Statistical Dialog Manager Design Tool for Health Screening and Assessments

Ugan Yasavur, Christine Lisetti, Napthali Rishe

School of Computing and Information Sciences
Florida International University
Miami, FL, 33199, USA
ugan.yasavur@fiu.edu, lisetti@cs.fiu.edu, rishen@cs.fiu.edu

Abstract. We focus on creating a programming tool which enables to create dialog managers for speech-enabled IVA applications in the standardized health screening and assessment domain. Our approach aims to bridge the gap between the intelligent virtual agents (IVA) and the spoken dialog systems (SDS) research communities for the delivery of standardized health interviews by embodied conversational agents (ECA).

Keywords: spoken dialogue system, partially observable Markov decision processes (POMDP), reinforcement learning, embodied conversational agent (ECA), intelligent virtual agents (IVA), brief intervention, behavior change, alcoholism

1 Introduction and Related Research

Latest progress in speech recognition technology, together with advances in the field of conversational intelligent virtual agents (IVA), have created new possibilities to develop a variety of useful applications to address contemporary healthcare challenges. Because current automatic speech recognizers (ASR) are still regarded as a noisy input channels, they need to be backed up with a mechanism to operate against noisy recognitions. In the spoken dialog systems (SDS) area, recent research has mostly concentrated on addressing this problem by employing stochastic and data-driven dialog management (DM) methodologies, namely reinforcement learning based approaches [8]. These approaches, however, have not been widely used by IVA researchers so far.

There are a number of IVA applications in health-related domains such as fitness and activity promotion applications [1], virtual support agents for post traumatic stress disorder [4], health interventions for drinking problems [3]. As an interaction modality, some researchers use ECAs with menu-based inputs [1, 3], others use speech as input modality with simple dialog management [5] and, Morbini et al. [4] uses free-text as input modality.

We aim to bridge the gap between SDS and IVA research, and to use findings from the SDS community for DM in the applications of IVAs in the health domain. The goal of our approach is to create a tool to design custom dialog managers which employs POMDPs as an underlying mechanism. Using our tool, a dialog designer can just specify the content-related information (e.g. question to be asked, information that needs to be provided) and connections between each question in the interview. To facilitate the process, we created an API to design dialog managers. Our tool can be used to create spoken dialog systems for initial screening of patients, conducting brief health interventions and information-providing applications.

2 Approach

Brief interventions for behavior change. We focussed our current approach on an IVA-delivered behavior change brief intervention for excessive alcohol consumption. For the content of the behavior change dialog, we strictly follow the pocket guide from National Institute on Alcohol Abuse and Alcoholism (NIAAA) for alcohol screening and brief interventions for youth [2], which is publicly available online. The initial screen and brief intervention has 3 steps: **1)** Step 1: Ask the two screening questions; **2)** Step 2: *a)* Guide Patient, *b)* Assess Risk **3)** Advise and Assist.

Dialog management. The dialog manager has to track a dialog state which usually contains some important dialog state attributes such as: ASR confidence level, grammar type, information about whether the received answer is confirmed or not, number of re-asks. Each state is mapped to a dialog action that is called *Dialog Policy*. To optimize the system, a reward function is designed. We can formalize the defined mechanism with the Markov decision processes (MDP) framework. MDP assumes that the entire state space is fully observable. However, it is partially observable in SDSs because of imperfect ASR outputs.

An SDS with a partially observable Markov decision processes (POMDP) model attempts to address the partially observable nature of SDS state spaces [7]. According to SDS-POMDP model, at each dialog turn, a user has a goal g in mind (e.g. provide an answer as to the number of days in week s/he drinks, or as to whether alcohol consumption causes any health problems). The system takes a dialog action a (e.g. how many days in a week do you drink alcoholic beverage?) and the user replies with action u ("I usually drink on weekends"). The speech recognizer outputs the N-best list of recognitions $\nu = \{\nu_1, \nu_2, ..., \nu_n\}$ with the estimated confidence scores indicating the likelihood of each recognition being correct, $P(u|\nu)$, while processing the audio. A history variable h keeps track of the relevant dialog history (e.g. receipt of each piece of information, confirmation status of each piece of information). Because ASR is a noisy sensor, g, u, h are not fully observable by the system. Instead, the system maintains a distribution b over these values. Given some existing distribution $b(g, h)$, and observations a and ν', an updated distribution $b'(g', h')$ can be computed [7, 6]:

$$b'(g', h') = k \sum_{\nu'} P(u'|\nu') \sum_{h} P(u'|g', h, a)P(h'|g', u', h, a) \sum_{g} P(g'|a, g)b(g, h) \tag{1}$$

where $P(u'|g', h, a)$ computes how likely are user actions; $P(h'|g', u', h, a)$ computes how the dialog history evolves; $P(g'|a, g)$ computes how the user's goal may change; and k is a normalizing factor.

POMDPs grow exponentially with the number of possible user goals, and it is not possible to calculate this update in real time. This means that POMDP usually suffer from scalability issues [7]. To overcome this problem, a distribution over the set of partitions of user goals $\{p_1, p_2, ..., p_n\}$ is maintained: each partition p_n indicates a collection of user goals, and each user goal can be belong to exactly one partition. The belief in a partition is the sum of the dialog states it contains.

It is assumed that the user's goal is fixed during the interaction, and that error-prone ASR confusions between recognitions that are not on the ASR N-best list are uniform. These two assumptions allow to compute [9]:

$$b'(g', h') = k \sum_{\nu'} P(u'|\nu') \sum_{h \varepsilon p'} P(u'|p', u', h, a) P(p'|p) b(g, h) \tag{2}$$

where $P(p'|p)$ shows the fraction of belief in p which p' would have if p were split into p' and $p - p'$: $P(p'|p) = b_0(p')/b_0(p)$ and $P(p - p'|p) = b_0(p - p')/b_0(p)$, where $b_0(p)$ is the prior probability of a partition p [9].

The partitioning is performed in the following way: first each recognition in the N-best list is compared to each existing partition; if user action can split the partition, the partition is divided. Then the belief in each partition (and dialog histories) is updated using Equation (2). To avoid exponential growth of the number of partitions, low confidence partitions are combined by summing up their beliefs. This approach usually allows to take into account 2-3 N-best recognitions [9]. This problem is addressed by applying incremental partition recombination for tracking dialog states by using a larger number of N-best recognitions [6]. We use the incremental partition combination approach in dialog state tracking [6].

A tool for representing patterns in health brief interviews as programatic objects. The goal of our approach is to create a tool to design custom dialog managers which employs POMDPs as an underlying mechanism. As a result, a dialog designer can just specify the content-related information (e.g. question to be asked, information that needs to be provided) and transitions between each question in the screening. To facilitate the process, we created an API to design dialog managers. Each question represented as an object which encapsulates the dialog policies and the transition information. In other words, each question object consists of a POMDP with transition information to successor step. Since the most of the dialog actions have similar purposes such as asking a question, confirmations, and re-asking a question, it is possible create parameterized patterns that are encapsulated in programmatic objects. Basically, each question object contains all the underlying basic functionality for each piece of information which can be customized. The questions objects can be considered as nodes of a graph, the transitions between questions can be considered as edges that require a key value to transit from one node to another. The key value is a piece of information that the system tries to get from a user in a particular question.

A dialog designer needs to instantiate a question object with at least 4 parameters: **1)** the question text, **2)** the semantic keys which are used to create edges from current node to successor nodes, **3)** the prior node which indicates which is the prior of the current node, and **4)** the semantic-key-to-connect indicates which edge of the prior node it should connect. Creating question objects which encapsulates POMDP mechanism is as easy as specifying some content-related parameters.

3 Conclusion and Future Work

We created dialog manager design tool for creating dialog managers for delivery of standardized health interviews, which will increase the accessibility of state of the art dialog management approaches to non-experts. We also adapted methodologies currently used by the SDS community to health dialogs. As future work, we plan to test and evaluate our system with real users.

References

1. Bickmore, T.W., Schulman, D., Sidner, C.: Automated interventions for multiple health behaviors using conversational agents. Patient education and counseling 92(2), 142–148 (2013)
2. of Health, U.D., Services, H.: Alcohol screening and brief intervention for youth: A practitioner's guide (2011)
3. Lisetti, C., Amini, R., Yasavur, U., Rishe, N.: I can help you change! an empathic virtual agent delivers behavior change health interventions. ACM Transactions on Management Information Systems (TMIS) 4(4), 19 (2013)
4. Morbini, F., Forbell, E., DeVault, D., Sagae, K., Traum, D.R., Rizzo, A.A.: A mixed-initiative conversational dialogue system for healthcare. In: Proceedings of the 13th Annual Meeting of the Special Interest Group on Discourse and Dialogue. pp. 137–139. SIGDIAL '12, Association for Computational Linguistics, Stroudsburg, PA, USA (2012), `http://dl.acm.org/citation.cfm?id=2392800.2392825`
5. Turunen, M., Hakulinen, J., Ståhl, O., Gambäck, B., Hansen, P., Rodríguez Gancedo, M.C., de la Cámara, R.S., Smith, C., Charlton, D., Cavazza, M.: Multimodal and mobile conversational health and fitness companions. Computer Speech & Language 25(2), 192–209 (2011)
6. Williams, J.D.: Incremental partition recombination for efficient tracking of multiple dialog states. In: Acoustics Speech and Signal Processing (ICASSP), 2010 IEEE International Conference on. pp. 5382–5385. IEEE (2010)
7. Williams, J.D., Young, S.: Partially observable markov decision processes for spoken dialog systems. Computer Speech & Language 21(2), 393–422 (2007)
8. Young, S., Gašić, M., Thomson, B., Williams, J.: Pomdp-based statistical spoken dialog systems: A review (2013)
9. Young, S., Gašić, M., Keizer, S., Mairesse, F., Schatzmann, J., Thomson, B., Yu, K.: The hidden information state model: A practical framework for pomdp-based spoken dialogue management. Computer Speech & Language 24(2), 150–174 (2010)