ORIGINAL PAPER

Let's talk! speaking virtual counselor offers you a brief intervention

Ugan Yasavur · Christine Lisetti · Naphtali Rishe

Abstract We developed a virtual counseling system which can deliver brief alcohol health interventions via a 3D anthropomorphic speech-enabled interface—a new field for spoken dialog interactions with intelligent virtual agents in the health domain. We present our spoken dialog system design and its evaluation. We developed our dialog system based on Markov decision processes framework and optimized it by using reinforcement learning algorithms with data we collected from real user interactions. The system begins to learn optimal dialog strategies for initiative selection and for the type of confirmations that it uses during the interaction. We compared the unoptimized system with the optimized system in terms of objective measures (e.g. task completion) and subjective measures (e.g. ease of use, future intention to use the system) and obtained positive results.

Keywords Spoken dialog systems · Reinforcement learning · Intelligent virtual agents and avatars · Conversational agents · Alcohol · Healthy lifestyle screening · Behavior change brief intervention

1 Introduction

Intelligent virtual agents (IVA)—also known as embodied conversational agents (ECA) or virtual humans (VH)—and spoken dialog systems (SDS) are two emerging fields of research which, *together*, could bring a revolution to human-computer interaction as we know it. Even though the term ECA includes the notion of spoken dialog, SDS and ECA communities still do not have a strong connection. While progress in the spoken dialog system area is complementary for the development of conversational embodied agents, latest findings in SDS research have not been commonly used by ECA researchers (and vice versa).

Indeed, although spoken dialog systems (SDS, henceforth) research has shown in the past few years that using reinforcement learning (RL) with MDPs for dialog management outperforms older hand-crafted rule-based approaches [13,56], intelligent virtual agent researchers have not yet integrated these results in their dialog systems. ECA-based systems usually involve spoken dialog (versus menu options to choose from), but their dialog management usually still relies on hand-crafted methods [2,30].

In this project, we bring together latest progress from the SDS community to the IVA community with the use of RL-based dialog management integrated with a 3D animated character (shown in Fig. 1). The 3D animated virtual character is an interface for a task-based spoken dialog to deliver brief alcohol interventions to people at-risk of health issues due to excessive alcohol consumption.

From a computer science perspective, our work aims at building a fully implemented system to be used as screening tools to help individuals at risk of health issues, and at evaluating the system in terms of both, users' (subjective) acceptance and dialog system's (objective) performance.

From a healthcare perspective, we aim at increasing access to effective evidence-based health interventions with a novel mode of delivery for computer-based health interventions - namely delivering health interventions with a virtual counselor. Our screening dialog system brings insight and awareness regarding *alcohol problems* by using the well established

U. Yasavur (✉) · C. Lisetti · N. Rishe
School of Computing and Information Sciences, Florida International University, Miami, FL 33199, USA
e-mail: uyasa001@cis.fiu.edu

C. Lisetti
e-mail: lisetti@cis.fiu.edu

N. Rishe
e-mail: rishen@cis.fiu.edu

Fig. 1 Multimodal embodied conversational agent interface

brief intervention (BI) counseling approach. BIs are short, well structured, one-on-one counseling sessions, focused on specific aspects of problematic lifestyle behavior. BIs are not only ideally suited for people who drink in ways that are harmful or abusive (which is the current domain of our work), but BIs have also been used successfully for a variety of target problem behaviors (e.g. overeating, lack of exercise). Therefore the results of our research will also have an impact on dialog systems for diverse behavior change interventions for healthy lifestyles.

In this article, we give an overview of the current brief intervention counseling style which our system is based on, as well as current progress on spoken dialog systems. We then describe our approach to build a spoken dialog system integrated with an intelligent virtual character to deliver a brief intervention for people at-risk regarding their alcohol consumption. We discuss the results of the evaluation of the system, and conclude with potential future directions for our field of research.

2 Related research

2.1 Brief interventions for at-risk behaviors

Excessive alcohol consumption is regarded as a very worrisome public health problem in the USA: with approximately 85,000 of directly or indirectly attributable deaths per year, excessive alcohol use is the 3rd leading lifestyle-related cause of death in the United States [29]. In 2006, there were more than 1.2 million emergency room visits and 2.7 million physician office visits due to excessive drinking [7]. Excessive alcohol use is also a risk factor for many health and social problems, including motor-vehicle crashes, violence, suicide, hypertension, unsafe sex, or unintended pregnancy. The economic costs of excessive alcohol consumption

in 2006 were approximately $223.5 billion [7]. To attempt to address these alarming statistics, healthcare research has led to the development and deployment of behavior change interventions that can be delivered efficiently in primary care offices.

Brief interventions (BI) are short, well structured, one-on-one counseling sessions, focused on specific aspects of problematic lifestyle behavior, and are ideally suited for people who drink in ways that are harmful or abusive. BIs can be delivered in 3–5 min [32] and (for alcohol consumption as a target) aim to moderate a person's alcohol consumption to reasonable levels and to eliminate harmful drinking behaviors. BIs have a simple approach: they assess an individual's patterns of behavior with respect to a problem behavior, provide tailored feedback, and raise an individual's awareness about the problematic behavior. BIs are the top ranked out of 87 treatment styles in terms of efficiency [27]. It is reported that even a few minutes of discussion about behavioral problems can be as effective as more extended counseling [1]. Many challenges are involved in delivering BIs to people in need, such as finding the time to deliver them in busy doctors' offices, obtaining the extra training that helps staff become comfortable providing these interventions, and managing the cost of delivering the interventions [33].

Patients are often encouraged to use computer programs developed based on BI content in the doctor's waiting room or at home, or to access the interventions through the Internet. Computer-based interventions not only offer privacy, but also the ability to complete the program anywhere, any time of the day [38,39,53]. Although computer-based interventions adapted from one-on-one brief interventions are reported to have positive effects on reducing patients' drinking level [18,39,53], they have high drop-out rates because their users loose interest with interacting with the system. One study showed, however, that the delivery of web-based interventions with virtual agents is promising in terms of increasing people's intention to use such an intervention versus an intervention delivered with text only [26]. That system however is not speech-enabled and the user interacts with mouse and keyboard entries.

We posit that these challenges in administering computer-based brief interventions, can be overcome with the use of spoken dialog systems delivered by an intelligent virtual agents (or embodied conversational agent) which, integrated together, aim at emulating face-to-face conversation, which is the focus of our current research.

2.2 Spoken dialog systems

Dialog systems can be classified into two main categories based on their dialog management technique, which can be either based on machine learning (e.g. based on reinforcement learning), or hand-crafted. Systems based on RL are

popular in the SDS community and are reported to work better than hand-crafted ones for *speech*-enabled systems [13,56] against noisy speech recognition. *Hand-crafted systems*, on the other hand, can be divided into three subcategories, with dialog management approaches using finite states [47], plans and inference rules [6,11] or information states [48].

RL-based dialog systems can learn dialog strategies in a given dialog state from their prior experiences. The idea of having a dialog manager (DM) that can learn interactively from its experience is a cost effective methodology given the alternative approaches: crafting system responses to all possible user's input using rules and heuristics [36]. At best, these rules are based on accumulated knowledge from a trial-and-error experience. At worst, they are based on intuition and limited experience of the designer. Either way, because it is extremely challenging to anticipate every possible user's input, hand-crafting dialog management strategies is an error-prone process that needs to be iteratively refined and tuned [36]. That iterative refinement of course requires substantial amount of time and effort.

The RL-based approach provides the opportunity to automate the design of dialog management strategies by having the system learn these strategies from received reward signals. Potential advantages of statistical dialog management approaches against hand-crafted approaches are listed by [22] as (1) a data-driven automatic development cycle, (2) provably optimal dialog action policies, (3) a principled mathematical model for action selection, (4) possibilities for generalization to unseen states, and (5) reduced development and deployment costs.

Approaches for dialog systems based on reinforcement learning (RL) use Markov decision processes (MDP) [44] or partially observable Markov decision processes (POMDP) frameworks [54,57] to develop robust dialog managers [13, 56]. While both MDPs and POMDPs require high amount of data for training, POMDPs usually suffer from scalability issues [55,58], and optimization algorithms usually become intractable with large number of states.

In this paper we used MDP approaches to avoid the mentioned problems associated with POMDPs. Unlike the classic dialog strategy learning approaches [24] in which the system literally has no knowledge for dialog action selection in the training stage, our system knows which actions make sense in each state, despite being non-optimal as in [44]. For example, taking a farewell action at the beginning of a dialog instead of greeting does not make sense. Our approach enables our system to learn dialog strategies faster from a small amount of dialog corpus than systems with absolutely no knowledge in the training stage. The ideas that are used in [44] inspired our system design decisions: as in the NjFun system [44], we tried to minimize the state space, and to learn dialog policies from real and small amount of data. We extended and adapted

some of these ideas, such as state representations and policy design, and applied them to practical health applications.

RL-based dialog systems are mainly used for slot-filling applications. The domain of the dialog is usually in the tourist information domain, such as finding information about restaurants [8,20,57], appointment scheduling [15], flight reservation [17], or museum guidance [37]. There has also been work conducted in question-answering [28] and tutoring domains [9]. The RL-based dialog management paradigm was also recently used for assistive technologies [25].

In the mentioned systems, each piece of information is accepted as a *slot* which needs to be filled to complete the task. The total number of slots that current systems can usually handle are less than 5, and the flow of a dialog is determined by the slots that need to be filled [44]. One of the important measures of success of task-based dialogs is their *task completion rate*. When the number of the slots that are needed to be filled increases, the likelihood of successful completion decreases.

We will discuss how our system can handle a substantially larger number of slots than has been attempted so far (we have 18 slots). We will also explain how the flow of the dialog does not depend on unfilled slots, but is rather decided dynamically during the interaction, which is also a novel approach. For example, if our system finds out that the person does not have a drinking problem, it just ends the interaction gracefully. Hence, the length (in terms of slots) and the flow of the dialog are not fixed.

Dialog strategy learning is the main concentration of RL-based dialog systems for the optimization process. Learning optimal dialog strategies can be performed in two ways: based on user simulations [14], or based on interactions with real users in exploration mode [28,44]. Since there is no data in our domain, we had to follow the latter way. It is worth mentioning that simulated users can have some shortcomings to realistically reproduce the behaviors of the diverse real users' behaviors [41].

2.3 Assistance technologies for health related dialogs

In a recent comprehensive literature review of active assistance technologies in health-related behavior change systems [21], covering articles written between January 2005 and January 2012, 41 health-behavior change systems are identified. Dialog systems and ECAs are identified as emerging technology themes in the behavior change systems field. It is reported that 19 studies out of 41 use dialog systems technology. Among those 19 systems, only 1 system uses speech as an input modality [23]. The remaining 19 dialog-based systems use text or menu-based choices as a style of communication. In the same literature review [21], it is reported that 8 of the systems use ECAs [4,5,10,49].

There are also dialog systems in other health-related domains such as health and fitness companions [49], or virtual support agents for post traumatic stress disorder [31]. Although, there has been growing interest to develop multimodal SDS which can converse, guide, assist or motivate users for different health related topics [4,12,31], to the best of our knowledge, there does not exist any spoken dialog system for the alcohol consumption domain, which is the focus of our research.

Furthermore, dialog management for health-related dialog systems have so far been mostly designed based on *finite state* dialog management mechanisms such as hierarchical transition networks [3,4], plan-based approaches [42], or information-state based approaches [31]. These systems usually do not have speech recognition integration. Interaction is usually conducted with menu-based choices, but the system utterances are delivered vocally via text-to-speech or prerecorded voice [4,26].

Other than systems that use menu-based interaction, there are systems that use a different input modality. SimCoach, for example, is a web-based system which uses text as input modality, and an interface with an anthropomorphic ECA which responds to users' questions with answers prerecorded by a human. Whereas human voices are still superior to synthetic ones, using pre-recorded utterances means that the sentences that the system can speak are fixed, as opposed to systems using text-to-speech engines (like ours) which provides the flexibility of adding new sentences for the system to utter automatically, i.e. without the need to prerecord new sentences. SimCoach, designed to provide support and healthcare information about post-traumatic stress disorder, incorporates information-state approach [48] with dialog moves with assigned reward values [31]. The Companions project includes three different systems in the domain of health and fitness [49], and all of them use speech as input modality. The dialog management approach in the Companion project is the information-state approach, and one system (the cooking companion) has an ECA interface.

In the remaining of the article, we discuss our system design and approach, and our evaluation experiment results.

3 Approach

An overview of our system architecture is shown in Fig. 3, and explained in details in the Sect. 3.5. In short, we use reinforcement learning for the dialog management, combined with a 3D animated character who converses with the user with text-to-speech (TTS) utterances (versus fixed prerecorded sentences). The domain is a task-based spoken dialog to deliver brief alcohol interventions to people, and identify whether they are at-risk of health issues due to excessive alcohol consumption.

3.1 Dialog structure for brief interventions

According to the clinician's guide for conducting brief interventions from the National Institute on Alcohol Abuse and Alcoholism (NIAAA) [34], a brief intervention for alcohol-related health problems can be delivered in three sequential steps:

– Step 1: *Screening* about alcohol *use*
– Step 2: *Assessing* for alcohol use *disorders*

  – Assessment of *abuse*
  – Assessment of *dependence*

– Step 3: *Advising* and assisting according to degree of alcohol problem

  – Advice for *at-risk* drinkers
  – Advice for drinkers with alcohol use *disorder*

To develop our dialog content, we follow the brief intervention guide for alcohol prepared by NIAAA [33]. The goal of our dialog system is to deliver alcohol screening and brief interventions based on this guide. Each step contains a set of questions.

In *Step 1*, there are 5 questions. The system asks these 5 questions, and if the user expresses that s/he is not consuming alcohol from time to time, the interaction is gracefully terminated by the system. Otherwise, the dialog manager continues to the second step.

In *Step 2*, in the Assessment of *abuse* stage, there are 4 questions to assess alcohol abuse indicators. It is enough to find one indicator of alcohol abuse (e.g. risk of bodily harm, relationship trouble) to move to the Assessment of *dependence* stage (e.g. keep drinking despite problems, not able to stick to drinking limits). If the system can not find any indicator of abuse with the 4 questions, it passes to the dependence stage. In the Assessment of *dependence* stage (still in *Step 2*), there are 7 questions.

It is enough to detect 3 dependence indicators to transit to *Step 3*, Advice for drinkers with alcohol use *disorder*. If the system does not detects 3 dependence indicators, it transits to Advice for *at-risk* drinkers. Therefore, the dialog branches to two separate steps in *Step 3*: (1) one for *at-risk* drinkers, and (2) one for drinkers with alcohol use *disorder*. In both branches, the system provides information related to the assessment of the system. If the system assessed that the user has an alcohol use disorder, it refers the user to treatment, asks the user if she or he is ready to change, and suggests a goal toward a change of drinking patters, based on the user's readiness. If the user is an at-risk drinker, it gauges his or her readiness to change, and provides feedback and information about the person's drinking. Therefore in both stages, the system provides factual information about

the person's drinking and suggested drinking limits, and asks what is the user's intention to change with a single question. In total there can be a maximum of 18 different questions in a single session.

A sample dialog between the system and the user is shown in Table 1. The dialog we presented in the table covers em

Step 1 Screening about alcohol *use*, and *Step 2* Assessment of *abuse* completely. In *Step 2* Assessment of *dependence*, there are 7 questions, and then the system branches to Step 3. Because of space limitations, we cannot present a longer dialog. The system uses questions recommended by NIAAA. It uses simple reflections for confirmations instead of explicit

Table 1 Sample dialog

S1:	Hi, My name is Lola, I am a virtual character created to have conversations with you and to teach you about healthy lifestyles. Today, I'm here to talk about drinking to figure out if you have any issues. May I ask you few questions about your drinking?	*Step 1*: *Use* screening
C1:	Yes, sure you can	
S2:	Do you sometimes drink beer, wine, or other alcoholic beverages?	
C2:	Well yes, I drink wine or beer sometimes!	
S3:	How many times in the past year have you had 5 or more drinks in a day?	
C3:	I think around once a week	
S4:	Actually, consuming more than 4 drinks on any single day is classified as heavy drinking. So it seems that you are sometimes going above heavy drinking limits? Is that right?	
C4:	Yes, that's correct	
S5:	How frequently do you have an alcoholic beverage each week?	
C5:	I think at least 3 days a week	
S6:	On a typical drinking day, how many drinks you have?	
C6:	I just drink... maybe a bottle of whisky	
S7:	Sorry, I could not understand your answer, can you please repeat?	
C7:	I said one bottle of scotch	
S8:	Ok, thanks for the information you have provided to me about your drinking. Next, I will try to determine some important effects of your drinking	
S9:	Have you ever taken risks under the influence, such as drinking and driving, operating a motor vehicle, or swimming?	*Step 2*: *Abuse* Assessment
C9:	No I have not	
S10:	I see, so you are careful to avoid high risk situations after or while having drinks, right?	
C10:	Yes, I try to be careful	
S11:	Have you continued to drink even though you knew it was causing you trouble with your family or friends?	
C11:	I think yes, I continue to drink while my family complains about my drinking	
S12:	Thanks for the information you've provided. In the next step, I will ask you questions about potential dependence-related aspects of your drinking. Is that okay?	
C12:	Yes I want to know, please go ahead	
S13:	Have you had times when you ended up drinking more than you meant to? Or kept on drinking for longer than you intended?	Step 2: *Dependence* assessment
C13:	Not usually, but... yes I guess I did	
S14:	Have you found that you have to drink much more than you once did in order to get the effect that you like?	
C14:	I think my tolerance has increased	
S15:	Sorry, I could not understand your answer. Do you need to drink a lot more sometimes to get the same effect?	
C15:	I meant yes, I need to drink more and more	
S16:	Ok, sorry but just be sure I got it right: you can not get the same effect as you used to with the same amount of alcohol. Is that right?	
C16:	Yes, that is correct	
S17:	Have you more than once wanted to stop or cut down on your drinking? Or tried more than once to stop or cut down but found you couldn't?	

Table 1 continued

C17:	Well, I think I tried but I was not that serious about stopping or cutting down	
S18:	I see, so you really never attempted to stop or cut down. Is that right?	
C18:	No, it is wrong	
S19:	Sorry for the misunderstanding, can you please repeat your answer? Have you more than once wanted to stop or cut down on your drinking but found you couldn't?	
C19:	Yes I did but not seriously	
S20:	Ok, now, when the effects of alcohol are wearing off, have you had trouble sleeping, or found yourself shaking, nervous, nauseous, or restless?	
C20:	I found myself sweating and my heart beating fast	
S21:	Sorry I could not understand your answer. Can you please repeat it?	
C22:	I said I found myself sweating and my heart beating fast	
S22:	I am sorry for asking again, but can you repeat your answer please?	
C22:	I said I am sweating and my heart beats fast when I drink too much	
S23:	Have you had a period when you spent a lot of time drinking? Or being sick or getting over the bad aftereffects of drinking?	
C23:	Yes several times	
S24:	Thank you very much for all your answers. Now I'd like to just take one or two more minutes of your time. Is that okay?	*Step 3*: *Disorder*
C24:	Yes, please go ahead	
S25:	On a scale from 1 to 10, how ready are you to make a change?	
C25:	I think 7	
S26:	Great! Just for your information, men who drink more than 4 standard drinks in a day or more than 14 per week, and women who drink more than 3 drinks in a day or more than 7 drinks per week, are at increased risk of alcohol related problems. Based on your answers, I recommend that you stick with this limit, or even abstain if you can. I also would like you to track your drinking patterns, and I could send you a calendar for that if you'd like me to. I'm also going to give you information about places near you where you can get help if you choose to. Finally, I would like to talk with you again in 3 months. Is that okay?	
S26:	Yes, it's okay	
S27:	Thanks for talking with me. I hope you've learned something useful about your drinking pattern. Good bye and let's talk again in 3 months!	

confirmations, and examples are shown in Table 1 with the S4 and S10 system dialog turns.

3.2 Reinforcement learning

We built our system using reinforcement learning. *Reinforcement Learning* (RL, henceforth) is a sequential decision making algorithm, where the RL agent interacts with its environment [46]. The environment is defined as: "anything that cannot be changed arbitrarily by the agent is considered to be outside of it and thus part of its environment" [46, p. 53]. Reinforcement learning treats the learning of dialog strategies as a sequential optimization problem, leading to strategies which are globally optimal [46].

Within the RL framework for dialog development, dialog strategies are represented as mappings from states to actions within Markov decision processes (MDP). In other words, a dialog strategy specifies, for each system state, what is the

next action to be taken by the system [24]. The MDP framework can be characterized by a 4-tuple (S,A,T,R), where:

- S is a finite set of states
- A is a finite set of actions
- T is a state-transition function such that $T(s, a, s') = \mathbf{P}(s'|s, a)$ which describes how the probability of performing action **a** in state **s** will lead to state **s'**
- **R**(s, a, s') is a local reward function such that $R(s, a) = \sum_a \mathbf{P}(a|s)\mathbf{R}(s, a, s')$, and the objective of the SDS is to maximize the gained reward.

3.3 Markov decision processes

To develop our system, we needed to represent the internal states of our system using MDP formalization as described earlier. States, actions and state transitions needed to be created. First, we elicited the state attributes. We considered the

Table 2 State attributes and values for *Step 1*: screening about alcohol *use*

Attribute	Values	Description
Greet (G)	0,1	Whether the system has greeted the user
Question (Q)	1,2,3,4	Which question is being queried
Confidence (C)	0,1,2,3, 4,5,6	0,1,2 for low, medium, and high confidence of speech recognizer. 3,4 for confirmed or not confirmed. 5 to indicate system is waiting for confirmation. 6 is for to indicate system transit to next question without confirmation
Value (V)	0,1	Is the value obtained for current question
Grammar (Gr)	0,1	What type of ASR grammar used, restrictive or dictation (non-restrictive) grammar.
Aux (Ax)	0,1,2	Multiple purpose attribute. Use to indicate number of ReAsks and semantic valence of the received answer. If it is 0, it indicates, it is not used in that state.

most important dialog attributes to represent the dialog state concisely. To avoid the data sparsity problem during training, we divided the whole system into 5 sections according to the BI guide steps. Hence, there are 5 MDPs in our system. For each step, we elicited related state features.

For each state, there are 5 common *attributes*: Question, Confidence, Value, Grammar, and Aux (see Table 2 for the descriptions of each attribute). There is usually 1 extra attribute to represent the step-specific requirement at each step, e.g. Greet because in the first step the system needs to greet the user first and get consent of the user to start the dialog. The full list of state features for *Step 1* are listed in Table 2.

We then designed possible *states*: for each question, there are 34 possible states. For *Step 1 Use*, *Step 2 Abuse*, *Step 2 Dependence*, *Step 3 At-risk*, and *Step 3 Disorders*, the number of states are 170, 136, 238, 68 and 68 respectively, i.e. 34 states multiplied by the number of questions in a step). Total number of states is 680. Each state represents the state of the conversation.

For example, dialog state *112101* indicates that the system has greeted the user (G = 1), the first question has been queried (Q = 1), the automatic speech recognizer (ASR) confidence level is high (C = 2), the answer value has been obtained (V = 1), the type of grammar is restrictive (Gr = 0), and that the polarity of the alcohol usage indicator is positive (Ax = 1).

An alcohol usage indicator is positive if consuming alcoholic beverages is above recommended limits, and negative

if consumption is below limits. The Aux attribute is used for multiple purposes, such as keeping track of the number of re-asked questions, or of the polarity of alcohol problem indicator. Aux is set to 0 when it is not used.

We refined our state representations by excluding the states that make no sense and manually checking each state. For example, state *040111* is a non-sensible state in *Step 1*. The system can not be in that state because the system needs to greet the user *first*, before it can ask questions. It is not possible to ask question 4 in the first step without greeting or without asking prior questions. Excluding non-sensible states yielded a very large state-space reduction.

We then created our *dialog actions* for each question. Our system uses 2 types of initiative dialog actions: system initiative, where the system asks close-ended questions (e.g. Do you sometimes drink beer, wine, or other alcoholic beverages?), and user initiative, where the system asks open-ended questions (e.g. Can you briefly talk about your alcohol consumption?). We refer to the system/user initiative terms in the same manner as they are widely referred to in the SDS community [43,44]. In the system initiative questions (where the expected answer is relatively restricted given the close-ended nature of the questions), the system uses a restrictive grammar for speech recognition (SR). In the user initiative, the system uses a non-restrictive SR grammar to handle user's answer to open-ended questions.

There are 9 possible actions for each question, which are grouped under 4 categories: (1) *Ask* actions are used when the system needs to ask a question to the user for the first time, which can be performed with the two types of initiatives; (2) *ReAsk* actions are used if the system can not understand the user's speech, which can also be done with two types of initiatives; (3) *Confirmation* actions are used to ask for confirmation as to whether the system understood what the user said.

The confirmation actions are system initiated by default, and are of 3 types: positive, negative and no confirmation. The positive type is used if the system receives an answer which reveals alcohol usage/abuse/dependence information. The negative type is used if the system receives an answer which indicates no drinking/abuse/dependence problem. The third type is no-confirmation, and it is used if the system decides to pass to the next question without confirmation (possible action at any question). *NotConfirmed* action is used if the user gives a negative answer to a confirmation action.

The available actions for the first question in *Step 1* screening about alcohol *use* are shown in Table 3. The first column is the name of the dialog action, and the second column is the system utterance. The actual name of the dialog action starts with the step information (e.g. S1), then the type of dialog action (e.g. ask), then the question being queried (Q1) and the initiative type (sys). So S1-AskQ1Sys stands for *Step1*

Table 3 Dialog actions For Question1 in Step 1

Dialog action	System utterance
S1-AskQ1Sys	Do you sometimes drink beer, wine, or other alcoholic beverages?
S1-AskQ1User	Can you briefly talk about your alcohol consumption?
S1-ReAskQ1Sys	Sorry, I could not understand your answer. Do you sometimes drink beer, wine, or other alcoholic beverages?
S1-RaAskQ1User	Sorry, I could not understand your answer. Can you briefly talk about your alcohol consumption?
S1-ConfQ1Pos	So you like to have alcoholic beverages from time to time, is that right?
S1-ConfQ1Neg	So you are recently not having any alcoholic beverages, is that right?
S1-NoConf	—
S1-NotConfirmedQ1Sys	I am sorry for the misunderstanding. Do you sometimes drink beer, wine, or other alcoholic beverages?
S1-NotConfirmedQ1User	I am sorry for the misunderstanding, can you briefly talk about your alcohol consumption?

(S1), the question type is *Ask*, the question being queried is *question one* (Q1) and the initiative type is *system initiative* (sys).

The number of available actions for each question is 9 (as for the first question shown in Table 3). Although the length of the dialog is not fixed, our system asks a maximum of 18 questions. There are 162 available actions (for asking questions, re-asking questions, and confirmations) for the system to select from in the longest dialog session (18 questions multiplied by the number of available actions). There are dialog actions which are used while transiting from one step to another step (e.g. from *Step 1* to *Step 2 Abuse*) and dialog actions for ending the conversation. There are 2 actions for giving feedback to at-risk drinkers and to drinkers with alcohol use disorder at the end of the each session. The total number of the dialog actions is 169.

After creating the dialog actions, we created the dialog policies. A dialog policy is a mapping of a state to sensible dialog actions. We mapped each state to 2 possible dialog actions based on the initiative or confirmation type. Table 4 shows exploratory dialog policies for Question 1 in Step 1.

As we mentioned earlier, for each question there are 34 states. State updates are performed based on user's dialog actions or on systems dialog actions in each dialog turn. In Table 4, only 30 state-actions mappings that are updated by the system dialog actions or user dialog actions are shown. The remaining 4 states are only updated based on user's dialog actions, which is why we did not include them in Table 4. The reason for this is that, if the system waits for the confirmation from the user (i.e. where C = 5 as shown in see Table 2), the system dialog actions can not be used to update a state. In other words, the remaining 4 states need to be updated by user's dialog actions. In Table 4, we only show the states that are updated by the system. However, the states in Table 4 are the result of the user's dialog actions since *Value Grammar, Confidence* and sometimes *Aux* are updated by user's dialog actions in each dialog turn. For example, when the user speaks to the system, the speech recognizer *Confidence* level

and *Value* attributes are updated based on the user's dialog action. Our system aims to learn approximately optimal dialog strategies for the initiative style and the confirmation type selection.

3.4 Modeling world with interconnected MDPs

To avoid the curse of dimensionality problem, we aimed at minimizing the number of system states used. Since the BI dialog requires many dialog turns between the system and a user, the number of available dialog strategies is very large, and can make learning optimal policies infeasible with limited training data. To alleviate this problem, we used separate MDPs for each phase.

We represent each step or phase of the BI with one MDP with local goals and reward functions. This approach divided the problem into 5 interconnected MDPs (shown in Fig. 2) but, in any interaction with the system, we use a maximum 4 MDPs, i.e. (1) Step 1; (2) Abuse; (3) Dependence; and (4) one MDP from Step 3 based on Abuse or Dependence problem. This approach also reduced the number of required state features for each step, thus reducing the number of states required.

Since there are two phases in *Step 2* (one for querying alcohol abuse and one for querying alcohol dependence), we represent Step 2 with two distinct MDPs (as shown in Fig. 2), which greatly reduces the number of exploratory policies (because it reduces the number of state features) without compromising fine-grained distinctions between dialog strategies. Because the two phases are independent from each other, representing each phase with a separate MDP is appropriate.

There are two separate MDPs for representing the two different phases in *Step 3*. One is used for representing the model for "At-risk" drinkers who do not have alcohol use disorder problems (i.e. no abuse nor dependence). The second one is used to identify drinkers with alcohol use disorders.

Table 4 Explaratory policies for Step 1 Question 1

States						Available actions per state
G	Q	C	V	Gr	Ax	
1	1	0	0	0	0	S1-ReAskQ1Sys, S1-ReAskQ1User
1	1	0	0	1	0	S1-ReAskQ1Sys, S1-ReAskQ1User
1	1	0	0	0	1	S1-ReAskQ1Sys, S1-ReAskQ1User
1	1	0	0	0	2	S1-AskQ2Sys, S1-AskQ2User
1	1	0	0	1	1	S1-ReAskQ1Sys, S1-ReAskQ1User
1	1	0	0	1	2	S1-AskQ2Sys, S1-AskQ2User
1	1	0	1	0	1	S1-ConfQ1Pos, S1-NoConf
1	1	1	1	0	1	S1-ConfQ1Pos, S1-NoConf
1	1	2	1	0	1	S1-ConfQ1Pos, S1-NoConf
1	1	0	1	1	1	S1-ConfQ1Pos, S1-NoConf
1	1	1	1	1	1	S1-ConfQ1Pos, S1-NoConf
1	1	2	1	1	1	S1-ConfQ1Pos, S1-NoConf
1	1	0	1	0	2	S1-ConfQ1Neg, S1-NoConf
1	1	1	1	0	2	S1-ConfQ1Neg, S1-NoConf
1	1	2	1	0	2	S1-ConfQ1Neg, S1-NoConf
1	1	0	1	1	2	S1-ConfQ1Neg, S1-NoConf
1	1	1	1	1	2	S1-ConfQ1Neg, S1-NoConf
1	1	2	1	1	2	S1-ConfQ1Neg, S1-NoConf
1	1	6	1	1	1	S1-AskQ2Sys, S1-AskQ2User
1	1	6	1	1	2	S1-Q1End, S1-Q1End
1	1	6	1	0	1	S1-AskQ2Sys, S1-AskQ2User
1	1	6	1	0	2	S1-Q1End, S1-Q1End
1	1	3	1	0	1	S1-AskQ2Sys, S1-AskQ2User
1	1	3	1	1	1	S1-AskQ2Sys, S1-AskQ2User
1	1	3	1	0	2	S1-Q1End, S1-Q1End
1	1	3	1	1	2	S1-Q1End, S1-Q1End
1	1	4	1	0	1	S1-NotConfirmedQ1Sys,S1-NotConfirmedQ1User
1	1	4	1	1	1	S1-NotConfirmedQ1Sys,S1-NotConfirmedQ1User
1	1	4	1	0	2	S1-NotConfirmedQ1Sys,S1-NotConfirmedQ1User
1	1	4	1	1	2	S1-NotConfirmedQ1Sys,S1-NotConfirmedQ1User

In conclusion, the system is modeled with 5 MDPs. In each MDP, there are multiple terminal states. Some terminal states terminate the Step (such as the consent state), and some terminal states provide transparent transitions to the start state (or start state distribution) of another MDP (see Fig. 2). At the same time, the agent receives a positive reward. The agent also receives immediate positive/negative rewards as showed in Fig. 2. For details on immediate rewards, please see Sect. 3.6. With this approach, learning the optimal dialog strategy for an entire dialog is reduced to learning optimal dialog strategy for each of the MDPs.

3.5 Agent and dialog strategy learning

As shown in Fig. 3, the *Agent* component of the system operates as an interface between other main components of the system. If the system asks a system initiative question, the *Speech Recognizer* component operates by using *Speech Recognizer Grammar Specification (SRGS) grammars*[1], and it outputs *Semantic Interpretation for Speech Recognition*[2] *(SISR)* tags. If the system uses non-restrictive grammar, it uses the *Semantic Parser* to parse the recognized speech. We use the Phoenix robust semantic parser [52], which requires to write context-free grammar style recursive grammars to extract relevant information from the user utterances.

Therefore the *Agent* component receives SISR tags (i.e. when the type of system dialog action is system initiative or closed questions), or Phoenix parse results (i.e. when the type of system dialog action is user initiative or open questions)

---

[1] http://www.w3.org/TR/speech-grammar/.

[2] http://www.w3.org/TR/semantic-interpretation/.

Fig. 2 Representation of world model With MDPs
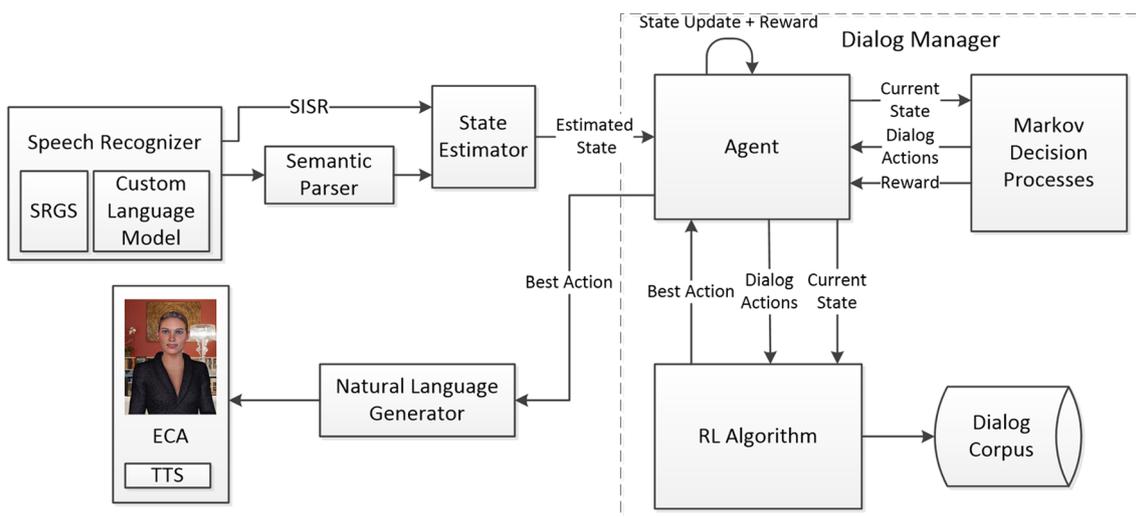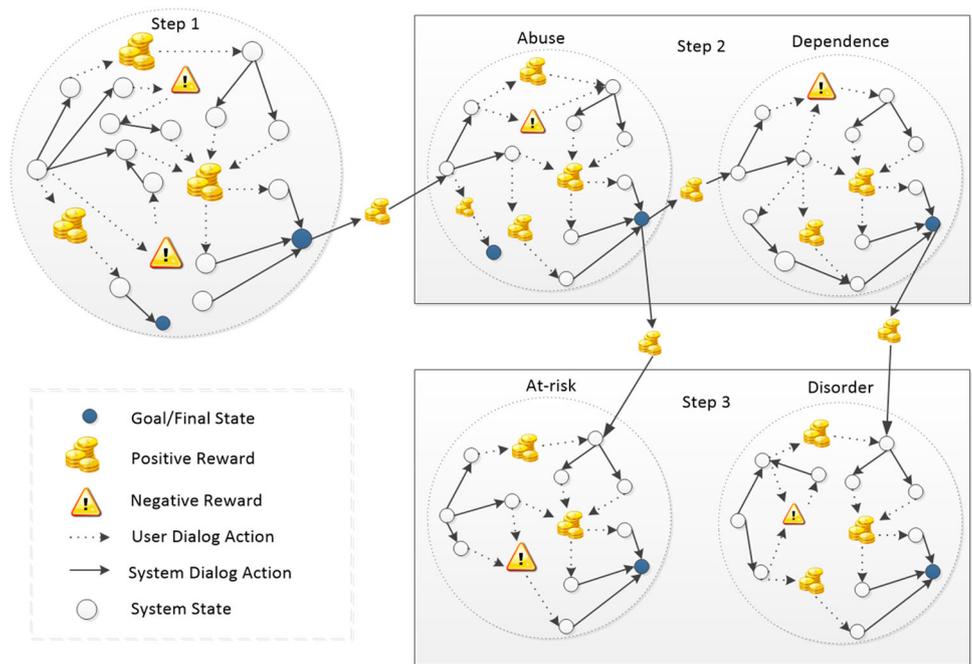
Fig. 3 System architecture and dialog manager

according to the initiative type, as semantic interpretations. The agent updates the system *Current State* and collects the *Reward* according to the reward function (see Sect. 3.6 for the reward function). It then queries the corresponding *Markov Decision Process* with the current state, and receives *Dialog Actions* and a *Reward* information for the current state, and there might not be any associated rewards.

A reward is received only if the *Current State* has an associated *Reward*. For example, the final state of each MDP has associated rewards. The agent sends the received *Dialog Actions* from the MDP and the *Current State* to the *RL algorithm*, and the RL algorithm selects the *Best Action—an action for which the agent received a maximum amount of*

reward in its prior experience— based on the *Dialog Corpus* (see Sect. 3.8) which is collected from real user interactions.The dialog corpus contains information about gained rewards at each step and accumulated rewards for a whole dialog session. The best action is the one that leads the agent to collect the maximum amount of reward. If the system is running in exploration/unoptimized mode, it selects dialog actions randomly among available actions in that state. Therefore, the best action selection does not happen in the unoptimized version which is usually used to collect training data (exploration mode).

The *Best Action* is passed to the *Natural Language Generator* component, which gives the final form of the sys-

tem response and passes the text to the *Text-to-Speech (TTS)* engine. The embodied conversational agent *ECA* utters the response with lip synching. After each dialog turn, the *Dialog Corpus* is updated by the *Agent* with the old dialog state, action, the new dialog state and the reward information. Actually, the corpus contains more information about each turn but the RL algorithm uses reward signals to select the best dialog actions in each state.

At the inception of the project, we did not have any data for optimizing the system for our domain of discourse (the domain of alcohol use). So we first used the system with an algorithm which selects a dialog action randomly among the available ones. Since we have mapped each state to sensible dialog actions, the system was able to deliver basic unoptimized functionality.

After having acquired the *Dialog Corpus* for the domain of alcohol abuse—which is itself a contribution as it can be reused—we used the RL algorithm to learn optimized dialog policies and select the best action according to available data (see Sect. 4.2).

Based on each of our MDPs, the expected cumulative reward *Q(s, a)* of taking *action a* from *state s* can be calculated in terms of Q-values of the next dialog states with the following equation [46];

$$Q^*(s, a) = R(s, a) + \gamma \sum_{s'} \mathbf{P}(s'|s, a) \max_{a'} Q^*(s', a'). \quad (1)$$

where $\mathbf{P}(s'|s, a)$ is the transition model and R(s, a) is the local reward function. The γ ($0 \leq \gamma \leq 1$) is the discount factor which is mainly used to indicate the importance of sooner versus later rewards.

The Q-values in Eq. 1 can be easily computed with a desired threshold using the Q-value version of the standard *Value Iteration algorithm* [46]. The algorithm updates iteratively the current value of Q(s, a) based on the current Q-values, and it stops when the update yields a difference that is below the threshold. Once the Value Iteration algorithm is completed, approximately optimal dialog strategies can be selected by the system, which are essentially dialog actions with the maximum Q-values. The optimized dialog strategy must collect the maximum amount of rewards from future users.

The biggest challenge of this approach is in collecting enough human-machine dialog data to learn an accurate model. To avoid the data sparsity problem, we used minimal state representations and approximated the true state of the system during the interaction. Since the length of the dialog is long, a large amount of data is required to optimize the system. As we describe in Sect. 4.2, we run the systems in two modes, training/exploration and testing. Training mode is for data collection, and in testing mode, the system uses

optimized dialog strategies based on the data collected in training. Therefore, Eq. 1 is used only for testing mode.

3.6 Reward function design

The reward function we use is designed based on the amount of information collected and the cost of collecting each piece of information. The agent gets a reward in each question: if the value is obtained in the first attempt with the *ASK* type of action, it gets +10 reward; if on the other hand the value is not obtained, the agent gets no reward. For each *Confirmation* action, if the obtained value is confirmed by the user, it gets +2, otherwise it gets −2. For each *ReAsk* action which could not result in obtaining the necessary information, the agent receives −3 reward, otherwise it receives +3 reward for the obtained value. If the obtained value is disapproved by the user, it deletes the previously gained reward. Therefore the agent gains a positive or negative reward for each question and dialog action. In addition to rewards gained per question, there are rewards in the MDPs which are associated with the final states. The system receives +15 reward if it is able to reach any of the final states in any MDP. For example, the successful completion of Step 1 gives the agent a +15 reward. In Fig. 2, we depict the immediate rewards and the rewards that are received from the goal states for each MDPs.

We have used this approach to perform strategy learning for each question. Since the system tries to obtain one piece of information in each question, learning the approximately optimal actions in each question is useful.

3.7 Speech recognition and language model

In our system, the operation mode of the speech recognizer[3] is adapted according to the dialog manager's action selection. If the dialog manager asks *system initiative* questions to the user, the system uses *Speech Recognizer Grammar Specification (SRGS) grammars*. Even though we refer to system initiative questions as closed questions, our SRGS grammar does not restrict the user to answer with short answers such as yes/no or a number. It can still understand unrestricted speech. If the system operates in system initiative mode, the Phoenix parser is not used. Instead *Semantic Interpretation for Speech Recognition (SISR)* tags are used. We created a grammar by first authoring it in Augmented Backus-Naur Form (ABNF), and then we converted it to SRGS by using the NuEcho[4] ABNF editor.

Our system uses our custom dictation grammar while it operates in *user initiative* mode. In user initiative mode, we load two types of grammars in the in-process speech recognizer. One is the SRGS grammar which is prepared for the

---

[3] Microsoft Speech Recognizer.

[4] http://www.nuecho.com/en/.

system initiative version of the current question. If the speech recognition result is based on dictation grammar, we use the Phoenix parser, otherwise we use SISR tags. Since the standard dictation language model is comprehensive, it does not work well in specialized domains. To address this problem, we created our own language model by using Windows Vista Dictation Resource Kit software. It is a tool which enables the creation of custom speech recognition dictation language models.

Language models help a speech recognizer decide upon the likelihood of a word sequence. Hence it is useful independently of the acoustics of the word sequences. A language model lets the recognizer make the right guess when two different sentences sound similar. For example, both of the following sentences sound similar: "Because of alcohol, I had hard problems" and "Because of alcohol, I had heart problems". With a language model on alcohol consumption, the recognizer knows that the first sentence is more likely to be what was said than the second one. Furthermore, a language model does not only give information about homonyms, it also gives statistical information about which word might appear after another, among other information. Therefore, if a language model consists of word sequences that are relevant in a specific context, it is very likely that it will operate better than a comprehensive language model for English.

To collect the data for the language model, we first collected data using the Mechanical Turk (MT) crowd sourcing website[5] after obtaining Internal Review Board approval for the study. We asked MT participants the same questions that our system in full mode would ask (after being built from the process described above and after we have acquired the language model). In the instructions, we requested them to role play a person who is having alcohol problems. Our instructions were:

> "*Imagine that you are recently having drinking problems and that you are talking with a health professional face-to-face about your drinking problems. The health practitioner asks you the questions on this page. Please answer as naturally as possible.*"

Because alcohol usage is a very common and universal social problem that everyone understands, MT users' answers were relevant. Once can note that we would not necessarily had collected meaningful answers had we asked MT users, for example, to imagine having some complex disorder such as schizophrenia, because most people do not know what behaviors are associated with this condition. Consuming alcohol in different quantities however, is an experience that many people can relate to, and therefore the answers that we collected were very relevant.

Participants answered the 18 questions. We created the language model from the responses of 447 MT workers. We preprocessed it (corrected spelling and grammar problems) before creating the language model. We improved the language model by adding sentences generated based on our SRGS grammars, and used this language model in our experiments. In the model, there are 7,599 utterances, the average length of an utterance is 11.82 words, there are 100,679 word tokens, and 5,423 distinct words.

We used our custom language model in our evaluation (see Sect. 4.2). We collected the training data from real user dialogs (described in Sect. 3.8) which includes sound files. We ran the speech recognizer on the collected sound files and compared recognitions based on the two language models. We performed quantitative analysis to compare the Microsoft standard dictation language model with our custom language model. We found that when we use our custom language model, the word error rate is approximately 17 % lower than the Microsoft standard dictation language model.

3.8 Dialog corpus

We created a very richly annotated XML-based dialog corpus from the test dialogs, whose size will continue to grow as we collect more data. The corpus is organized turn by turn. Each turn element contains: step and state information, question asked by the system, initiative type, best speech recognition, grammar type, semantic value or result of the Phoenix parser, N-best recognitions with confidence score, reward gained from the question, cumulative reward and sound files. Each XML log file contains sequences of dialog turns for one dialog session.

4 Evaluation

4.1 Participants and procedure

University students represents a very appropriate sample for target population for brief interventions. The latest report of NIAAA on college drinking indicated that alcohol problems are very prevalent among college students [35], and 19 % of college students (ages 18–24) meet the criteria for alcohol abuse or dependence[6]. The use of brief interventions with college students to educate students about drinking and increase their awareness is very common [35]. As a result of many studies, the NIAAA report on college drinking emphasized that "increased alcohol screening and brief interventions are feasible and appropriate for identifying and addressing harmful drinking among college students".

---

6 From the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSMIV), American Psychiatric Association.

In addition, using computer and web-delivered interventions is very well studied in college settings [40,50,51]. For example, Saitz et al. [40] tested the feasibility of providing online alcohol screening and brief intervention to more than one-half of an entire freshman class. The students were contacted through e-mail and invited to take the brief intervention. The researchers found that, in general, unhealthy alcohol use—ranging from risky drinking to alcohol abuse and dependence—decreased following the intervention. Hence, although we are not assessing the impact of the system on heath/drinking outcomes (which would require a randomized clinical trial outside the scope of this study), our target population is very appropriate for participating in brief interventions.

For the evaluation of the system, 89 subjects were recruited from volunteer university students through fliers and emails. From 89 participants, 62 of them were males and 27 of them were females; 51 of them were native speakers and 38 of them were non-native speakers, which realistically represents the diversity of the population in the Miami, Florida area.

Participants sat in front of a PC computer running the systems (some the training system and some the testing system as described below), and responded in English to the questions asked by the embodied conversational agent shown in Fig. 1. The computer was equipped with a USB sound card and a Sennheiser ME 3-ew microphone.

It is important to note that we did *not* perform any user training nor speaker adaptation for speech recognition.

After obtaining an oral consent approved by the University Internal Review Board, we gave the following instructions to each subject before using the system for both experiments:

– You will be asked questions about your drinking behavior with an avatar/virtual character. You may or may not have any alcohol related problems, but we just want you to role-play a person who is having drinking-related problems and give relevant answers to each question.
– Try to speak clearly and loudly enough.
– Wait until the avatar stops speaking before you answer.

4.2 Objective evaluation results

In the *first phase* of the study, for the first 52 subjects, the system operated in training/exploration mode and selected random dialog actions from the available ones in each state (see Sect. 3.5 for discussion). In the second phase, the remaining 37 subject used the system in testing mode. Since, we mapped each state to sensible dialog actions, the system could deliver basic, but expectedly unoptimized functionality. The goal of the first phase was to collect training data to optimize the system for initiative and confirmation type selection.

Table 5 Task completion rate: training versus testing

Evaluation measure	Training	Testing	△	p value
Self-report completion	0.1538	0.5675	0.4137	0.0402
Real completion	0.03846	0.4594	0.42094	0.0434
Step 1: Assessment	0.3461	0.7297	0.3836	0.0371
Step 2: Abuse	0.3076	0.6216	0.3139	0.1058
Step 2: Dependence	0.1923	0.6216	0.4293	0.0300

In the *second phase* of the experiment, the users used the optimized system. Even though the number of subjects is not very large to compute the optimal dialog strategies, it was sufficient to compute approximately optimal dialog strategies. We observe the positive effects of optimization while testing the optimized system.

4.2.1 Task completion evaluation

In Table 5, we present the results of our *task completion evaluation*: Column 1 "Evaluation Measure" is the type of evaluation; Column 2 "Training" is the mean of task completion measure obtained for the training system; Column 3 "Testing" is the mean of task completion for the optimized system; Column 4 "△" shows the difference between testing and training averages; and Column 5 "p value" is the statistical significance value obtained using the standard two-sample t test over subject means.

We show the average values of binary task completion across 52 training dialogs and 37 testing dialogs. At the end of the each interaction, we asked questions to each subject. One of them was "Did you complete the intervention?". If they completed the intervention, the binary completion value was +1, otherwise it was −1. The task completion reported (and perceived) by the subjects is referred to in Table 5 as *Self-Report Completion*.

The additional *Real Task Completion* measure is defined because perceived task completion and real task completion are different. Real task completion indicates whether or not the system obtained all the answers for each question it asked. The perceived (self-report) task completion is different because, if the system can not obtain the answer in three attempts, it skips that question without having an answer and the user is not aware of it.

Three other task completion metrics show the real task completion for each step. The training and testing blocks show averages of binary task completion for each individual version of the system. Since the difference between "real completion" and completion rates for *Step 3* is negligible, we do not report it.

Each row shows a different task completion information and compares the two versions of the system. The first row

is the *Self-Report Task Completion* (perceived) for the whole intervention. The difference between the two versions is statistically significant ($p = 0.0402 < 0.05$)[7]. As mentioned above, the perceived task completion refers to when the subject could complete the intervention, even though there may exist some questions which the system could not obtain answers to, but the user was not aware of it.

The second row shows the *Real Task Completion*, which means that the system did obtain an answer for each question asked. The mean values are lower than self-report completion because the system was able to complete sessions by skipping questions. For example, according to the NIAAA guide for brief interventions which we followed (see details above) [33], it is enough to obtain 1 abuse indicator with the 4 questions which query alcohol abuse. If the system could not obtain an answer to the first three question but obtained an answer to the forth one, the user could still complete the session but from the system's perspective, there are questions which it could not obtain answers to. The difference between the training and the testing system for real task completion is statistically significant ($p = 0.0434 < 0.05$).

The difference in task completion rate for the *Step 1: Assessment* is statistically significant ($p = 0.0371 < 0.05$) for the training and testing versions. Step 1 contains five questions, and since the dialog length is short, a higher task completion rate is expected for both of the versions.

The difference in task completion rate for the *Step 2: Abuse* is not statistically significant ($p = 0.1058 > 0.05$). This is because of the length of the this step. However, as mentioned in the NIAA guide for brief interventions [33], it is sufficient to find a 1 abuse indicator to pass to the "Step 2 Dependence" step.

The difference in task completion rate for *Step 2: Dependence* is statistically significant ($p = 0.0300 < 0.05$). This step is long and the system needs to identify three indicators by using 7 questions. The task completion rates for each sub-steps of *Steps 3: Advise* converge to real task completion rate because it is the end of the intervention. Since the difference between real completion and completion rates for *Step 3* is negligible, we did not report it.

The task completion rate in the training dialogs is 58 %, and for the optimized system it is 77 %, an improvement of task completion rate of 19 %. Although the results we obtained are statistically significant for most of our task-completion criteria, for data hungry reinforcement learning algorithms with a large number of system states, a larger number of subjects will allow us to draw conclusions about the optimality of the learned policies. However, as shown in Fig. 4, we compared Q values for each episode. An episode
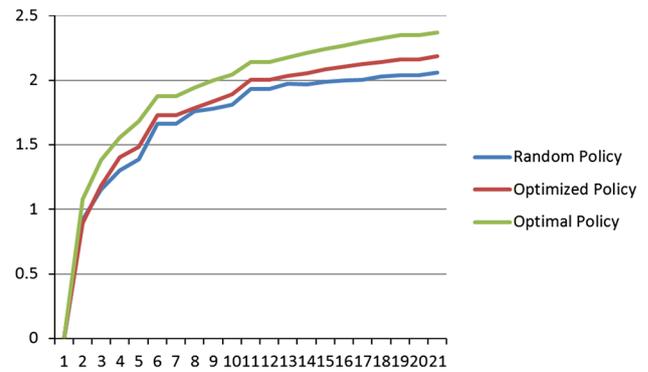
Fig. 4 Q values for each episode, the X axis shows the number of episodes and the Y axis shows the log-scale Q values

can be defined as completing one question and passing to the next question. Completion of a question does not mean that the system obtained the information it was trying to get. As discussed earlier, it is possible for the system to transit to the next question without having obtained the information, and in that case, the system receives negative reward. We described the details of the reward function in Sect. 3.6. We show the improvement of Q-values for each episode in Fig. 4. We have 21 episodes because we have 18 questions, plus transitions between MDPs. As shown in Fig. 4, the optimized policy performed better, even though it is not optimal. We have to note that *optimal policy* represents the highest reward that the system can achieve, whereas the *random policy* and the *optimized policy* represent the average score that the system collected in training and testing operation modes, respectively.

4.2.2 Dialog evaluation

In addition to task metrics, we looked at *Dialog Metrics* to measure the number of turns for successful completions, and the number of words per turn.

The average length of a dialog is 31.9 turns, the shortest completed dialog is 24 turns and the longest one is 43 turns. The length of the dialog is significantly larger than similar RL-based systems [13,44,57]. For Step 1, Step 2 abuse, Step 2 Dependence and Step 3, the average length of the dialog are respectively: 9.6, 4.8 and 13.4, and 4.1. The average number of words used or recognized in each turn is 3.3.

4.3 Subjective evaluation results

4.3.1 User's experience

After the subjects completed the intervention, the subjects answered a survey aimed at evaluating the user's experience with the system. The survey has two parts, the first part has 4 yes/no questions and the second part is a 34-item question-

---

[7] Conventionally, a p value less than 0.05 is considered statistically significant, a p value less than 0.10 is considered indication of a statistical trend.

Fig. 5 Subjective evaluation

Fig. 6 Assessment—negative (1), neutral (4), positive (7)
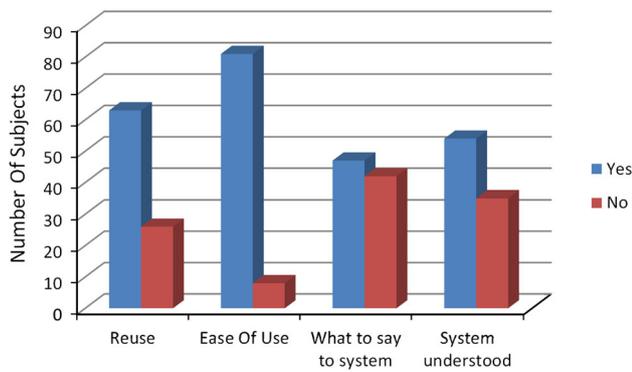
naire about the subject's assessment and experience with the system.

In the first part, we asked questions about reuse "Would you use the system in future?", and ease of use "Is the system easy to use and is it easy to understand how to use the system?", and "Did the system understood what you said" and "Did you know what to say to the system in each turn". Since these 4 questions are not directly related with dialog strategies and we want to see the complete picture, we did not compare test and training systems.

Our evaluation of the subjective aspects shown in Fig. 5 demonstrates that acceptability of the system by users is very high in terms of *Ease of Use* (81 Yes versus 8 No) and *Intention to Reuse* (63 Yes versus 26 No) the system. The *What to say to system* shows that sometimes users do not know how to answer the system questions. We believe that the reason can be that when the system is in user initiative mode (open questions), the subjects may not be sure to what extend they should provide details. The *System understood* criteria shows that most of the users think that the system understood what they said. We postulate that this is achieved with our ample use of confirmation questions that the system utters when not sure.

4.3.2 Subjective assessment of speech interfaces

In the second part of the subjective assessment, we used a 34-item questionnaire named *Subjective Assessment of System Speech Interfaces (SASSI)* [16]. It is a widely used evaluation questionnaire in the SDS community. The subjects answered a randomized list of SASSI questionnaire on a 7-point Likert scale. The SASSI questionnaire queries 6 aspects of the user's assessment and experience with the system. These aspects are *Accuracy*, *Likeability*, *Cognitive Demand*, *Annoyance*, *Habitability*, and *Speed* of the system.

The items in *Accuracy* are related to whether the system recognizes user's input correctly and does what the user expects. The items in *Likeability* include statements about
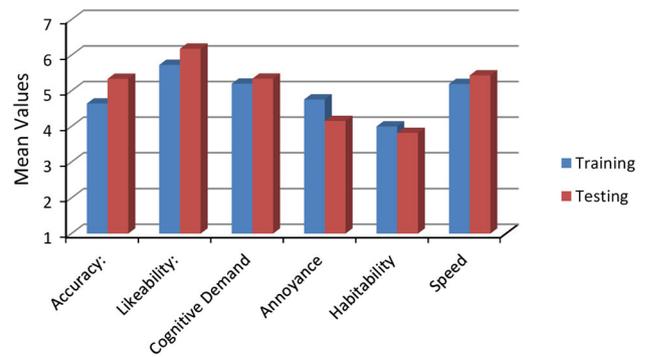
the opinion and feelings of the user about the system. *Cognitive Demand* summarizes the level of effort needed to use the system and the user's feelings arising from this effort. The *Annoyance* includes statements such as "the interaction with the system is repetitive/boring/irritating". *Habitability* contains statements related to whether the user knows what to say and knows what the system is doing. The *Speed* contains only 2 items related to the speed of the system. We compared two versions of the system (training and testing) for the SASSI evaluation. As discussed earlier, **52** subjects used the training system and **37** subjects used the testing version of the system.

We show the results in Fig. 6. In the 7-point Likert scale, 1 is the lowest negative score (strongly disagree), 4 is neutral score (neither agree nor disagree) and 7 is the highest score (strongly agree). We actually compared two versions of the system but our goal was also to assess the overall performance of the system for speed and habitability categories, because both versions of the system do not have any difference in terms of features which are assessed by speed and habitability measures. To be consistent, we compared habitability and speed measures, as we did for other subjective measures. The results for habitability and speed correlate our viewpoint, because the mean values are very close, as showed in Fig. 6.

In Table 6, we show mean values for each evaluation category for both versions of the system, the difference between mean values, and p values. We obtained p values by performing the standard two-sample t test. Column 1 is the type of evaluation; Column 2 is the mean of the evaluated subjective category for training; Column 3 is the mean of the evaluated subjective category training (optimized) system; Column 4 shows the difference between test and training averages; and Column 5 is the statistical significance value obtained using the standard t test.

Accuracy of the system improved in the test version: the results show that there is a statistical significance between the two versions ($p = 0.0360 < 0.05$). This result indicates that the optimized system can select better dialog strategies then the training system which randomly selects dialog strategies.

Table 6 Subjective evaluation categories: training versus testing

Evaluation measure	Training	Testing	△	p value
Accuracy	4.6435	5.3363	0.6928	0.0360
Likebility	5.7286	6.1830	0.4544	0.0928
Cognitive Demand	5.2000	5.3435	0.1435	0.6394
Annoyance	4.7596	4.1540	0.6056	0.0472
Habitability	4.000	3.8200	0.1800	0.6302

Likeability of the system improved slightly in the test version. As can be seen, both versions of the system have very high scores for likeability. It is possible to draw two conclusions: first the acceptance rate of the system is high; second, although the difference between the two versions is not statistically significant ($p = 0.0928 > 0.05$), the optimized behavior of the system provides more desirable interactions.

The mean values of *cognitive demand* and *habitability* are very close for the training and testing versions (see Table 6). Therefore p values are not statistically significant. However, we can infer that the required cognitive demand is slightly higher than neutral level for both versions. Habitability of the system is almost neutral for both of the versions.

We believe that there is a connection between *accuracy* and *annoyance* categories, because if the number of re-asks and confirmation increases, the annoyance level might increase. For the test version, the reported annoyance level decreased and the result is statistically significant ($p = 0.0472 < 0.05$). Since the accuracy also increased for the test version, it might have had a significant impact on the decrease of annoyance.

5 Implications

Health screening and assessment dialogs are different than dialogs that are found in information-seeking applications usually studied by SDS researchers. The main goal of brief behavior change interview dialogs is to collect initial screening information, educate patients, increase their awareness about potential problem behaviors and, if needed, refer the patient to a treatment. This is usually the plan of standardized health interviews (e.g. [19,34,45]) by national or international health institutions. So the system has to conduct the conversation according to that plan. The system usually needs to ask one question at a time and in a specific order, while the flow of the dialog adapts according to the received answers. The length of the dialog is also longer than current information-seeking dialogs.

Our work have several implications. Our reduced state space representation with multiple MDPs enables to learn approximately optimal dialog policies with a relatively low amount of data. Even though we designed the system for

brief alcohol interventions, the approach that we use is easily applicable to any other similar health interviews (e.g. eating behaviors, exercising behaviors, use of drugs). Indeed, brief interventions are adaptable and useful for a variety of life-style related issues that target one specific problematic behavior.

Secondly, our collected dialog corpus will help the development of future data-driven research projects in the health domain.

Thirdly, we connect this work with the notion of intelligent virtual agents (IVA). Whereas we focussed our current discussion on our efficient approach for a spoken dialog-based interaction, our work is directly linked with our research on the graphical animation of the intelligent virtual agents that deliver the spoken intervention. In a recent study [26], we showed that empathic virtual agents that deliver computer-based behavior change interventions are much more engaging than the currently available text-only computer-based interventions. We created a model of empathic communication for an IVA to deliver behavior change interventions: in brief, the agent can sense the user's facial expressions and answers, and adjusts its non-verbal responses accordingly (e.g. express concern or encouragement) to deliver its messages. Whereas there are debates about the impact of virtual characters communicating empathically with humans, our results showed that people are 31 % more likely to use our empathic agent system compared to using the same intervention content delivered instead with text-only. We are currently in the process of integrating and evaluating our empathy agent model with the dialog manager discussed in this article.

Lastly, the performance of our system has also convinced medical and healthcare personnel to conduct randomized clinical trials to evaluate health outcomes and potentially deploy our system in clinicians' waiting rooms and community centers. Whereas computer scientists might think that the healthcare profession could be threatened by the creation of such virtual counselor technologies, they are instead quite enthusiastic about getting technological assistance to address some of the nations' current epidemics (e.g. obesity, overweight, which put people at risk of a variety of chronic conditions such as diabetes, cardiovascular diseases, among others). Virtual counselors have many advantages, including increased accessibility to cost effective health interventions for people in need, increased anonymity and therefore self-disclosure of at-risk behaviors, which in turn leads to better healthcare, among many others [26].

6 Conclusion and future research

We created a spoken embodied conversational system which uses the Reinforcement Learning (RL) paradigm for dialog

management. The system is able to learn dialog strategies for initiative and confirmation selection. Our contributions to the SDS domain include the creation of a RL paradigm to the completely new domain of behavior change—where our dialog length is 4–5 times longer and where the nature of the dialog is less restricted than spoken dialog systems operated in tourist information domain.

We contributed to the healthcare domain with the first system to use speech as an input medium with a RL-based approach. Our initial evaluation showed that the dialog managers that are optimized with RL have the potential to reach optimal behavior, given enough training data.

Our future research will involve extending our evaluation with more training data, and testing the optimized system with a larger number of subjects. Our system currently takes into account the best recognition of the speech recognizer. We plan to use partial observability concepts to deal with uncertainty, which stems from speech recognizer hypotheses: future versions may work with N-best speech recognitions instead of best speech recognition.

References

1. Babor TF, Grant M (1992) Programme on substance abuse: project on identification and management of alcohol-related problems. Report on phase II, an randomized clinical trial of brief interventions in primary health care
2. Bickmore T, Gruber A (2010) Relational agents in clinical psychiatry. Harv Rev Psychiatry 18(2):119–130
3. Bickmore TW, Puskar K, Pfeifer LM, Sereika SM (2010) Maintaining reality: relational agents for antipsychotic medication adherence. Interact Comput 22(4):276–288
4. Bickmore TW, Schulman D, Sidner CL (2011) A reusable framework for health counseling dialogue systems based on a behavioral medicine ontology. J Biomed Inf 44(2):183–197
5. Bickmore TW (2005) Establishing and maintaining long-term human-computer relationships. ACM Trans Comput Hum 12(2):293–327
6. Bohus D, Rudnicky AI (2009) The ravenclaw dialog management framework: architecture and systems. Comput Speech Lang 23(3):332–361
7. Bouchery EE, Harwood HJ, Sacks JJ, Simon CJ, Brewer RD (2011) Economic costs of excessive alcohol consumption in the US, 2006. Am J Prev Med 41(5):516–524
8. Chandramohan S, Geist M, Pietquin O (2010) Sparse approximate dynamic programming for dialog management. In: Proceedings of the 11th annual meeting of the special interest group on discourse and dialogue, SIGDIAĽ10. Association for Computational Linguistics, Stroudsburg, pp 107–115
9. Chi M, Van Lehn K, Litman D (2010) Do micro-level tutorial decisions matter: applying reinforcement learning to induce pedagogical tutorial tactics. In: Intelligent tutoring systems. Springer, Berlin, pp 224–234
10. de Rosis F, Novielli N, Carofiglio V, Cavalluzzi A, De Carolis B (2006) User modeling and adaptation in health promotion dialogs with an animated character
11. Ferguson G, Allen JF et al (1998) Trips: an integrated intelligent problem-solving assistant. In: Proceedings of the national conference on artificial intelligence. Wiley, New York pp 567–573
12. Ferguson G, Quinn J, Horwitz C, Swift M, Allen J, Galescu L (2010) Towards a personal health management assistant. J Biomed Inf 43(5 Suppl):S13–S16
13. Frampton M, Lemon O (2009) Recent research advances in reinforcement learning in spoken dialogue systems. Knowl Eng Rev 24(4):375–408
14. Georgila K, Henderson J, Lemon O (2006) User simulation for spoken dialogue systems, learning and evaluation. In: In Interspeech/ICSLP
15. Georgila K, Wolters MK, Moore JD (2010) Learning dialogue strategies from older and younger simulated users. In: Proceedings of the 11th annual meeting of the special interest group on discourse and dialogue, SIGDIAĽ10. Association for Computational Linguistics, Stroudsburg, pp 103–106
16. Hone KS, Graham R (2000) Towards a tool for the subjective assessment of speech system interfaces (SASSI). Nat Lang Eng 6(3,4):287–303
17. Henderson J, Lemon O, Georgila K (2008) Hybrid reinforcement/supervised learning of dialogue policies from fixed data sets. Comput Linguist 34(4):487–511
18. Hester RK, Squires DD, Delaney HD (2005) The drinker's checkup: 12-month outcomes of a controlled clinical trial of a stand-alone software program for problem drinkers. J Subst Abuse Treat 28(2):159–169
19. Humeniuk R, Henry-Edwards S, Ali R, Poznyak V, Monteiro MG, World Health Organization et al (2010) The assist-linked brief intervention for hazardous and harmful substance use: a manual for use in primary care/prepared by R. Humeniuk et al. 2010
20. Jurčíček F, Thomson B, Young S (2012) Reinforcement learning for parameter estimation in statistical spoken dialogue systems. Comput Speech Lang 26(3):168–192
21. Kennedy CM, Powell J, Payne TH, Ainsworth J, Boyd A, Buchan I (2012) Active assistance technology for health-related behavior change: an interdisciplinary review. J Med Internet Res 14(3)
22. Lemon O, Pietquin O et al (2007) Machine learning for spoken dialogue systems. In: Proceedings of the European conference on speech communication and technologies (Interspeech'07), pp 2685–2688
23. Levin E, Levin A (2006) Evaluation of spoken dialogue technology for real-time health data collection. J Med Internet Res 8(4):e30
24. Levin E, Pieraccini R, Eckert W (1998) Using markov decision process for learning dialogue strategies. In: Proceedings of the 1998 IEEE international conference on acoustics, speech and signal processing, vol 1. IEEE Press, New York, pp 201–204
25. Li W, Glass J, Roy N, Teller S (2013) Probabilistic dialogue modeling for speech-enabled assistive technology. In: Proceedings of the fourth workshop on speech and language processing for assistive technologies. Association for Computational Linguistics, Grenoble, pp 67–72
26. Lisetti C, Amini R, Yasavur U, Rishe N (2013) I can help you change! an empathic virtual agent delivers behavior change health interventions. ACM TMIS 4(4):19
27. Miller WR, Wilbourne PL (2002) Mesa grande: a methodological analysis of clinical trials of treatments for alcohol use disorders. Addiction 97(3):265–277
28. Misu T, Georgila K, Leuski A, Traum D (2012) Reinforcement learning of question-answering dialogue policies for virtual museum guides. In: Proceedings of the 13th annual meeting of the special interest group on discourse and dialogue. Association for Computational Linguistics, pp 84–93

29. Mokdad AH, Marks JS, Stroup DF, Gerberding JL (2004) Actual causes of death in the united states, 2000. JAMA 291(10):1238–1245

30. Morbini F, DeVault D, Sagae K, Gerten J, Nazarian A, Traum D (2012) FLoReS: a forward looking, reward seeking, dialogue manager. In: 4th international workshop on spoken dialog systems, Paris

31. Morbini F, Forbell E, DeVault D, Sagae K, Traum DR, Rizzo AA (2012) A mixed-initiative conversational dialogue system for healthcare. In: Proceedings of the 13th annual meeting of the special interest group on discourse and dialogue, SIGDIAL'12. Association for Computational Linguistics, Stroudsburg, pp 137–139

32. Moyer A, Finney JW, Swearingen CE, Vergun P (2002) Brief interventions for alcohol problems: a meta-analytic review of controlled investigations in treatment-seeking and non-treatment-seeking populations. Addiction 97(3):279–292

33. NIAAA (2006) Niaaa alcohol alert no. 66: Brief interventions

34. NIAAA (2007) Helping patients who drink too much: a clinician's guide

35. NIAAA (2007) What colleges need to know now: an update on college drinking research

36. Paek T, Pieraccini R (2008) Automating spoken dialogue management design using machine learning: an industry perspective. Speech Commun 50(8):716–729

37. Papangelis A, Kouroupas N, Karkaletsis V, Makedon F (2012) An adaptive dialogue system with online dialogue policy learning. In: Proceedings of the 7th hellenic conference on artificial Intelligence: theories and applications, SETN'12. Springer, Berlin, pp 323–330

38. Portnoy DB, Scott-Sheldon LJ, Johnson BT, Carey MP (2008) Computer-delivered interventions for health promotion and behavioral risk reduction: a meta-analysis of 75 randomized controlled trials, 1988–2007. Prev Med 47(1):3–16

39. Riper H, Spek V, Boon B, Conijn B, Kramer J, Martin-Abello K, Smit F (2011) Effectiveness of e-self-help interventions for curbing adult problem drinking: a meta-analysis. J Med Internet Res 13(2):e42

40. Saitz R, Palfai TP, Freedner N, Winter MR, Macdonald A, Lu J, Ozonoff AL, Rosenbloom DL, Dejong W (2007) Screening and brief intervention online for college students: the ihealth study. Alcohol Alcohol 42(1):28–36

41. Schatzmann J, Georgila K, Young S (2005) Quantitative evaluation of user simulation techniques for spoken dialogue systems. In: 6th SIGdial workshop on DISCOURSE and DIALOGUE

42. Schulman D, Bickmore TW, Sidner CL (2011) An intelligent conversational agent for promoting long-term health behavior change using motivational interviewing. In: Association for the advancement of artificial intelligence (AAAI) spring symposium series. Association for the advancement of artificial intelligence, pp 61–64. http://www.aaai.org

43. Singh S, Kearns M, Litman DJ, Walker MA et al (2000) Empirical evaluation of a reinforcement learning spoken dialogue system. In: AAAI/IAAI, pp 645–651

44. Singh S, Litman D, Kearns M, Walker M (2002) Optimizing dialogue management with reinforcement learning: experiments with the NJFun system. J Artif Intell Res 16:105–133

45. Steinberg KL, Roffman RA, Carroll KM, McRee B, Babor TF, Miller M, Kadden R, Duresky D, Stephens R (2005) Brief counseling for marijuana dependence: a manual for treating adults. Center for substance abuse treatment, substance abuse and mental health services administration, Rockville

46. Sutton RS, Barto AG (1998) Reinforcement learning: an introduction, vol 1. Cambridge University Press, Cambridge

47. Sutton S, Cole R, De Villiers J, Schalkwyk J, Vermeulen P, Macon M, Yan Y, Kaiser E, Rundle B, Shobaki K et al (1998) Universal speech tools: the CSLU toolkit. In: Proceedings of the international conference on spoken language processing (ICSLP). Sydney, pp 3221–3224

48. Traum D, Larsson S (2003) The Information state approach to dialogue management. Kluwer Academic, New York, pp 325–353

49. Turunen M, Hakulinen J, Ståhl O, Gambäck B, Hansen P, de la Cámara R, Smith C, Charlton D, Cavazza M (2011) Multimodal and mobile conversational health and fitness companions. Comput Speech Lang 25(2):192–209

50. Walters ST, Hester RK, Chiauzzi E, Miller E (2005) Demon rum: high-tech solutions to an age-old problem. Alcohol Clin Exp Res 29(2):270–277

51. Walters ST, Neighbors C (2005) Feedback interventions for college alcohol misuse: what, why and for whom? Addict Behav 30(6):1168–1182

52. Ward W (1991) Understanding spontaneous speech: the phoenix system. In: Proceedings of the acoustics, speech, and signal processing. ICASSP-91, 1991 international conference, ICASSP'91. IEEE Computer Society, Washington, DC, pp 365–367

53. White A, Kavanagh D, Stallman H, Klein B, Kay-Lambkin F, Proudfoot J, Drennan J, Connor J, Baker A, Hines E, Young R (2010) Online alcohol interventions: a systematic review. J Med Internet Res 12(5):e62

54. Williams JD (2008) The best of both worlds: unifying conventional dialog systems and pomdps. In: INTERSPEECH, pp 1173–1176

55. Williams JD, Young S (2007) Partially observable markov decision processes for spoken dialog systems. Comput Speech Lang 21(2):393–422

56. Young S, Gasic M, Thomson B, Williams JD (2013) Pomdp-based statistical spoken dialog systems: a review. Proc. EEE 101(5):1160–9219, doi:10.1109/JPROC.2012.2225812

57. Young S, Gašić M, Keizer S, Mairesse F, Schatzmann J, Thomson B, Yu K (2010) The hidden information state model: a practical framework for pomdp-based spoken dialogue management. Comput Speech Lang 24(2):150–174

58. Young S, Gašić M, Keizer S, Mairesse F, Schatzmann J, Thomson B, Yu K (2010) The hidden information state model: a practical framework for pomdp-based spoken dialogue management. Comput Speech Lang 24(2):150–174