Building an On-Demand Avatar-Based Health Intervention for Behavior Change

Christine Lisetti, Ugan Yasavur, Claudia de Leon, Reza Amini, Napthali Rishe
School of Computing and Information Sciences
Florida International University
Miami, FL, 33199, USA

Abstract
We discuss the design and implementation of the prototype of an avatar-based health system aimed at providing people access to an effective behavior change intervention which can help them to find and cultivate motivation to change unhealthy lifestyles. An empathic Embodied Conversational Agent (ECA) delivers the intervention. The health dialog is directed by a computational model of Motivational Interviewing, a novel effective face-to-face patient-centered counseling style which respects an individual’s pace toward behavior change. Although conducted on a small sample size, results of a preliminary user study to assess users’ acceptance of the avatar counselor indicate that the system prototype is well accepted by 75% of users.

1 Introduction
Recent epidemics of behavioral issues such as excessive alcohol, tobacco, or drug use, overeating, and lack of exercise place people at risk of serious health problems (e.g. obesity, diabetes, cardiovascular troubles). Medicine and healthcare have therefore started to move toward finding ways of preventively promoting wellness rather than solely treating already established illness. Health promotion interventions aimed at helping people to change lifestyle are being deployed, but the epidemic nature of these problems calls for drastic measures to rapidly increase access to effective behavior change interventions for diverse populations.

Lifestyle changes toward good health often involve relatively simple behavioral changes that can easily be targeted (as opposed to major life crisis that usually require psychotherapy): for example lower alcohol consumption, adjust a diet, or implement an exercise program. One of the main obstacle to conquer change of unhealthy habits, is often ‘just’ a matter of finding enough motivation. There has recently been a growing effort in the health promotion sector to help people find motivation to change, and a number of successful evidence-based health promotion interventions have been designed.

Our current research interests have brought us to consider one such counseling style for motivating people to change – namely Motivational Interviewing (MI) (Miller and Rollnick 2002; Miller and Rose 2009). MI is a recently developed style of interpersonal communication which was originally conceived for clinicians in the field of addiction treatment, and which has been since then proven useful for medical and public health settings (Emmons and Rollnick 2001), as well as for concerned others. It is aimed at helping clients (or people in general) to find the motivation to change their lifestyle with respect to a specific unhealthy pattern of behavior. Although originally conceived as a preparation for further treatment, enhancing motivation and adherence, research revealed that change often occurred soon after one or two session of MI, without further treatment, relative to control groups receiving no counseling at all (Miller and Rollnick 2002).

At the same time as novel behavior change interventions have developed, simulated human characters have become increasingly common elements of user interfaces for a wide range of applications such as interactive learning environments, e-commerce and entertainment applications such as video games, virtual worlds, and interactive drama. Simulated characters that focus on dialog-based interactions are called Embodied Conversational Agents (ECAs). ECAs are digital systems created with a physical anthropomorphic representation (be it graphical or robotic), and capable of having a conversation (albeit still limited) with a human counterpart, using some artificial intelligence broadly referred to as an “agent” (Cassell and J. Sullivan, S. Prevost, & E. Churchill 2000). With their anthropomorphic features and capabilities, they provide computer users with a somewhat natural interface, in that they interact with humans using humans’ natural and innate communication modalities such as facial expressions, body language, speech, and natural language understanding (Magnenat-Thalmann, N. and Thalmann 2004; Predinger and Ishizuka 2004; Allen et al. 2006; Becker-Asano 2008; Boukricha and Wachsmuth 2011; Pelachaud 2009). We posit that given the well-defined structure of MI (discussed next), MI-based interventions can be automated and delivered with Embodied Conversational Agent (ECA). We describe the design and implementation of an MI-based automated intervention aimed at helping people with some excessive alcohol consumption to become aware and potentially motivated to change unhealthy patterns of behavior.
Related Research

Several avatar-based health intervention or counseling applications have been developed and evaluated. Silverman and colleagues coupled a simulation package with an animated pedagogical heart-shaped avatar to promote health behavior change in heart attacks (Silverman et al. 2001). Anthropomorphic pedagogical agents are used in a system called DESIA for psychosocial intervention deployed on hand-held computers in order to teach caregivers problem solving skills (Johnson, LaBore, and Chiu 2004). FitTrack uses an ECA to investigate the ability of an avatar-based system to establish and maintain a long-term working alliance with users (Bickmore and Picard 2005). The Virtual Hospital Discharge Nurse has been developed to explain written hospital discharge instructions to patients with low health literacy. It is available to patients on their hospital beds and helps to review material before discharging them from the hospital. It indicates any unresolved issues for the human nurse to clarify (Bickmore, Pfeifer, and Jack 2009).

More recently MI has been identified as particularly useful for health dialog systems (Lisetti 2008; Lisetti and Wagner 2008). Since then some researchers have explored the use of 2-D cartoon or comic-like characters to provide MI-related material to patients or to train health care personnel in MI (Schulman, Bickmore, and Sidner 2011; Magerko et al. 2011). Although the avatars employed in these mentioned systems have shown some promising acceptance by their users, because they are 2D, they lack dynamic expressiveness which is key for communicating facial affect (Ehrlich, Schiano, and Sheridan 2000), itself essential in establishing the MI requirement of empathic communicative style (Miller and Rollnick 2002).

Motivational Interviewing (MI) has been defined by Miller and Rollnick (Miller and Rollnick 2002) as a directive, client-centered counseling style for eliciting behavior change by helping clients to explore and resolve ambivalence. One of MI central goals is to magnify discrepancies that exist between someone’s goals and current behavior. MI basic tenets are that a) if there is no discrepancy, there is no motivation; b) one way to develop discrepancy is to become ambivalent; c) as discrepancy increases, ambivalence first intensifies; if discrepancy continues to grow, ambivalence can be resolved toward change. In the past few years, adaptations of motivational interviewing (AMIs) have mushroomed with the purpose to meet the need for motivational interventions within medical and health care settings (Burke, B. L. and Arkowitz, H. and Dunn 2002). In such settings, motivational issues in patient behavior are quite common, but practitioners may only have a few minutes to deal with them. Some of these AMIs can be as short as one 40-minute session.

Health Counselor System Architecture

We based our system on one such AMI referred to as the Check-Up that specifically targets excessive drinking behaviors and with which people can reduce their drinking by an average of 50% (Miller, W., Sovereign, R. and Krege 1988; Hester, R. K., Skire, D.D., and Delaney 2005). The check-up (along with modifications of it) is the most widely used brief motivational interviewing adaptation, where the client (usually alcohol or drug addicted) is given feedback, in a MI “style” based on individual results from standardized assessment measures (Burke, B. L. and Arkowitz, H. and Dunn 2002).

The On-demand Health Counselor System Architecture we developed can deliver personalized and tailored behavior change interventions using the MI communication style via multimodal verbal and non-verbal channels. The system is developed in the .NET framework as a three-tier architecture which uses third party components and services.

The system architecture is composed of the main modules shown in Figure 1, which we describe in details later. In short, during any interaction, the user’s utterances are processed by the Dialog Module which directs the MI sessions and elicits information from the user; the user’s facial expressions are captured and processed by the Affective Module, in an attempt to assess the user’s most probable affective states and convey an ongoing sense of empathy via a Multimodal Avatar-based Interface (using 3D animated virtual character with verbal and nonverbal communication); Psychometric Analysis is performed based on the information collected by the Dialog Module, and its results are stored in the User model which is maintained over multiple sessions to offer a dynamically Tailored Intervention in the form of sensitive feedback or specific behavior change plans.

Figure 1: On-demand Avatar-based Counselor Architecture.
Dialog Module
The Dialog Module evaluates and generates dialog utterances using three key components: a Dialog Planner, an MI-based Dialog Engine and a collection of Psychometric Instruments. The interaction of the client with the system is based on a series of dialog sessions, each of which having a specific assessment goal to identify different aspects of the user’s behavioral problem (if any) and in this article we focus on excessive intake of alcohol. Each session is based on psychometric instruments which the Dialog Planner uses as a flexible plan so that the dialog-based interaction can be dynamically conducted based on the knowledge that the system has acquired about the user (maintained in the user model), and on the user’s current entries and affective states.

Psychometric Instruments: We use a collection of well-validated Psychometric Instruments commonly used by therapists to assess an individual’s alcohol use in order to design each assessment session (Miller and Rollnick 2002). Each psychometric instrument contains a set of questions which represent the plan for an assessment session. For example one instrument (DrInC, 50 items) is used to identify consequences of the client’s drinking problem with five different parameters (e.g. interpersonal relationships, work) and these are stored in the user model (see below).

Dialog Planner: This component of the dialog module decides what type of utterance is to be generated based on the previous interaction in that session, i.e. it decides the next stage of the session. For instance, when the system needs to assess the client’s awareness of the consequences of his/her drinking, the The Drinker Inventory of Consequences (DrInC) psychometric instrument is used as the plan for the session, and each question is assigned one of the five topics of the DrInC (e.g. inter-personal, intra-personal). To measure consequences in each area, there are sets of non-sequentially positioned questions for each area and these questions are conceptually related with each other. The dialog planner aims at detecting discrepancies between received answers for questions in the same area. If the planner detects discrepancy in the client’s answers, then the MI-based Dialog Engine is invoked to generate some MI-style dialog, such as reflective listening (explained next). If no such discrepancy nor ambivalence in the user’s answers is detected, then the dialog planner continues with the assessment session.

MI-based Dialog Engine: This engine applies a finite set of well-documented MI techniques known under the acronym (OARS): Open-ended questions, Affirmations, Reflective listening, and Summaries, which are applied toward goals concerning specific behaviors (e.g. excessive alcohol use, drug use, overeating). The engine adapts each static psychometric instrument to MI style interaction and currently implements 1) affirmations (e.g.“You have much to be proud off”); 2) Summaries (“Ok, so let me summarize what you’ve told me today about your experience during the last 90 days. You have had more than 20 drinks, during the last 90 days you have missed work about once every two weeks, you have gotten into arguments with friends frequently at parties,...”); and 3) reflective listening.

Reflective Listening is a client-centered communication strategy involving two key steps: seeking to understand a speaker’s thoughts or feelings, then conveying the idea back to the speaker to confirm that the idea has been understood correctly (Rogers 1959). Reflective listening also includes mirroring the mood of the speaker, reflecting the emotional state with words and nonverbal communication (which we discuss in the Affect Module below). Reflective listening is an important MI technique which the system simulates to help the client notice discrepancies between his or her stated beliefs and values, and his or her current behavior. We currently implement simple reflection to elicit ambivalence (e.g. – Client: “I only have a couple of drinks a day.” – Counselor: “So you have about 14 drinks per week.”).

We also use Double-sided reflection to summarize ambivalence, which can help to challenge patients to work through their ambivalence about change. Statements such as “on the one hand, you like to drink and party, but on the other hand, you are then less likely to practice safe sex” can clearly summarize a patient’s ambivalence toward change (Botelho 2004). Other double-sided reflections such as “Ok, so your wife and children complain daily about your drinking but you do not feel that your ability to be a good parent is at stake, is that right?” can magnify the discrepancies between the user’s current situation and what they would like it to be.

Psychometric Analysis
As shown in Figure 1, psychometric analysis is concerned with processing Psychometric Data collected in the dialog module using the collection of psychometric instruments (described earlier). Based on instructions in each instrument, the Score Evaluator module calculates the score of the user for a particular measuring instrument, according to normative data (MATCH 1998). The result is stored in the user model (see next) in order to identify specific aspects of the drinking problem. For instance, the instrument assessing drinking patterns can locate the user on a spectrum of alcohol consumption (low, mild, medium, high, or very high, where a score above 8 indicates mild or higher problems). When provided feedback about this score (by the Tailored Feedback module), the user can make an informed decision about whether or not to continue with the offered self-help program. Respecting the MI spirit, the user is never coerced and can at any moment exit the intervention and the system gracefully terminates the conversation.

Tailored Intervention
User Model: The user model enables the system to tailor its intervention to a specific user. Tailoring can be achieved in multiple ways: by using the user’s name referred to as personalization, by using known characteristics of the user such as gender referred to as adaptiveness, and by using self-identified needs of the user referred to feedback-provision.

The user model collects the age, gender, weight, height and first name or nickname to preserve anonymity. It also gathers the user’s ethnicity to implement patient-physician
ethnic concordance with diverse avatars matching the ethnicity of the user because it is known that racial concordance increases the quality of patient-physician communication and some patient outcomes (e.g. patient positive affect) (Cooper et al. 2003). In addition to static demographic information, the user model builds and dynamically maintains the profile of each user as it changes over time. The user model includes different variables in six different areas: motivation, personal attributes, frequency of drinking, risk factors, consequences and dependence. Each area consists of multiple variables to assess specific aspects of the problem from different viewpoints. These psychometric data are calculated based on psychometric instruments (see above).

Because the intervention is designed to assist the user to change unhealthy behavioral patterns, it can be used over multiple sessions as on-demand follow-up sessions. As the user continues to interact with the system over time, the user model keeps track of the fluctuating user’s progress toward change in terms of: the user’s current level of ambivalence, by identifying and weighing positive and negative things about drinking; the user’s readiness to change, from pre-contemplation to maintenance (Prochaska and Velicer 1997); as well as the result of other specialized self-report questionnaires (e.g. the user’s awareness of the consequences of drinking).

**Tailored Feedback Engine:** This engine generates tailored feedback based on the user model. Several studies showed that tailored health communication and information can be much more effective than non-tailored one because it is considered more interesting and relevant, e.g. (Noar, Benac, and Harris 2007). Because the feedback module reveals discrepancies between the user’s desired state and present state, it is likely to raise resistance and defensiveness. The style in which the feedback is given therefore needs to be empathetic enough to avoid generating defensiveness. A number of pre-identified sub-dialog scripts are used for the various topics of feedback, e.g. self-reported drinking, perceived current and future consequences of drinking, pros and cons of current consumption patterns, personal goals and their relation to alcohol use, readiness to change drinking behavior. The user receives feedback about the assessed self-check information, which is put in perspective with respect to normative data (e.g. “according to the current national norm, 10% of people your age drink as much as you do”).

**Behavior Change Planner:** This component helps to create personalized behavior change plans for clients who are ready to change, and is only entered by users who have received feedback from the Tailored Feedback Engine. The first step of the behavior change planner is to assess the user’s readiness to change along a continuum from “Not at all ready” to “Unsure” to “Really ready to change”. Depending upon the level of readiness, different sub-modules are performed.

**Avatar-based Multimodal User Interface**

The user interface of the system is web-based with an embedded anthropomorphic Embodied Conversational Agent (ECA) or avatar, which can deliver verbal communication with automatic lip synchronization, as well as non-verbal communication cues (e.g. facial expressions).

**Avatar System:** We integrated a set of features that we consider necessary for health promotion interventions in our current prototype (using a third party avatar system called Haptek): 1) a 3-dimensional (3D) graphical avatar well-accepted by users as documented in an earlier study (Lisetti et al. 2004); 2) a selection of different avatars with different ethnicity (e.g. skin color, facial proportions) shown in Figure 3 (Nasoz, F. and Lisetti 2006); 3) a set of custom-made facial expressions capable of animating the 3D avatar with dynamical believable facial expressions(e.g. pleased, sad, annoyed shown in Figure 2) (Paleari, Grizard, and Lisetti 2007); 4) a Text-To-Speech (TTS) engine able to read text with a natural voice with lip synchronization; and 5) a set of mark-ups to control the avatar’s TTS vocal intonation and facial expressions in order to add to the agent’s believability (Predinger and Ishizuka 2004).

**Affective Module**

Our on-demand avatar counselor aims at delivering health promotion messages that are not only tailored to the user (via the user-model), but also empathetic. Because discussing issues about at-risk behaviors are highly emotional for people to talk about (e.g. embarrassment, discouragement, hopefulness), our context or domain application is particularly suited to inform the design of empathic characters. Empathy is considered the most important core condition in terms of promoting positive outcomes in psychotherapeutic contexts (Bellet and Maloney 1991) and it can account for as much as 25%–100% rate of improvement among patients (Miller and Rollnick 2002; Rogers 1959).

Currently, the **Empathy Engine** is limited to deciding which facial expressions the avatar needs to display. Our current system focusses on mimicking the user’s facial expressions, which is considered as a type of primitive empathy (Boukricha and Wachsnnth 2011). In the **Facial Processing** component, pictures of the user’s face are taken at fixed intervals during the user’s interaction with our avatar-based system, and these are analyzed and labeled with categorical emotion labels (neutral, sad, happy, surprised, and angry) using the face recognition engine from (Wolf, Hassner, and Taigman 2011). Since health counselors only carry out
partial mimicry to retain a positive overtone during the interaction, we model semi-mimicry so that the avatar displays a sad expression if the user’s expression is identified as sad, and smiling or neutral otherwise.

4 Avatar Evaluation Results

In a recent survey study, eleven individuals with different levels of technicality were asked to interact with the system and evaluate their experience based on functionality and on how much the users liked interacting with the avatar.

Specific results such as the willingness of individuals to discuss alcohol consumption with an embodied conversational agent (ECA) showed that 75% of users felt as comfortable or more comfortable interacting with the avatar counselor than they would with a human counselor: 37.5% felt as comfortable, 37.5% felt even more comfortable, and a remaining 25% said it was less comfortable than with a human.

Although the small group that was uncomfortable with the ECA argued that body language – currently lacking – was important to obtain feedback, the groups in favor pointed out that because the avatar could not judge nor embarrass them, they felt more positive and comfortable. These results are in line with research findings that patients that engage in behavior that can put them at risk (e.g. excessive drinking, unsafe sex, over-eating) tend to report more information to a computer interviewer than to its human counterpart (Servan-Schreiber 1986). The subjects in favor of the ECA interaction further pointed out the benefit of the system in terms of availability when human counselors are not present, or simply not needed at all times.

In a general context, the statistic of whether individuals would consider an ECA to discuss personal topics was 73% positive. According to 64% of individuals, the implementation of this system using animated avatars made the interaction engaging. This fact can become substantial to address high attrition rates found in textual computer-based health interventions (Cooper et al. 2003).

Lastly, modeling empathy for the patient was a key point when individuals were asked how they felt about the level of expression of emotions of the avatars. A 45% agreed on the suitability of the expression of emotions while a 36% said the avatar’s facial expressions were not intense enough. When the last group of individuals was asked why the level of expressions was not appropriate, they argued the need for more positive emotions such as smiles and body movement.

Although the number of subjects was small, these results seem to indicate that using avatars for health promoting interventions might lead to user’s higher engagement, lower attrition rates than in text-only computer-based health systems, and overall increased divulgence of sensitive issues due to increased confidentiality.

5 Conclusion

The demand for health promotion interventions that are tailored to an individual’s specific needs is real, and the technological advance that we described to deliver MI behavior change interventions with avatars an attempt to address that need. From our current promising results, we anticipate that avatar health counselors will be very well accepted by users once the field has matured from further research on the topic. We hope this research will contribute to promoting user’s engagement with health intervention program, and, ideally, increases positive outcomes.

6 Acknowledgements

We would would like to thank Eric Wagner for introducing us to MI and how it is used in his NIAAA-funded adolescent behaviors and lifestyle evaluation (ABLE) program. This research was supported in part by NSF grants HRD-0833093.

References


Approach.


