the previous belief state, the most recent action and the most recent observation, and returns an updated belief state. The policy π selects actions based on the system's current belief state. Two of the main tasks of a POMDP are computing belief states and finding an optimal policy (i.e., optimal dialogue strategy). These two tasks are explained in [4].

The first work that applied POMDP for the dialogue management problem was the robot home-assistant application [5]. The work following this track is [7, 8]. All these approaches focus on spoken dialogue systems.

3 A POMDP approach to affective dialogue modeling

We select the factored POMDP [2] for representing our affective dialogue model. The state set and observation set are composed of six features. The state set is composed of the `user's goal` (G_u), the `user's emotional state` (E_u), the `user's action` (A_u), and the `user's dialogue state` (D_u) [7]. The observation set is composed of the `observed user's action` (OA_u) and the `observed user's emotional state` (OE_u). Depending on the complexity of the application's domain, these features can be represented by more specific features. For example, the user's emotional state can be encoded by continuous variables such as *valence* and *arousal*, and can be represented using a continuous-state POMDP [3]. The observed emotional state might be represented by a set of observable effects such as response speech, speech pitch, speech volume, posture, and gesture [1].

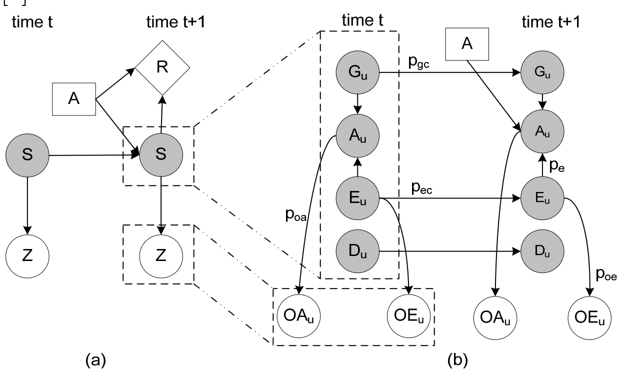

Fig. 1. (a) Standard POMDP, (b) Two time-slice of factored POMDP for the ADM

We are at this moment working with finite-state discrete-time POMDPs. Fig. 1b shows our affective dialogue model (ADM). The features of the state set, action set, observation set, and their correlations form a two time-slice Dynamic Bayesian Network (2TBN). The 2TBN in Fig. 1b is built for our route navigation example that will be presented in Section 4. We can easily modify this 2TBN for representing other correlations, for example the correlation between the user's goal and emotional state. Parameters $p_{gc}, p_{ec}, p_e, p_{oa}$, and p_{oe} are used to produce the transition and observation models in case no real data is available, where p_{gc} and p_{ec} are the probabilities that the user's goal and emotion change; p_e is the probability of the user's action error induced by emotion; p_{oa} and p_{oe} are the probabilities of the observed action and observed emotional state errors.

The reward model depends on each specific application. Therefore, it is not specified in our general affective dialogue model.

4 Example: Route navigation in an unsafe tunnel

We illustrate our affective dialogue model described in Section 3 by a simulated toy route navigation example. An accident happened in a tunnel. A rescue member (denoted by "the user") is sent to the unsafe part of the tunnel to evacuate some injured victims. Suppose the user is in one of three locations (a, b, c). The user interacts with the system. It is able to produce the route description when knowing the user's current location. Furthermore, the system can detect the user's stressful state (*stress* or *nostress*) and uses this information to act appropriately. In this simple example, the system can `ask` the user about his current location, `confirm` a location provided by the user, `show the route description` of a given location, and stop the dialogue by choosing `fail` action.

The POMDP for this problem is represented by $S = \langle G_u \times A_u \times E_u \times D_u \rangle = \langle \{a, b, c\} \times \{a, b, c, yes, no\} \times \{stress, nostress\} \times \{firstturn, nofirstturn\} \rangle$, $A = \{ask, confirm\text{-}a, confirm\text{-}b, confirm\text{-}c, rd\text{-}a, rd\text{-}b, rd\text{-}c, fail\}$, and $O = \langle OA_u \times OE_u \rangle = \langle \{a, b, c, yes, no\} \times \{stress, nostress\} \rangle$. The full flat-POMDP model is composed of 61 states (including a special `end` state), eight actions, and ten observations.

The transition and observation models are generated from the 2TBN (Fig. 1b). We assume that the observed user's action only depends on the true user's action (i.e. $P(oa_u|a_u) = (1 - p_{oa})$ if $oa_u = a_u$, otherwise $P(oa_u|a_u) = 1/4 \times p_{oa}$). The observed user's emotional state is computed in a similar way. We use two criteria to specify the reward model, helping the user obtain the correct route description as soon as possible and maintaining the dialogue appropriateness [7]. Concretely, if the system `confirms` in the first turn, the reward is -2, the reward is -5 for action `fail`, the reward is 10 for action `rd-x` where $g_u = x$ $(x \in a, b, c)$, otherwise the reward is -10. The reward for any action taken in the absorbing *end* state is 0. The reward for any other action is -1.

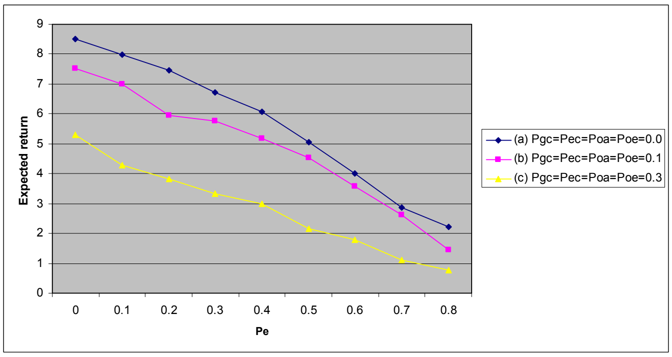

Fig. 2. *Expected return vs. the user's action error induced by stress p_e*

The expected return of the optimal policy (Fig. 2) is computed using the Perseus [6] which is an approximate POMDP algorithm that requires two inputs, a number of belief points and a maximum runtime value. We found 1000 belief points and a runtime of 60 seconds be a good choice for testing our problem. The probability of the user's action error being induced by stress p_e changes from 0 (stress has no influence to the user's action selection) to `0.8` (the user is highly stressed and acts almost randomly). Three lines in Fig. 2 are: no observation error ($p_{oa} = p_{oe} = 0$); low observation error ($p_{oa} = p_{oe} = 0.1$); and high observation error ($p_{oa} = p_{oe} = 0.3$). All these lines show that the expected return of the optimal policy depends on p_e.

5 Conclusions and future work

We have presented a POMDP approach to affective dialogue modeling and illustrated our affective dialogue model by a simple example. The 2TBN representation allows integrating the features of states, actions, and observations in a flexible way. We have also shown that even if the observation is perfect, the expected return of the optimal dialogue strategy depends on the correlation between the user's emotion state and the user's action.

Three important issues we plan to tackle are: (1) scaling up the model with larger state, action, and observation sets for real-world dialogue management problems; (2) extending the model representation, especially by adding more specific features related to the user's goal and emotion and specifying their correlations; and (3) collecting and generating both real and artificial data to build and train the model as well as to validate the model design.

Acknowledgements: This work is part of the ICIS program (http://www.icis.decis.nl). ICIS is sponsored by the Dutch government under contract BSIK 03024.

References

1. Eugene Ball. A bayesian heart: Computer recognition and simulation of emotion. In Paolo Petta Robert Trappl and Sabine Payr, editors, *Emotions in Humans and Artifacts*, chapter 11, pages 303–332. The MIT Press, 2003.
2. Craig Boutilier and David Poole. Computing optimal policies for partially observable decision processes using compact representations. In *AAAI/IAAI, Vol. 2*, pages 1168–1175, 1996.
3. Alex Brooks, Alexei Makarenko, Stefan Williams, and Hugh Durrant-Whyte. Planning in continuous state spaces with parametric pomdps. In *IJCAI Workshop Reasoning with Uncertainty in Robotics*, July 2005.
4. Anthony R. Cassandra, Leslie Pack Kaelbling, and Michael L. Littman. Acting optimally in partially observable stochastic domains. In *Proceedings of the Twelfth National Conference on Artificial Intelligence (AAAI-94)*, volume 2, pages 1023–1028, Seattle, Washington, USA, 1994. AAAI Press/MIT Press.
5. Nicholas Roy, Joelle Pineau, and Sebastian Thrun. Spoken dialogue management using probabilistic reasoning. In *Proceedings of the 38th Annual Meeting of the Association for Computational Linguistics*, 2000.
6. Matthijs T. J. Spaan and Nikos Vlassis. Perseus: Randomized point-based value iteration for pomdps. *Journal of Artificial Intelligence Research*, 24:195–220, 2005.
7. Jason D. Williams, Pascal Poupart, and Steve Young. Factored partially observable markov decision processes for dialogue management. In *4th Workshop on Knowledge and Reasoning in Practical Dialog Systems*, 2005.
8. Bo Zhang, Qingsheng Cai, Jianfeng Mao, and Baining Guo. Spoken dialog management as planning and acting under uncertainty. In *Proceedings of Eurospeech*, 2001.