A Tool to Promote Prolonged Engagement in Art Therapy: Design and Development from Arts Therapist Requirements

Jesse Hoey*
School of Computer Science
University of Waterloo
Waterloo, Canada, N2L 3G1
jhoey@cs.uwaterloo.ca

Krists Zutis
School of Computing
University of Dundee
Dundee, Scotland, DD1 4HN
kristszutis@computing.dundee.ac.uk

Valerie Leuty
Dept. of Occupational Science and Occupational Therapy
University of Toronto, Canada, M5G 1V7
valerie.leuty@utoronto.ca

Alex Mihailidis
Dept. of Occupational Science and Occupational Therapy
University of Toronto, Canada, M5G 1V7
alex.mihailidis@utoronto.ca

ABSTRACT

This paper describes the development of a tool that assists arts therapists working with older adults with dementia. Participation in creative activities is becoming accepted as a method for improving quality of life. This paper presents the design of a novel tool to increase the capacity of creative arts therapists to engage cognitively impaired older adults in creative activities. The tool is a creative arts touch-screen interface that presents a user with activities such as painting, drawing, or collage. It was developed with a user-centered design methodology in collaboration with a group of creative arts therapists. The tool is customizable by therapists, allowing them to design and build personalized therapeutic/goal-oriented creative activities for each client. In this paper, we evaluate the acceptability of the tool by arts therapists (our primary user group). We perform this evaluation qualitatively with a set of one-on-one interviews with arts therapists who work specifically with persons with dementia. We show how their responses during interviews support the idea of a customizable assistance tool. We evaluate the tool in simulation by showing a number of examples, and by demonstrating customizable components.

Categories and Subject Descriptors

I.2.1 [**Computing Methodologies**]: Artificial Intelligence—*Applications and Expert Systems*; H.5.2 [**Information Systems**]: Information Interfaces and Presentation—*User Interfaces*

General Terms

Theory, Design, Human Factors

Keywords

Markov decision process, user modeling, computer vision, art therapy, dementia

*This work was completed while the first author was with the School of Computing, University of Dundee, Scotland

ASSETS'10, October 25–27, 2010, Orlando, Florida, USA.

1. INTRODUCTION

A growing area of activity in health technology is support systems for persons with cognitive or physical disabilities, who want to continue to live independently in their own homes. Such systems are typically engineered for a certain task to provide guidance, assistance, or emergency response. However, this approach is labour intensive, and the resulting systems tend to have no capacity to adapt over time or to different users or tasks. In this paper, we discuss an approach to this problem: a customizable, ubiquitous modeling technique that can adapt to users over time. The idea to is allow non-technical persons, such as caregivers, or therapists, to customize applications for their clients.

To allow for customization and adaptivity, we propose an abstract model of autonomous assistance based on the partially observable Markov decision process, or POMDP. The POMDP uses decision theory to reason about what actions the *agent* can take to optimize over a user-specified utility function. The POMDP can explicitly deal with uncertain effects of actions, and uncertainty in sensing.

This paper considers one class of tasks where autonomous, customizable assistance is paramount: assisting persons with dementia and their therapists during art therapy sessions. There is increasing evidence that leisurely activities decrease dementia risk [6], and that engagement with visual artworks has benefits for the promotion of quality of life in older adults [16, 3]. However, many older adults have difficulty motivating themselves to engage in a creative activity for a reasonable period of time. These difficulties are compounded when the older adult has a progressive illness, such as Alzheimer's disease (AD). The cognitive difficulties that characterize dementia (the most common form of which is AD) include trouble following instructions, remembering steps in a process, staying engaged, and making choices.

Arts therapists primarily work in residential, hospital and day care settings, although exceptions exist [9]. Individuals remaining in their own homes spend long periods with no occupation, as carers are often busy with daily routines. These periods reduce the ability to engage with the creative process, and can result in the person lacking motivation and desire to participate in independent activities [17]. While this engagement can be provided by a dedicated therapist, there is a lack of such therapists to support the increasing number of persons with a progressive illness and who are remaining in the home. Perhaps more importantly, a large benefit of engaging elderly persons with the arts is to enable them to do so independently.

The contributions of this paper are twofold. First, we demonstrate an abstract model of automated assistance based on the partially observable Markov decision process, and show how this model can be applied to art therapy for dementia. Second, we show, through a user-centered design approach culminating in a set of interviews with creative arts therapists, how it is possible to give non-technical users control over an intelligent assistant. Trials are planned for fall 2010 with the same group of arts therapists and their clients (persons with dementia). This paper describes a key step in the design process leading to these trials. It is critically important to complete this step prior to client testing, as obtaining useful feedback from persons with dementia directly is challenging, and must be approached with care.

This paper overviews published results from a survey and focus group of arts therapists, followed by a definition of an abstract model of assistance. Section 4 presents a prototype device, Section 5 gives results of interviews and simulations, and Section 6 describes related work.

2. SURVEY AND FOCUS GROUP

A survey of 133 practicing arts therapists was conducted in the early spring of 2009. The respondents were primarily from Canada and the UK, and represented a broad spectrum of specialty areas (e.g. mental illness, disabilities, dementia), specialty populations (e.g. children, older adults), and preferred techniques (e.g. visual art, music, etc.). There were 46 respondents (35%) who specifically worked with older adults who have dementia. Key themes that emerged from a design ethnographic analysis of the survey results were elaborated on in a focus group with seven arts therapists held at a Toronto-based long-term care facility in April 2009 [10]. The following main themes and concepts were extracted:

There was **enthusiasm** for technology in art therapy, and **adaptivity** to match a client's current abilities was considered a key element. **Inertia** can be provided by the device to help with loss of focus. Clients often cannot maintain their own engagement for even a small length of time while a therapist is attending to another client. An automated device can assist by providing help to keep a client engaged, thus serving as a therapist's ***inertial tool***. **Customizability** of the devices is very important. Therapists must control the client application, both from an artistic standpoint (they have novel ideas for client activities based on their experiences with a particular client), and from a technical standpoint (to ensure safety and reliability of the interaction between client and device). Therapists require the ability to customize in order to help users feel more comfortable with the software and to tailor the application to a user's needs and abilities. The customization must also allow for the differing practices of arts therapists. **Engagement** tracking is a key element for the device, as this is primarily what the arts therapists attempt to optimize. Therapists use a variety of techniques to track engagement and mood, but primary ones are facial expression and gaze.

3. AUTOMATED ASSISTANCE

Key to our approach is the ability for a therapist to customize an *adaptive* application for a particular client. The adaptivity is provided by a sophisticated AI back-end (see Section 4.2). We do not want to require therapists to deal with these technical aspects, but we want to give them access to the functionality. Therefore, in order to allow our users to customize the devices, we must provide methods for therapists to indirectly change the definition of the POMDP controller. Our system allows them to do this by creating a separation between *generic* control algorithms that work for any application, and *specific* control systems that interact with a user and a particular customized interface. This separation requires us to create an abstraction of the task of assistance within an art therapy application.

We start with the primary goal of the device, which is to maintain a person's *engagement*, defined as the level of attention of a user focused specifically on an activity.

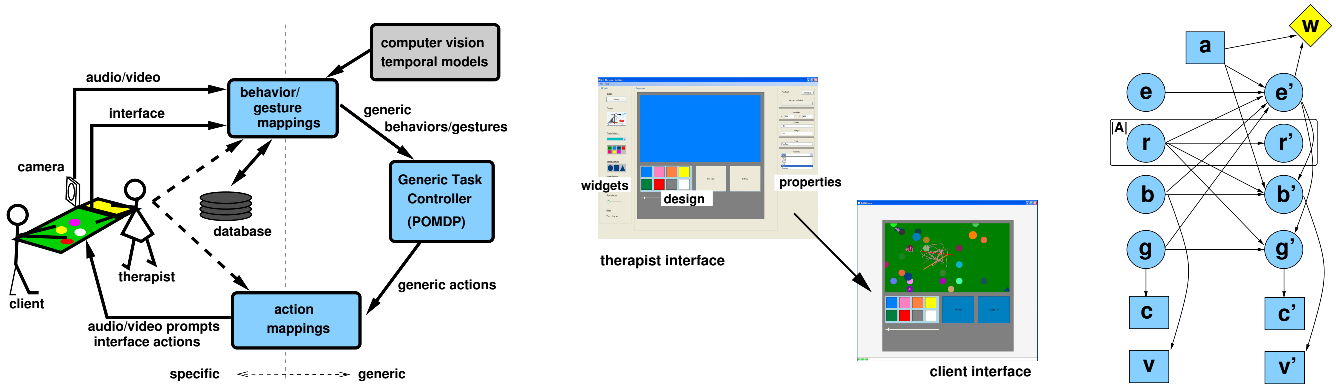
We define categories of user *behaviors* based on the influence of *engagement* on those behaviors. We define the *involvement* of a behavior as the amount of engagement it requires. For example, we can define behaviors as being *interactive* if they require (are indicative of) high engagement, or *inactive* for situations of low engagement.

Finally, we define categories of system actions (e.g. prompts) based on their influence on *engagement*, and their cost to the user in terms of interruptions and disturbances. We define the *interactivity* of an action as the amount of involvement it requires from a user. For example, a system that pops up a dialog box requiring input is highly interactive, whereas a change to an icon in a corner is mildly interactive. The trade-off is that a very interactive prompt may get a disengaged user involved, but may be a disruptive action for an already engaged user, causing them to disengage. We now describe how these concepts can be integrated into a decision theoretic model.

4. PROTOTYPE TOOL DEVELOPMENT

4.1 Overview

Figure 1(a) shows a schematic of the system. On the left, we see a user doing some visual art-work on a horizontally mounted screen. A therapist is present, and has a separate interface that could be on a mobile device. The outputs of the device are a video stream from a web-camera watching the user's face, and the interface actions (finger movements on the screen). These are passed to a therapist-defined set of behavior/gesture mappings, that map each behavior of a user on the screen to a category of *involvement*. The generic controller (Section 4.2) then uses this to return a



(a) (b) (c)

Figure 1: (a) Schematic of the device, showing the separation between specific (left) and generic (right) components; (b) Therapist interface and final client application; (c) The POMDP model as a 2-time slice decision network. Observed variables are shown with a rectangular box. Variables are (e)ngaged, (r)espond, (b)ehaviour, (g)esture, a(c)tivity, (v)ideo, system (a)ction, and re(w)ard.

generic action (an *interactivity* level) estimated to be best at maintaining the client's *engagement*. This is returned to the application using a set of output mappings as a particular action for the system to take at that level of interactivity.

The therapist uses the interface shown in Figure 1(b) on the left to design the layout of the interface, to configure the features (functions) that are available for the user, and to configuring the system actions. Once done, the therapist exports the newly created application for use by the client (Figure 1(b) on the right). The therapist has control over the interface itself, the behavior/gesture mappings, and the action mappings (Section 4.3). The therapist can also consult logs of previous experiences in the database, and can use a set of canned computer vision algorithms or temporal models while defining the behavior/gesture mappings.

4.2 POMDP Controller

The tool uses a probabilistic, decision theoretic model: a partially observable Markov decision process (POMDP) consisting of a finite set S of states; a finite set A of actions; a stochastic transition model $\Pr : S \times A \rightarrow \Delta(S)$, with $\Pr(t|s,a)$ denoting the probability of moving from state s to t when action a is taken, and $\Delta(S)$ is a distribution over S; a finite observation set O; a stochastic observation model with $\Pr(o|s)$ denoting the probability of making observation o while the system is in state s; and a reward assigning $R(s,a,t)$ to state transition s to t induced by action a.

The system actions cause stochastic state transitions, with different transitions being more or less rewarding (reflecting the relative utility of the states and actions). States cannot be observed exactly. Instead, the stochastic observation model relates observable signals to the underlying state. The POMDP can be used to monitor beliefs about the system state using standard Bayesian tracking/filtering. Finally, a *policy* can be computed that maps *belief states* (i.e., distributions over S) into choices of actions, such that the expected discounted sum of rewards is (approximately) maximized.

The generic assistance controller is a class of POMDP models, as shown as a Bayesian decision network in Figure 1(c). We are using a factored POMDP representation in which the state space is represented as the cross product of a set of variables. The actions the system can take are to do nothing (a_0), or to perform some intervention ($a_1, \ldots, a_M$), where intervention a_i has interactivity level $i \in 1 \ldots M$.

The state space contains four factors or variables. Two of these factors relate to the task.

1. The **behavior**$\in\{$*interactive,active,intermittent,inactive*$\}$ is an interface involvement level inferred from observations of their finger interactions (**activity** observation). The observation model $P($**activity**$|$**behavior**$)$ defines the measurement noise.
2. The **gesture** variable is a set of gestures that indicate that a user is engaged with the device, and are inferred from the video stream. For example, the **gestures** could be gaze directions ($\{$*looking,not looking*$\}$), indicating if a person is looking at the screen. The **gestures** and the observation model $P($**video**$|$**gesture**$)$ are activity dependent, although for many applications the gaze direction will be sufficient.

The other two factors relate to internal, affective or mental states of the user and define a user model:

1. The user's **engagement**$\in\{$*yes,confused,no*$\}$ is the key element of this model. A user can be engaged (*yes*), or disengaged partially (*confused*) or completely (*no*). The dynamics of engagement, and it's effect on behaviors, gestures, and task are user and activity dependent.
2. The user's responsiveness to the system's actions, **respond**, is factored into a set of variables giving responsiveness to each of the system's action interactivity levels.

This user model defines a range of user types (e.g. responsive/engaged). For each user type, the model dynamics will give rise to a different strategy on the part of the device, allowing for adaptivity as the user interacts. Further, monitoring the device's *belief* in the user's state allows for a quantitative measure of a user's engagement and responsiveness that can be used for assessment.

Note that the precise discretization of each variable (e.g. the engagement in three levels), and the labels attached to these levels is not important to the monitoring/assistance system. The POMDP tracks a *belief* state, which is a continuous measure of a person's engagement, as inferred from their interactions on the screen. The discrete levels allow us to more easily represent this belief state on a machine, but the actual definition of, for example, engagement is only through the interactions between variables, which are ideally customizable by therapists.

The dynamics of the POMDP hinges on the user's **engagement**. The user's engagement changes dynamically over time as a function of the system's actions, and their previous behaviors. For example, if the user is disengaged, but is looking at the screen, and the system does something of interest, then the user may become engaged with some probability. On the other hand, if a user is already engaged, and the system gives a prompt or changes the interface, the user may become confused. The user's responsiveness comes into play when the system takes an action. If they are responsive to the interactivity level of the action, the effect of the action is to increase their engagement. Otherwise, the system action has little or no effect.

The reward function is based solely on the user's engagement, with +10,-6,-2 if the user is engaged, confused or not engaged, respectively. System actions are slightly costly (cost of 0.5), but only if the user is engaged. This models the effect of an action reducing feelings of independence in a user if they are already engaged. Note that this cost is separate from the indirect costs incurred if a user is prompted when engaged, possibly leading to confusion. Costs/Rewards are arbitrary up to a linear scale factor, and so are chosen to reflect approximately the relative value of prompts, but would ideally be elicited from therapists.

To compute an approximate policy, we used the SymbolicPerseus package[1]. It implements a factored, structured point-based approximate solution technique based on the Perseus algorithm [18]. The POMDP controller thus computed takes actions that lead to behaviors defined as more *involved* by the therapist. The POMDP is optimising over the user's *engagement*, but this is reflected in the amount of *involvement* of the behaviors. However, the POMDP must trade-off *involvement* and *engagement* against *interactivity*.

4.3 Application Interface

The designer tool presents the therapist with a screen split into three sections: *widgets*, *design*, and *properties* (see Figure 1(b)). The therapist can drag any widget onto the design area and then configure it to their needs with a mouse (e.g. resize with a click-and-drag). The properties screen allows configuration of additional options, such as background colour and specific options for each widget. For example, the button widget can optionally be made a *line*, *pen*, or *smudge* button. Other widget choices include colour palettes, shape selectors, and an image selector that allows personalized images to be added by the client.

The designer application has an *action manager* window, in which the therapist selects and customizes system actions. These actions range from adding shapes or images to the canvas, to animating buttons, or playing audio files. Action customizations are, for example, the sizes and locations of shapes that will be added by the system. The therapist can record or upload audio prompts and can change the *interactivity level* $\in$ {*low, high, stimulate*} for each action. An example of a configuration would be to set the interactivity level of *add circles* to *low*, and *play audio prompt* to *stimulate*. The application is then exported as a set of XML files.

The tool uses computer vision algorithms to detect if a client is looking at the screen. Images are taken from a standard web camera, and a Haar-like feature recognition method is used to detect a single face in the scene. Haar-like feature detectors extract features from a rectangle in the scene, and compare these features to a pre-trained database of known valid features (trained classifiers). We used an implementation in the OpenCV library [1] that has been trained on a wide variety of human faces. When a valid face has been found in the scene, a second Haar-like feature detector is applied to the detected face region in order to detect a pair of eyes. Only faces with eyes are reported as correct detections.Examples are shown in Figure 3.

User behaviors on the screen, also indicators of engagement, are classified into levels of *involvement* $\in$ {interactive, active, intermittent, inactive}. A user who is interacting with a recently added shape is *interactive*. Otherwise, the user behavior is categorized as *inactive*, *intermittent*, or *active* depending on the amount of time the user spends touching the screen over a five second interval.

5. EVALUATION AND RESULTS

5.1 Simulated Examples

Figure 2: An example sequence. Behavior can be interactive (I), active (A), intermittent (M), and inactive (N). Engaged can be no (N), confused (C), and yes (Y). LowRespond can be yes (Y) or no (N). The face detector output and the user interface can also be seen. The generic (POMDP) and specific (configured by therapist) actions are along the bottom. The marginal belief states are shown.

We now present two demonstrations in the laboratory with a subject without dementia acting according to defined situations. These are meant only to illustrate the functionality of the device and user monitoring, not as end-user

[1]code available at www.cs.uwaterloo.ca/∼ppoupart/software

tests, and so can only be evaluated qualitatively by seeing if the system beahviours look reasonable given the context. Figures 2 and 3 are samples from the interactions, showing the face detection, the user interface (drawing canvas on the touch screen), the user's action, the POMDP interpretation of the user's behavior, the POMDP estimate of the user's engagement, and responsiveness or gesture, as well as the generic POMDP action (interactivity level) and the specific action actually performed. The POMDP variables are shown as marginal beliefs, in which the bars represent the probability the user is in that state, irrespectively of all other variables. Figure 2 shows the simulated client responding to a low interactivity prompt at T=5 (behavior=(I)nteractive) that results in an increase in the system's estimate his responsiveness to such prompts. The client is engaged for the next 10 seconds, then stops all on-screen activity after T=10. When the system detected the user disengaging (Engaged=N at T=15 and 20), the POMDP attempted a "High" interactivity prompt. During the design phase, the therapist mapped the "Animate Button" system action to the "High" interactivity POMDP prompt, thus, a button on the user's interface was animated. Figure 3 shows a simulated client looking away from the screen (Gesture=NL at T=35) that results in a descrease in engagement, and the system gives a "Stimulate" interactivity prompt to get his attention.

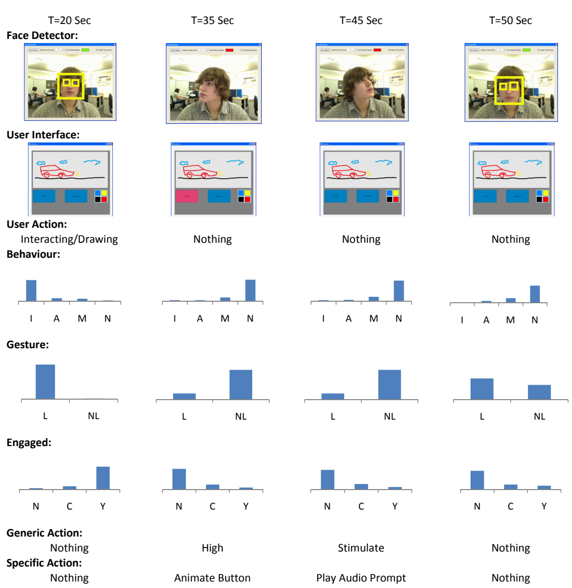


Figure 3: An example sequence of a 50 second interaction in which the user looks away. Gesture can be looking (L) and not looking (NL).

5.2 Arts Therapists Interviews

5.2.1 Procedure and Aims

The aim of interviews with arts therapists was to evaluate the current prototype and to explore the changes that were needed. Participants were arts therapists from Sunnybrook Hospital recruited via email due to their participation in the focus group. For those who agreed to participate, a follow-up email was sent to set up a meeting time. The interview time and location were provided at this time, and participants were given the opportunity to ask questions about the nature of the study beforehand. This process continued until six art therapists and one music therapist were recruited.

Ethics approval was obtained through the University of Toronto Research Ethics Board. At the time of each interview, the investigators introduced their roles, outlined the purpose of the study, and obtained verbal consent to proceed. All interviews were video recorded. Each interview began with a 15-20 minute demonstration of the system described above, and was followed by a semi-structured interview, approximately 40 minutes in length. A Dell Precision M65 laptop, a Logitech Quick Cam Pro 9000 webcam, and a 20" touch screen were used. A Nokia N95 mobile phone was also used to demonstrate a portable device web interface feature. Participants were asked about software features that should be added, removed, or modified, from either the client or therapist interface, and were asked to describe typical engagement patterns of their clients. If time permitted, extra features, such as the potential for multi-touch screens and an online interface, were explored. Any questions raised by the participants were addressed by the researchers.

All interviews (n=7) were transcribed verbatim. Videos and notes were reviewed in order to identify both the content and context of interview data. A coding scheme was developed based on a modified version of the direct content analysis of [5], in the context of categories predetermined from the survey and focus group results. The scheme was used to extract key themes from the transcripts. Information from the focus groups in phase one of the project was used to begin identifying key concepts in the current transcripts. This included any feature of the current software which participants felt should be added, removed, or modified. Interview questions were targeted towards these predetermined categories. A coding scheme of key themes was developed in the context of these predetermined categories, based on an analysis of two individually analysed and randomly selected transcripts. Peer debriefing and member checking were used in order to explore themes that may have been missed and to establish the validity of the results among the participants who offered up the original data.

A brief overview and description of the project were given to the therapists, and the current version of the system was demonstrated. Each therapist was given a description of the dual application design, and the therapist's role in creating the client's art therapy application. A full demonstration of the software was then given, showing the application creation stage, as well as launching and interacting with the created application. The participant was also invited to test the system and experiment with the available functions. The participant was then asked questions about the current system, and the changes and additions needed. The therapists were also asked how they would implement certain functions to best fit the therapist's needs. The interview process was concluded by asking for further feedback about the system.

The majority of the data collected was extracted from the interview process, however, the process of observing the therapist's interactions with the prototype system also gave useful insight into some usability, hardware, and design aspects of the system. Additionally, further insight into the environment that the software would be used, and general insight into Alzheimer's, the elderly, and art therapy was gained by observing an actual Art Therapy group session.

5.2.2 Qualitative Feedback

The majority of the feedback received from the therapists was very positive. The therapists were happy with the dual application design, and thought that the system would be easy to use with little training. Most therapists also remarked that the system represented a new and exciting tool for art therapy. Two significant themes were related to the customization of the system actions and the engagement monitoring, along with demands for new interface features.

A number of therapists responded positively to the customizable intelligent prompting process. One responded *"Oh, you choose how it's going to prompt them! I see, that's kind of neat,"* and commented *"that seems pretty logical."* Another therapist, while discussing audio prompt customization, also seemed satisfied with the system, and remarked *"I like how customizable that is, it's a great idea!"* Three out of seven participants expressed concern regarding the permanency of the prompts added to the client's canvas. If a client is not interacting with the canvas he/she is not necessarily disengaged, and adding an image or shape onto this client's art may be upsetting. Providing the option to accept or decline prompts was suggested. *"One thing that might be handy, is that when a line [prompt] is broadcast suddenly up there as an attention grabbing thing, that it could be a choice. Maybe it would flash or something like that and they could hit it if they wanted to keep it."*

Prompting the client in ways that do not disrupt his/her canvas directly was also suggested. *"I guess I have a problem with, if they are creating this and doing it, I would not want to tamper with their image, I would want to respect that. If there was a way it [the system prompts] could do something with the screen around it, perhaps change the background colour of the sketchbook..."* A general eraser or undo button was also mentioned by two out of seven participants.

When asked how to better classify between generic POMDP decisions and specific actions, one therapist stated that *"you've already built in the choice of when to use them, and assigning Low, Medium and High to them - that seems appropriate"*, and then added that *"it's fairly self explanatory; seems comfortable"*. The therapists showed clear approval of mapping the generic interactivity levels to the specific application prompts, such as adding shapes or playing sounds.

While the therapists were satisfied with the ability to separate the generic POMDP decisions with the specific system actions, they noted that more control would be useful. Particularly, they wished to have control of the AI back end during the therapy session. The therapists noted that they would like to have the ability to pause the AI decision making process or experiment with different settings throughout a single session, possibly using the portable device interface.

The second major theme was related to the accurate measurement of client engagement. If a client is engaged with the art therapy device, the current system assumes that he/she will be interacting with the device physically and that his/her face will be looking towards the screen. All seven participants were prompted to discuss how to accurately measure client engagement. The interviews showed differences over whether it was useful to have the computer measuring engagement or whether this should be up to the arts therapist. Therapists reported that engagement varied across different users, and that just because a person is not interacting with the screen, they may not be disengaged, *"they may be thinking, or if they have a disease such as Parkinson's, they may just be frozen."* Engagement is very individualized: *"one person looking at the screen may be very engaged, while another may not be at all."* Three out of seven participants explicitly stated that they would be uncomfortable trusting a computer to accurately measure engagement, and all participants agreed that the system could not possibly pick up on all the exceptions to the set levels of disengagement. *"Measuring engagement is all about timing and being aware and present with the resident. It is more intimate than the computer could measure."* One therapist agreed that being able to control the intelligent assistive system by configuring engagement settings during the design phase would further improve the system. This indicates that the customization of the behavior mappings is an important element to be improved upon. It is also possible to allow for customization of the POMDP that describes what it means to be engaged (see Section 5.3).

A feature that four out of seven participants agreed might be a more accurate way of measuring engagement was eye tracking. Currently, the system detects the face and the eyes, but the movements of the eyes are not recorded.

5.2.3 Hardware Requests

A common request among the therapists was for multi-touch functionality. Currently, the single-touch system allows the elderly user to interact with the computer in a fairly intuitive and simple manner. However, the therapists believed that adding an even more intuitive interfacing system would allow the clients to use the system with much more ease and less anxiety. Also, another application of a multi-touch system is to program the ability to compensate for accidental hand gestures.

A common request included the addition of other hardware, such as pens, paintbrushes, or other realistic art tools. By using real art tools rather than a finger, some of the therapists believed that the level of interaction and realism can be increased. Another feature that was suggested was for a hardware system that tracks the elderly user's biological data, such as heart rate, temperature, and skin resistance.

5.3 Customizability Simulations

Some therapists were concerned that the specific methods for assistance that had been offered were too narrowly defined. However, the intelligent assistance model we have proposed is inherently customizable, and in this section, we demonstrate a set of simple *tuning knobs* for the system that will allow it to be further adapted to particular users "on the fly." We focus in particular on three aspects of the interaction: the passivity of the system, the activity of the client, and the eye-contact of the client.

The passivity of the system is the simplest adaptation and governs overall how much prompting the system will be doing. The client-centered aspects are more subtle. As uncovered in the interviews, therapists indicated that clients will externalize their engagement at different levels. Some will spend more time while engaged looking at the screen and interacting, whilst others will spend more time looking away and not interacting (possibly considering, thinking, or engaging with their art-work in non-interactive ways). We therefore define, for each simulated client, an eye-contact level l and an activity level a. A larger value of $\{l, a\}$ means the simulated client spends more time {looking at,interacting with} the screen when they are engaged, respectively.

To this end, we implemented a simple simulator that maintains a level of engagement at each time step using the following update equation

$$e[t] = draw(e[t-1] * (1-\beta) + (1-e[t-1]) * r[a]) \quad (1)$$

where $e[t]$ is the engagement at time t, $\beta = 0.2$ is a *disengagement* rate, and $r[a]$ is the level of responsiveness to prompt interactivity level a, set to $\{0.45, 0.3, 0.2, 0.05\}$ for $a = \{high, med, low, none\}$, and $draw(x)$ is a random binary draw (a coin flip giving 0 or 1) from a binomial $\{x, 1-x\}$. The simulator then generates a gesture$\in\{$*looking,not looking*$\}$ and a behavior$\in\{$*interactive,active,intermittent,inactive* based on the eye-contact level l and an activity level a:

$$gesture = draw(e * p_c^+[l] + (1-e) * p_c^-[l]) \quad (2)$$

$$behavior = draw(e * p_y^+[a,b] + (1-e) * p_y^-[a,b], b) \quad (3)$$

where $p_c^+[l], p_c^-[l]$ gives the probability the simulated client is looking at the screen if they are/aren't engaged, respectively, when they have eye-contact level l, and $p_y^+[a,b], p_y^-[a,b]$ is the probability the simulated client has behavior b if they are/aren't engaged, respectively, when they have activity level a. These probabilities are chosen to cover a broad range of interaction styles. $draw(\vec{x}, b)$ is a random draw from the distribution x (after normalization) over variable b (e.g. a roll of a weighted |behavior|-sided dice). The simulator is executed for 50 runs of 100 steps, and the percentages of each type of prompt is averaged over the runs.

We use two types of POMDP model: adapted and non-adapted. Non-adapted models are the standard model (Section 4.2), while the adapted models have modified dynamics or rewards based on the characteristics of the simulated clients. Figure 4 shows the percentage of interactions in which the various levels of prompt are given as a function of the passivity, eye-contact and activity levels for the adapted (left column) and non-adapted (right column) models.

The system passivity is adjusted by changing the cost of the prompts. The resulting policies then have a variable prompting rate, as shown in Figure 4(a-left), where the percentage of interactions in which a prompt is given is shown as a function of the passivity level. The passivity level thus provides a *tuning knob* that allows the therapist to select a more or less passive system during a simulated client's interaction.

The simulated activity level is modeled as a multiplicative factor $\alpha(a)$ that changes the dynamics of the behavior it more active (more behaviors$\in\{$*interactive,active,intermittent*$\}$ when engaged). The eye-contact level is modeled as a multiplicative factor $\eta(a)$ that changes the dynamics of the gesture by making the simulator look at the screen more when they are engaged. Figure 4(b) and (c) show the percentage of prompts for activity and eye-contact levels, respectively, for adapted (left column) and non-adapted (right column) models. We can see that, as the activity level decreases, the non-adapted model starts making more highly interactive interventions, since it believes the simulated client has lost engagement. The adapted model, on the other hand, correctly infers that this decrease in activity is due to the user's activity level, and decreases it's overall level of prompting and its interactivity. Similarly, as the eye-contact level decreases, the adapted system decreases the amount of prompts it is giving, whereas the non-adapted system does not.

We have therefore described three *tuning knobs* that can

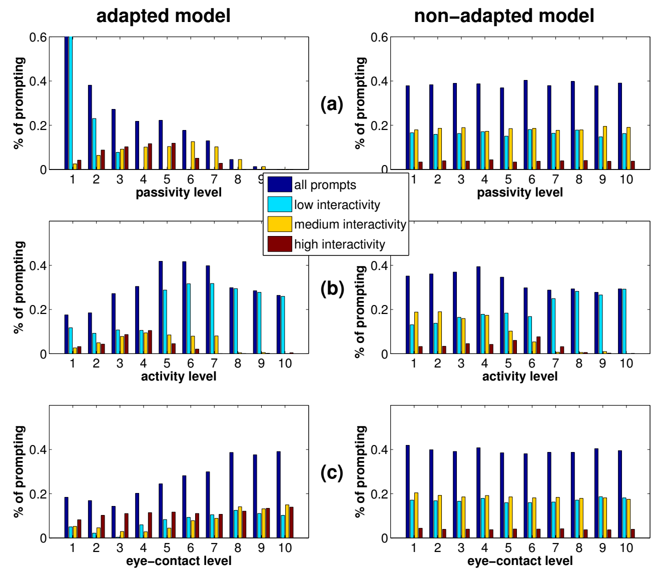


Figure 4: Percentage of prompting as a function of (a) the passivity level of the system, (b) the activity level of the client, and (c) the eye-contact level of the client for adapted (left column) and non-adapted (right column) models.

be used by therapists to adapt the interaction style of the system on the fly. These *tuning knobs* correspond roughly to the user types therapists' have indicated to be of most interest during the interviews.

6. RELATED WORK

The system we describe is similar to the COACH for handwashing assistance [11], which uses an overhead camera to monitor a user and a POMDP to estimate the user's progress. Audio-visual cues are delivered to assist the user in completing the task. The main difference from the decision making perspective is that handwashing is a very structured task, with only few ways to accomplish it, and with goals based on physical outcomes (e.g. hands clean), as well as on user states (e.g. user independence). The creative arts task, on the other hand, is very weakly structured, with goals depending only on user affective states (e.g. engagement).

There are several other intelligent systems currently being developed for the older adult population. These include the Aware Home Project [13] and the Assisted Cognition Project [7]. These projects are similar to ours in that they incorporate AI and a decision-theoretic approach. The Autominder System [14], one aspect of the Nursebot Project, applies a POMDP in the planning and scheduling aspects of the system. However, these systems are mainly used as scheduling and memory aids, and do not incorporate user attitude modelling or planning for prompting.

POMDPs provide a rich framework for planning under uncertainty and can be used to robustly optimize the course of action of complex systems despite incomplete state information due to poor or noisy sensors. For instance, in mobile robotics [14], and vision-based systems for assistive technology [11], POMDPs can be used to optimize controllers that rely on the noisy information provided by sensors.

There has been some research published on the use of technology in art therapy contexts. For example, investigations on the use of video [4] and digital photography [19] as new tools in psychotherapy have yielded very positive results. With respect to computer-based art therapy (i.e. using technology as part of the art therapy process), only a limited number of technologies have been developed. Examples include Jabberstamp [15], which is a tool that allows children to add voices to their drawings. However, these applications are not practical for older adults with dementia because of required high levels of cognition and fine motor skills. Collie and Cubranic (1999) developed a computer based art therapy program for use in tele-health [2], but their system required an arts therapist to be physically present, and was not developed or tested with older adults with dementia. The majority of past computer-based applications have focussed on art therapy assessment, such as for judging the main colour in a drawing [8].Touchscreen devices for visual artworks have recently been developed [15, 12], however not for persons with dementia, and not as a tool for arts therapists specifically.

To date there has been no work on computer-based systems that can prompt and monitor a user's participation in creative arts akin to the ones proposed in this paper. In addition, it should be noted that software tools that have been developed as "general" art technologies have not been included in this summary as they were not developed specifically as therapy tools.

7. CONCLUSION

We have described a novel tool for arts therapists who work with clients with dementia. The tool was built using a user-centered design process with a survey, focus group, and set of interviews with arts therapists. The tool is also customizable by therapists, allowing them to build user-specific applications for clients, and to create an intelligent system that can help a user for short periods of time.

Future work involves the design and implementation of a new device based on the requirements outlined in this paper, and the testing of this device (scheduled for Sept. 2010). Beyond that, we plan to integrate some additional features suggested by therapists, such as monitoring of biometrics, using multi-touch, and using hardware art tools.

8. ACKNOWLEDGMENTS

This work was supported by the American Alzheimer's Association under the Everyday Technologies for Alzheimer's Care program. We thank Patrick Olivier, Dan Jackson, Laurel Young, Jen Boger, and our therapist participants.

9. REFERENCES

[1] G. Bradski and A. Kaehler. *Learning OpenCV*. O'Reilly Media, 2008.

[2] K. Collie and D. Cubranic. An art therapy solution to a telehealth problem. *The Journal of the American Art Therapy Association*, 16(4):186–193, 1999.

[3] J. Harlan. Beyond the patient to the person aspects of autonomous functioning in individuals with mild to moderate dementia. *The American Journal of Art Therapy*, 28(4):99–106, 1990.

[4] P. Hartwich and R. Brandecker. Computer-based art therapy with inpatients: Acute and chronic schizophrencis and boderline cases. *The Arts in Psychotherapy*, 24(4):367–373, 1997.

[5] H. Hsieh and S. Shannon. Three approaches to qualitative content analysis. *Qualitative Health Research*, 15(1277), 2005.

[6] A. Karp, S. Paillard-Borg, H.-X. Wang, M. Silverstein, B. Winblad, and L. Fratiglioni. Mental, physical and social components in leisure activities equally contribute to decrease dementia risk. *Dementia and Geriatric Cognitive Disorders*, 21(2):65–73, 2006.

[7] H. Kautz, L. Arnstein, G. Borriello, O. Etzioni, and D. Fox. An overview of the assisted cognition project. In *AAAI Workshop on Automation as Caregiver*, Edmonton, 2002.

[8] S. Kim. Computer judgment of main color in a drawing for art psychotherapy assessment. *The Arts in Psychotherapy*, 35:140–150, 2008.

[9] S. McElroy, A. Warren, and F. Jones. Home-based art therapy for older adults with mental health needs: Views of clients and caregivers. *Art Therapy: Journal of the American Art Therapy Association*, 23(2):52–58, 2006.

[10] A. Mihailidis, S. Blunsden, J. Boger, B. Richards, K. Zutis, L. Young, and J. Hoey. Towards the development of a technology for art therapy and dementia: Definition of needs and design constraints. *The Arts in Psychotherapy*, 2010.

[11] A. Mihailidis, J. Boger, M. Candido, and J. Hoey. The coach prompting system to assist older adults with dementia through handwashing: An efficacy study. *BMC Geriatrics*, 8(28), 2008.

[12] K. Muise and J. Yim. A touching harmony: Midas in artistic practice. In *Proceedings of the 2nd international Conference on Tangible and Embedded interaction*, pages 213–214, Bonn, Feb 2008. ACM.

[13] E. Mynatt, I. Essa, and W. Rogers. Increasing the opportunities for aging in place. In *Proc. of ACM Universal Usability Conference*, Arlington, VA., 2000.

[14] M. E. Pollack. Autominder: A case study of assistive technology for elders with cognitive impairment. *Generations: The Journal of the American Society on Aging*, 30(2):67–69, 2006.

[15] H. Raffle, C. Vaucelle, R. Wang, and H. Ishii. Jabberstamp: embedding sound and voice in traditional drawings. In *Proceedings of the Conference on Interaction Design and Children (IDC)*, Aalborg, Denmark, June 2007. ACM Press.

[16] J. Rusted, L. Sheppard, and D. Waller. A multi-centre randomized control group trial on the use of art therapy for older people with dementia. *Group Analysis*, 39(4):517–536, 2006.

[17] M. Seligman, editor. *Helplessness: On Depression, Development, and Death*. W.H. Freeman, San Francisco, 1975.

[18] M. T. J. Spaan and N. Vlassis. Perseus: Randomized point-based value iteration for POMDPs. *J. of Artificial Intelligence Research*, 24:195–220, 2005.

[19] R. Wolf. Advances in phototherapy training. *The Arts in Psychotherapy*, 34:123–133, 2007.