# A Planning System Based on Markov Decision Processes to Guide People with Dementia Through Activities of Daily Living

Jennifer Boger, Jesse Hoey, Pascal Poupart, Craig Boutilier, Geoff Fernie, and Alex Mihailidis

*Abstract*—**Older adults with dementia often cannot remember how to complete activities of daily living and require a caregiver to aid them through the steps involved. The use of a computerized guidance system could potentially reduce the reliance on a caregiver. This paper examines the design and preliminary evaluation of a planning system that uses Markov decision processes (MDPs) to determine when and how to provide prompts to a user with dementia for guidance through the activity of handwashing. Results from the study suggest that MDPs can be applied effectively to this type of guidance problem. Considerations for the development of future guidance systems are presented.**

*Index Terms*—**Markov decision process, dementia, autonomous guidance, activities of daily living**

## I. INTRODUCTION

### A. Scope of Paper

THIS paper presents the design and preliminary evaluation of a planning system that is a critical component of a larger information technology system to guide people with dementia through activities of daily living (ADLs). This planning system is built using a Markov decision process (MDP), a decision-theoretic model capable of taking into account both uncertainty in the effects of its actions and tradeoffs between competing short-term and long-term objectives when making decisions. The planning system discussed in this paper is designed to be integrated with the tracking system presented in [1], and a prompting system that is still under development. This paper evaluates the planning system independently of these other systems through the administration of an efficacy study involving professional caregivers.

J. Boger, J. Hoey, and A. Mihailidis are with the Intelligent Assistive Technology and Systems Lab, Dept. of Occupational Therapy, University of Toronto, Toronto, ON, M5G 1V7, Canada (e-mail: jen.boger@utoronto.ca; jhoey@cs.toronto.edu; alex.mihailidis@utoronto.ca).

P. Poupart is with the Dept. of Computer Science, University of Waterloo, Waterloo, ON N2L 3G1 Canada (e-mail: ppoupart@cs.uwaterloo.ca).

C. Boutilier is with the Dept. of Computer Science, University of Toronto, Toronto, ON, M5S 3G4, Canada (e-mail: cebly@cs.toronto.edu).

G. Fernie is with the Toronto Rehabilitation Institute, Toronto, ON, M5G 2A2, Canada (e-mail: fernie.geoff@torontorehab.on.ca).

### B. Problem Definition

There are currently 18 million people worldwide who are diagnosed with dementia, with numbers predicted to rise to 35 million by 2050 [2]. These estimates reflect a combination of an increasing number of older adults and the prevalence of dementia doubling every five years in patients over the age of 65 [3].

Older adults with dementia often forget how to complete ADLs, such as handwashing, dressing, and toileting, and rely on assistance from a caregiver. When the care recipient encounters difficulties in ADL completion, a caregiver provides cues to the care recipient for the next step required to progress in the activity. If the level of dementia worsens, as occurs with conditions such as Alzheimer's disease, the caregiver experiences greater feelings of burden as a result of increasing demands of caregiving duties. Generally, there comes a point where the caregiver feels s/he can no longer cope and the care recipient is permanently transferred to a long-term care facility.

Assistive information technology has the potential to delay institutionalization by alleviating some caregiver duties while restoring partial autonomy to the care recipient [4,5]. One prospective application of assistive technology is through partial compensation of the memory loss that often accompanies dementia in older adults. A specific example of this application is assistance with ADL completion.

### C. Previous Devices to Support the Completion of Activities of Daily Living

There have been several different types of aids designed to increase the independence of people with significant memory impairments by support of ADL completion. An example is the memory wallet, which was filled with pictures familiar to the user that served as cues for remembering tasks and people [6]. A touch screen program has been developed by Hoffman et al. [7] where the user is presented with a series of pictures of his/her surroundings and "touches" his/her way through a sequence of photos depicting step-by-step guidance through an activity. The electronic memory aid developed by Oriani et al. [8] allows a user or caregiver to pre-record messages, such as reminders on how to complete a task, which are played back to the user at prescribed times. Another electronic, handheld system developed by Levinson [9] uses classical (deterministic) planning algorithms to compute a "best" plan

for completion of an activity, and provides step-by-step guidance through tasks in the form of visual and audio cues. The Autominder System developed by Pollack et al. [10] uses dynamic Bayesian networks as an underlying domain model to coordinate pre-planned events in an attempt to ensure that scheduled tasks are executed without interfering with each other or with other activities, such as watching television. Pineau *et. al.* [11] have recently used a variant of partially observable Markov decision processes (POMDPs) to design the high level control system for "Nursebot", an artificially intelligent robot designed to assist elderly people with daily activities. The robot primarily provides intelligent reminders regarding specific activities (like Autominder) but also engages in a certain degree of social interaction.

While compensating for losses in memory function, these devices are still impractical for the more severely impaired population as they require user feedback, such as a button press or dialogue, to operate. This is an unreasonable expectation of this population as they are unlikely to remember how or why to respond to vague stimuli, such as a beeping alarm. It is also unreasonable to expect the caregiver to be continually interacting with the device, as this adds to his/her already extensive list of caregiving duties. If they are to be effective, devices aimed at aiding people with dementia must be able to operate autonomously, without any explicit feedback from the care recipient or the caregiver.

### D. Overview of Guidance System Project

The planning system described in this paper is designed to be a part of a larger guidance system that the authors and their collaborators are currently developing. The guidance system will unobtrusively monitor older adults with moderate-to-severe dementia and provide autonomous guidance to assist in the completion of ADLs, in particular, handwashing. The activity of handwashing was chosen for three reasons: 1) handwashing is a problematic activity for the moderate-to-severe dementia group; 2) handwashing is deemed relatively safe for clinical trials; and 3) it is anticipated that technology developed to model the activity of handwashing will be generalizable to other ADLs.

The guidance system consists of three sub-systems; sensing, planning, and prompting. The sensing system uses a video camera and a computer vision to track the position of the user's hands and the position of objects that are relevant to activity completion (e.g. the soap and the towel). Our previous work [1] describes an automated sensing system that will eventually be integrated with the planning system. In this paper, we simulate the action of the sensing system using a human operator. Specifically, a human operator input the position of the hands for use by the planning system (as discussed further below). The planning system determines the prompt to be given based on the input provided by the sensing system. The prompting system communicates the selected prompt to the user. In this paper, the prompting system was simulated by a human caregiver (who read the prompts provided by the planning system, as detailed below) to ensure

integrity of the prompts heard by participants in the efficacy study.

## II. DEVELOPMENT OF THE PLANNING SYSTEM

### A. Planning System Criteria

The following criteria were used to guide the design of the planning system. The planning system must be able to:

1) operate without explicit feedback from the user or his/her caregiver,

2) have a framework that is generalizable to other ADLs,

3) detect user progress through the activity,

4) capture enough of the washroom environment to appropriately guide the user through handwashing (i.e. correct identification of the next step in the task as well as timing and repetition of prompts), and

5) handle user regression/departure from an appropriate sequence of steps required for handwashing completion.

Criteria 3), 4), and 5) are the focus of this paper. Section III presents our efficacy study, designed to show how the MDP model fulfills these criteria. Results from the efficacy study are presented in Section IV and discussed in Section V. While criterion 1) is not directly addressed in this paper — the vision and planning systems were not tested directly, hence some human input was required — the prompting policy constructed by the planning system does not require explicit user or caregiver input or feedback. Criterion 2) must be answered by future research, as discussed in Section V.B.

### B. Markov Decision Processes (MDPs)

MDPs have been widely used in both operations research [12] and artificial intelligence [13] to model and solve decision-theoretic planning problems—essentially providing a model of a system's interactions with its environment and allowing one to construct appropriate policies to guide the system's control of the environment. MDPs are attractive as a model for ADL assistance for several reasons. First, they can capture the underlying stochasticity of a domain. As such, the MDP framework naturally lends itself to a problem such as handwashing, where the outcomes of actions taken by the system are uncertain (e.g. the user may be prompted to turn the water off, but may dry his/her hands instead). Second, an MDP allows one to account for multiple, potentially conflicting objective criteria, both short- and long-term. For example, in ADL assistance technology, we would like to allow the user as much independence as possible (minimal prompting to the user and summoning of the caregiver); but at the same time, we want to ensure the activity is completed successfully. As a result, the MDP model has the potential to address all of the design criteria presented in Section II.A. For an in-depth discussion of the MDP concepts presented in this paper, refer to [12,13].

An MDP consists of the following components. A finite set of states $S$ denotes the set of possible joint configurations of

the environment and system relevant to the prediction of the effects of actions and objective satisfaction. A finite set of actions *A* correspond to the actions available to the system that influence state. A transition function *P: S x A → Δ(S)* captures system dynamics, and reward function *R: S x A→ℜ* represents various objective criteria. In our model, a state is comprised of a combination of instantiations of each of the variables used to define the status of the planning system and its environment (these variables are discussed in detail below). Actions are simply the various prompting (and other) choices available to the system. If the system is in state *s* and takes action *a*, it will transition to future state *s'* with a known probability *P(s,a,s')*. Rewards and costs (i.e. negative rewards) are incurred by the system for taking certain actions in particular states: *R(a,s')* denotes the reward received for taking action *a* in states *s*.

Given such a model of a domain, solving an MDP requires that one construct a policy that maximizes the long-term, discounted expected reward. More precisely, we focus our attention on *stationary policies* of the form *π: S → A*, where *π (s)* denotes the action to be taken when the system is in state *s*. For any state *s*, any such policy induces a distribution over the sequence of rewards to be received. Our goal is to construct a policy that maximizes the expected discounted value of such a reward stream over the infinite horizon: $E(\sum_{t=0}^{\infty}\beta^t R_t \mid \pi, s)$.

Here $R_t$ denotes the reward received at time *t* and *0≤β<1* is a discount rate that ensures the sum of rewards is bounded.[1] The optimal policy can be constructed using any of a number of classic algorithms, based on either linear or dynamic programming [12]. The optimal policy can be computed offline (before the system is installed). Once we have the mapping associating specific actions with system states, the policy can be implemented online with minimal computation.

One difficulty with using standard dynamic programming algorithms for solving MDPs in a domain like ours is state space size. Since our states are defined using a number of distinct variables as we see below, our state space is exponential in the number of systems variables. In our case, this leads to over 22 million states, rendering explicit representation of the model (dynamics and reward function), explicit representation of the policy, and solution by enumeration of the state space, all infeasible. Fortunately, our domain has considerable structure, which we can leverage in two different ways. First, *dynamic Bayesian networks (DBNs)* [13,14] allow compact representation of the model by exploiting various independences among features. Namely, independent relationships between variables do not need to be explicitly enumerated, significantly reducing the size and complexity of the model. Second, algebraic decision diagrams [15], capture regularities in the local distributions used by the DBN model, to simplify the representation of the model.

---

[1] The use of an infinite horizon accounts for the fact that the time at which an activity will be completed cannot be bounded a priori (if it will be cut off after a fixed period of time, a finite horizon could be used, though the infinite horizon model is still applicable). Discounting also associates greater value to quicker task completion.
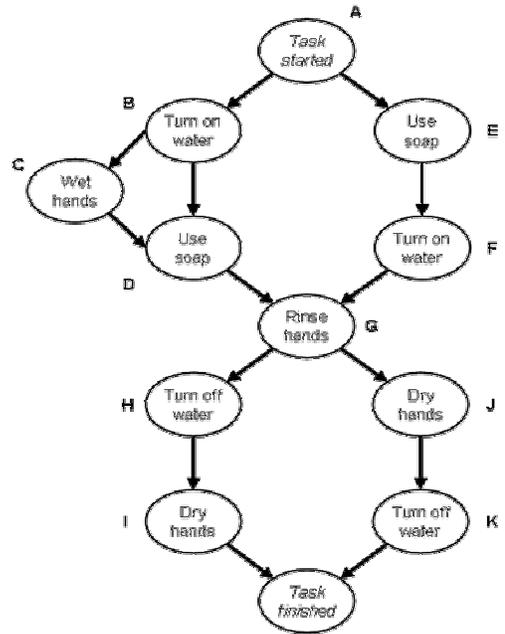


Fig. 1. Plan graph for the activity of handwashing (adapted from [16]). The step "wet hands" was considered optional as liquid soap was used. Arcs representing possible paths for regression have been omitted for simplicity.

We can then use a specialized dynamic programming method, SPUDD, which exploits this structure to construct an optimal policy without explicit state space enumeration [15].

### C. Planning System Design

Fig. 1 describes the steps and pathways used to describe the activity of handwashing. The critical steps in the handwashing activity are represented by nodes. The direction of the arrows represents progression through the handwashing activity from one (step) node to the next. Since there is more than one "correct" way to wash one's hands, the plan graph in Fig. 1 captures different acceptable sequences of steps by representing several alternate pathways. Any pathway the user follows from node A to either node I or K corresponds to a successful execution of the handwashing task. The handwashing activity is defined by the following variables, actions, and dynamics.

*1) Planning System Variables:* Two <u>environment</u> variables are used to describe the observable values of the handwashing environment. *HP* denotes the user's hand position (one of *sink, tap, towel, soap, water, away*) and *WF* denotes water flow (*on* or *off*). The different *HP* regions can be seen in Fig. 2. The towel and soap regions are associated with the objects themselves, and will therefore move if these objects move. The model is designed to make decisions based solely on input corresponding to these environment variables, obviating the need for explicit user or caregiver interaction. While both *HP* and *WF* are dictated by a human operator in the trials reported in this paper, in the full systems, these values are determined by the vision system (and, as such, are subject to noise, as we discuss further in Section V.B).
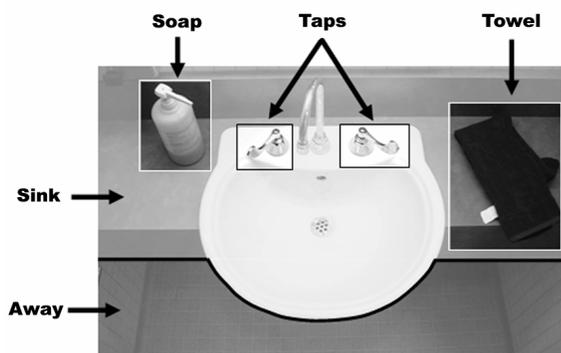
Fig. 2. Definition of *HP* variables (note "*water*" is located under the faucet and is only applicable when water is flowing from the tap)



Fig. 3. Example of how the probability of *WF* being "*on*", *P(WF=on)*, varies with the number of wait times (where *NW* = 2.5 s) when the previous action was to: give no prompt, prompt the user to turn the water on at the minimal level, and prompt the user to turn the water on at the specific level.

<u>Activity status</u> variables capture the user's progression through the activity. *PS* (plan step: *A to K*) denotes the last successfully completed step, corresponding to the letters beside the step nodes in Fig. 1. If *PS* has a value of I or K, then the user has successfully completed handwashing. *MPS* (maximum plan step: *A to K*) denotes the maximum value of *PS* reached during the user's current attempt at the activity. If the value of *PS* does not occur later in the plan graph than the value of *MPS*, the user has regressed in the activity. *MPSrepeat* (maximum plan step repeat: *yes, no*) represents whether or not a user has previously visited the plan step currently being attempted. *MPS* and *MPSrepeat* are reset on completion of the activity. *MPS* and *MPSrepeat* are necessary for the calculation of the reward function, which is described later in this paper.

<u>History</u> variables capture a history of the user's behaviour and the system's actions. These variables are *NP, PL, LP, NW, Prog,* and *Reg*. *NP* refers to the number of prompts (*zero, one, two, three+*) given during the current plan step and is reset to zero when *PS* changes. *PL* denotes the prompt level (*minimal, moderate, specific*) of the last prompt given, while *LP* denotes the direction of the last prompt that was given (*water on, water off, use soap, wet hands, rinse hands, dry hands*). *NW* refers to the number of time steps where no prompting action has been taken (*zero, one, two, three, four+*). *NW* increases when there is a user-driven, non-progressive change in state (e.g. the user moving his/her hands away from the towel when it is time to dry his/her hands) or when a defined amount of time has elapsed without a change in state, and is reset to whenever *PS* changes (either though completion of the step or regression in the activity) or *Prog* has the value *yes* (see below), to provide the user with a "fresh start" each time a step is attempted. *Reg* denotes how many times the user has regressed in the activity (*zero, one, two, three+*). The value of *Reg* increases if the step the user is currently attempting is earlier in the plan graph (Fig. 1) than the previously completed step. Finally, *Prog* (*yes, no*) indicates whether the user is progressing within the step s/he is attempting. The values of *Prog* and *NW* are determined by changes in the user's hand position. Specifically, if a user's hand moves towards the next area relevant to the step s/he is attempting, *Prog* equals *yes* and *NW* is reset to zero. If *NW* increases because a set amount of time
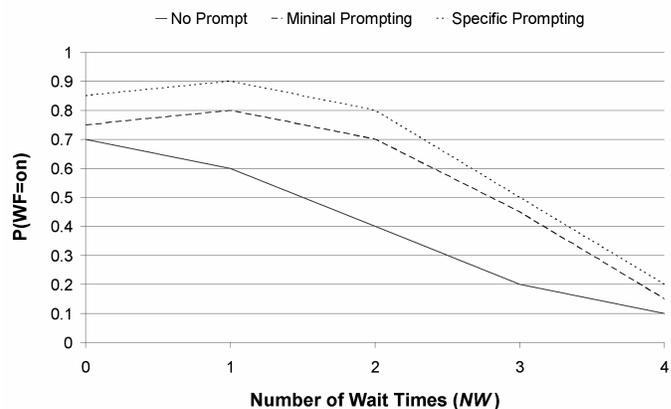
has elapsed without a change of state or because the user's hand moves to an area that is not conducive to step completion, *Prog* is set to *no*.

These variables (through all combinations of instantiations of their values) results in a state space size of 22,302,720 states.

*2) Planning System Actions:* There are twenty actions available to the planning system. The system can prompt the user to attempt six different subtasks, namely turn the water on, turn the water off, use the soap, rinse hands, wet hands, or dry hands. Each of these prompts can be given at a minimal (e.g. "Use the soap"), moderate (e.g. "Use the pink soap"), or specific (e.g. "John, use the soap in the pink bottle") level, for a total of eighteen possible prompting actions. Minimal prompts are designed to gently cue the user, whereas specific prompts are designed to get the user's attention as well as provide the user with more details on how to complete a step. As such, specific level prompts are pre-recorded for each user so that they include his/her name. The system can also choose to take the action of doing nothing (i.e. simply observe the user) or to call the caregiver to intervene.

*3) Planning System Model Dynamics:* The dynamics of the planning system were manually specified by one of the authors using prior knowledge of the domain, gained through extensive observation of a professional caregiver guiding ten subjects with dementia through handwashing [17]. Future work includes learning the parameters from data, as discussed in Section V.B. Because of the size of the state space (even exploiting sparsity in the transition matrices), it is impossible to specify all transition parameters in explicit form. Fortunately, the DBN representation we use allows the complete model to be specified with far fewer parameters, and provides for a very natural decomposition of the transition distribution function.

The only stochastic dynamics involve the environment variables *HP* and *WF*. All other variables are updated deterministically as a function of changes in the state of these variables or elapsed time. (For example, whether a particular

plan step has been completed, calling for an increase in the variable *PS*, is a function of the current plan step and the two environment variables).

For all steps in the activity, appropriate prompting actions increase the estimated probability of successful step completion as did giving prompts at higher levels of detail. An example of model dynamics is shown in Fig. 3, which depicts the probability of the user turning on the water, *P(WF=on)*, as a function of time and the system prompting action, at a specific state (namely, when *PS = A, WF=off, NP=0*; note that this probability is independent of the status of any other variables when *NP=0*, a fact that is exploited by the DBN representation). If no prompt is given, this probability decreases over time. This reflects the decreasing likelihood of the user ever turning on the water if s/he doesn't do so fairly quickly. The effects of minimal and specific prompting on this probability distribution are also shown in Fig. 3. Prompting makes it more likely that the patient will turn the water on. Waiting for one time step after prompting to turn on the water gives the user some time to comprehend and respond to the prompt. After this, there is a decline in *P(WF=on)* reflecting the probability that the user has forgotten the prompt or has otherwise lost his/her focus about the current step. Higher levels of prompting detail are assumed to result in higher user compliance, and therefore higher values of *P(WF=on)*.

Again, despite the fact that the model contains over 22 million states (thus requiring roughly half a trillion state transition parameters for each pair of states), manual enumeration of the transition probabilities was possible by exploiting the DBN structure using intuitions such as those described above.

*4) Reward Function:* The reward function employed by the planning system can be seen in Table 1. This reward function exploits a standard additive decomposition, decomposing *R(s,a)* into a reward associated with each state and a cost (negative reward) associated with each action (with *R(s,a)* being the sum of the two. Because rewards are associated with specific state features rather than individual states, the reward function can be specified compactly as described in the table.

A large reward is given when handwashing is considered complete (i.e. when *PS* equals *I* or *K*). A smaller reward is given the first time a new plan step is reached (the first arrival at a specific plan step occurs when *MPSrepeat* equals *no*).

Table 1: Definition of reward function, *R(s,a')*, for MDP model. The reward given to the system is dependent only on the action taken and the value of the variables *planStep* and *MPSrepeat*. The reward is the sum of the state reward and the action cost.

| Action | planStep | MPSrepeat | Reward Value |
|---|---|---|---|
| None | - | - | 0 |
| *Minimal* detail level | - | - | -3 |
| *Moderate* detail level | - | - | -5 |
| *Specific* detail level | - | - | -7 |
| Call caregiver | - | - | -1000 |
| - | A,B,C,D,E,F,G,H,J | No | 3 |
| - | I,K | No | 300 |
| - | - | Yes | 0 |

Structuring the reward function this way encourages the policy to complete the entire activity, but in cases where the completion odds are very low, to try to make as much progress as possible. Conditioning rewards for progress on *MPSrepeat=no* (as opposed to the *PS* variable) deters possible cyclical regression in guidance of the activity (so attempting to guide a person to complete a specific step *multiple* times will not be rewarded.)

There is a small cost associated with prompting that is proportional to the level of detail of the prompt. By making this cost proportional to prompt specificity, we reward prompting that encourages increased user independence slightly more than prompting that is very specific. The overall effect of this (as embodied in the optimal policy) is that the system generally begins prompting the user at the minimal level, resorting to more detailed levels of prompting only if the user is not responding. This strategy is intended to provide the user with as much independence as possible.[2]

In general, summoning the caregiver is penalized, as executing this action is assumed to result in activity completion with aid from the human caregiver. Since our aim is to complete the task without human intervention if possible, this cost should be set high enough to deter this choice unless the probability of the user completing the activity in a reasonable period of time is sufficiently low, or of the costs of prompting are sufficiently high. The cost should generally be set lower than the value associated with task completion however, since the net reward attached to calling the caregiver at any point in time should be positive (otherwise the system will get greater value by simply doing nothing "forever"). Because of our experimental design, however, it was important that the caregiver never be called (since, as discussed below), we evaluate our system through a comparison with a human caregiver (which will not have this action in his/her repertoire). Thus the large penalty of -1000 used made it impossible for the call caregiver action to occur in the optimal policy. This would not be the case the deployed guidance system, since calling the caregiver is an important option.

The values shown in Table 1 were determined in an iterative process in which the reward function was successively altered until desired performance was attained, as qualitatively assessed through a series of simulated trials. The system was rewarded at each time step based on the action the system took and the values of *MPSrepeat* and *planStep*. Our reward function was designed to promote user independence, completion of overall task, and minimal regression by the user. One focus of our ongoing research is to involve caregivers in the refinement of this reward function through the evaluation of the relative utility of possible outcomes. This study serves as a first step towards this goal.

---

[2] Naturally, the fact that the optimal policy starts out using less specific guidance depends critically on the systems dynamics as well. If minimal prompting had a very small probability of successfully inducing the user to complete a task step, it would be forgone in favor of an immediate, more specific prompt. This is discussed further in Section V.B.

## D. Computing the Policy and Using the Planning System

The optimal policy is computed using a well-known algorithm called value iteration [12]. We used a factored, structured implementation of value iteration called SPUDD, which exploits the compact DBN representation of dynamics and the additive structure of the reward function, to compute the policy [15]. The optimal policy was computed in 64 minutes and was represented with an algebraic decision diagram with 3284 internal nodes (in contrast, the explicit policy representation would have required over 22 million table entries to reflect the optimal action for each state).

The planning system operates as follows. Each time a change in the state is detected, the system consults the optimal policy and executes the action specified for the current state. To detect state changes, the current values of the state variables are deterministically calculated based on the action taken and the previous variable values, except for hand position ($HP$) and water flow ($WF$). The values of $HP$ and $WF$ are normally obtained from sensors (i.e. video camera and water flow sensor), but for the purpose of evaluating the planning system in isolation and to achieve full observability, these were supplied by a human observer.

## III. DEVELOPMENT OF EFFICACY STUDY

### A. Objective of Efficacy Study

The efficacy study evaluated the performance of the planning system according to the criteria in Section II.A. In the study, professional caregivers evaluated the effectiveness of the guidance strategies employed by the planning system and those of a human caregiver. Participants viewed videos of simulated handwashing scenarios and subsequently rated the effectiveness of the prompting s/he saw by completing a questionnaire. Both quantitative and qualitative data were captured. Trials were simulated (i.e. using an actor, as also conducted by [18]) at this point because the authors felt the technology should be proven and refined as much as possible before conducting clinical trials with real dementia subjects.

Section III.B describes how the video trials recorded for the efficacy study were used to evaluate how well the MDP system detected the user's progress through the handwashing activity (criterion 3). Section III.B also describes how the system handles user regression, illustrating it behaviour with a sample sequence (criterion 5). Section III.C describes how the efficacy study was designed to capture the appropriateness of MDP system guidance (criterion 4).

### B. Construction and Composition of Video Clips

A researcher impersonated an individual with moderate-to-severe dementia washing his hands. These handwashing scenarios were either guided by a human caregiver (human-guided) or by the MDP planning system (MDP-guided) and were recorded as video clips using a camcorder mounted over the sink area. Both the researcher and the caregiver had extensive experience  interacting with people with dementia (over 30 years exposure apiece). In the human-guided

scenarios, the caregiver was free to guide the actor through the activity of handwashing as she wished, provided that guidance was strictly verbal and she did not enter into the field of view of the camera. In the MDP-guided scenarios, the status of the actor's hand position and water flow was input by a researcher directly to the MDP policy. The corresponding optimal action defined by the MDP was displayed as a script on a monitor, which the caregiver read aloud to the actor. The caregiver was instructed to read the entire prompting script verbatim as soon as it was displayed. A caregiver was used to read prompts, as opposed to using pre-recorded prompts, in an effort to ensure that the voice guiding the subjects in both the human-guided and MDP-guided scenarios sounded as similar as possible. When the optimal action was "nothing", the monitor remained blank. Each scenario began when the actor was positioned at the sink and ended when his hands were successfully washed. To gauge how well the planning system addressed criterion 3) (defined in Section II.A), the operator checked to ensure that the visual observations of the user and the sink environment matched the appropriate $PS$ value (depicted by Fig. 1) for each MDP scenario.

Fig. 4 depicts snapshots of one of the MDP-guided scenarios used in the efficacy study. In this case, the handwashing subject is able to complete the steps of turning on the water ($PS$=B) and turning the water off ($PS$=K) independently. The subject ignores the prompt given to him at t=71s to dry his hands and regresses in the activity to $PS$=D by applying soap instead. The planning system copes with this regression (criterion 5 from Section II.A) by identifying that the subject has returned to $PS$=D, and prompts him to rinse his hands. The majority of the time the prompting screen remained blank, as portrayed at t=6s and t=55s, allowing the subject time to attempt each step on his own. We discuss this example further in Section V.A.

Of the 33 recorded handwashing scenarios, six were randomly selected for evaluation in the efficacy study. Three of the scenarios were human-guided and three were MDP-guided. In order to examine prompting strategies while ensuring that interviews would be an hour or less, scenarios included in the study had to: 1) be less than three minutes in length (deemed to be an appropriate amount of time from handwashing trials by [19] with nine subjects with dementia, which ranged between 50 seconds and four minutes); 2) not show undue amounts of the actor's head or face, for privacy reasons; 3) demonstrate interaction between the actor and the prompter (i.e. the actor was not overly independent in activity completion); and 4) not have unnecessarily distracting behaviour (e.g. the actor throwing the towel at the caregiver at the end of a scenario).

### C. Questionnaire Design

Professional caregivers (participants) were asked to fill out a questionnaire after each video clip was viewed, rating the effectiveness of the prompting s/he had observed. The questionnaire was designed to capture both qualitative and quantitative data. The quantitative portion of the evaluation

| Time (s) | User status | PS | Prompt |
|---|---|---|---|
| 0 | | A | Time to wash your hands, John |
| 6 | | B | |
| 14 | | B | Put some soap on your hands |
| 37 | | D | Rinse your hands under the water |
| 55 | | G | |
| 71 | | G | Pick up the towel on your right |
| 80 | | D | John, rinse your hands |
| 104 | | G | Use the towel on your right |
| 170 | | K | Thank you John, we're all done here |

Fig. 4. Sequence of events for a MDP-guided handwashing scenario used in the efficacy study, with accompanying PS references and MDP planning system prompts (note that the PS value refers to the last step that was completed)

was accomplished through the completion of a five-point Likert rating scale [20] of five aspects of effectiveness. These aspects were:

Identification - The prompt(s) given to the patient appropriately identified the next task in handwashing.

Detail - The level of detail of the prompt(s) was appropriate for completion of the task.

Time - The patient was given an appropriate amount of time to attempt the task before being prompted.

Repetitions - The number of repetitions of the prompts was appropriate for completion of the task.

Overall Effectiveness - The patient was guided effectively.

The *Identification*, *Time*, and *Repetition* questions were designed to gauge criterion 3), namely how appropriately the planning system prompted the user. *Detail* was not considered part of the evaluation of this criterion, as the prompts are scripted and therefore a reflection of the wording of the predetermined prompting script as opposed to decision making by the system. *Overall Effectiveness* was also not included as this was considered to reflect the effectiveness of the system rather than the appropriateness of the prompting.

The qualitative portion of the study consisted of a space at the bottom of each questionnaire page, where participants were encouraged to make any comments s/he had. After viewing all the video clips, participants were asked to fill out an "Evaluation Summary" sheet, which asked participants if they thought the prompting in the videos was believable (and if not, why not), if there were any specific aspects in the prompting strategies that could be improved upon, as well as any other comments the participant had.

### D. Administration of Efficacy Study

There were 30 participants in the efficacy study, all of whom provided professional care (including guidance through ADLs) to people with dementia in long term care facilities at the time of the study. The researcher did not inform the participants that computer guidance was involved in any way. Each participant was interviewed individually, with the exception of participants S9 and S10 who were interviewed at the same time. During each interview, the researcher explained the study procedure to the participant, had the participant sign consent, and then proceeded to show video clips of the handwashing scenarios and administer the accompanying questionnaires. Video clips were shown to participants on a laptop computer, operated by the researcher. Participants were shown a full handwashing scenario and asked to fill out a questionnaire. S/he was then shown each step in the scenario s/he just saw (e.g. turning on the water, using the soap, etc.) and asked to fill out a questionnaire after watching each step. This procedure was repeated for the remaining five scenarios for a total of 31 questionnaires per participant. Scenarios were presented in a non-repeated, random order to ensure each participant viewed a unique sequence of scenarios. Participants were invited to ask questions at any point during the interview. Interviews were conducted at the participant's institution of employment and were approximately one hour in length.

## IV. RESULTS

### A. Quantitative Results

An analysis of variance (ANOVA) was run to compare the ratings between the two guidance types for each aspect of effectiveness queried by the efficacy questionnaire. The results of this analysis can be seen in Fig. 5. As the human-guided scenarios achieved higher ratings than the MDP-guided scenarios for every aspect with high statistical confidence, it can be surmised that the overall prompting strategy employed by the human caregiver was considered by the participants in the study to be significantly more effective than the strategy employed by the planning system. There were no significant guidance*scenario [$F(2,58)=2.89$ $P_\alpha<0.76$] or guidance*aspect [$F(4,116)=0.80$ $P_\alpha<0.52$] interaction effects detected.

Of 150 rating comparisons (five aspects for two guidance types rated by thirty subjects), there were eleven instances where MDP-guidance was given a higher rating than human-guidance (specifically, three instances for Identification, two for Detail, four for Timing, two for Repetition, and zero for Overall effectiveness) and seventeen instances where the ratings were equivalent (three instances for Identification, four for Detail, three for Timing, five for Repetition, and two for Overall effectiveness). Eleven of the thirty participants recognized that there was an actor in the clips and six of thirty thought the scenarios were not believable. Four of the participants who considered the scenarios to be unbelievable also recognized that an actor was involved.

### B. Qualitative Results

The following themes were the most common in commentary feedback from the caregiver participants, with the number of participants who made a relevant comment listed in brackets after the cited theme. Participants thought it was important when prompting a patient to consider:

- the language construction of the prompts (22),
- allowing the patient enough time for step completion (17),
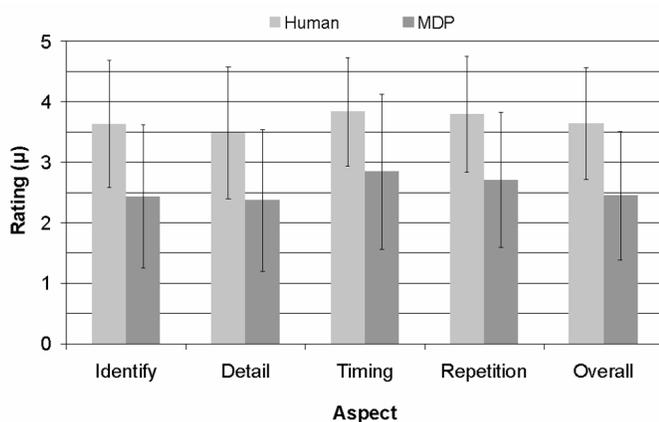- turning the water off before asking a patient to dry his/her hands (13),



Fig. 5. Mean effectiveness rating per aspect of efficacy study for human and MDP guidance (Differences were significant with a $P_\alpha<0.0001$)

- the effectiveness of having visual cues to accompany verbal ones (13),
- using prompts that are tailored to each individual patient's abilities (12),
- the use of positive feedback, a friendly voice, and being polite when prompting (11),
- turning the water on before asking a patient to use the soap (7), and
- checking the water temperature before the patient washes his/her hands (5).

None of the comments suggested that the MDP-guided scenarios were more or less believable than the human-guided ones. There were also no comments made regarding any perceived differences in the tone of voice of the prompter between the MDP-guided and human-guided scenarios.

## V. DISCUSSION

### A. Interpretation of Results

The higher ratings attained by the human-guided scenarios are particularly significant considering that none of the participants were told that there was computer guidance involved in any way. Since there were no specific comments about the believability or tone of voice in the MDP-guided scenarios, it is postulated that the discrepancies in the average rating of the effectiveness aspects were a result of the strategies used by the different guides in the scenarios. However, although the human-guided scenarios achieved higher ratings than the MDP-guided ones, this does not necessarily indicate that the planning system performed poorly. When the ratings are compared, the participants considered MDP-guidance of the subject in the scenarios to be as good as or better than human-guidance in 28 of 150 instances (18.7 % of the ratings). Therefore the study indicates that, while participants felt the MDP-based planning system did not guide users as effectively as the human caregiver, the MDP-guided scenarios were still comparable in effectiveness to the human-guided ones. We now examine how this data supports the design criteria proposed in Section II.A.

One of our claims was that the MDP system could appropriately guide the user through handwashing (criterion 4). We feel that the relative performance of the MDP prompting system to that of a professional caregiver shown in Fig. 5 justifies this claim with modest success. This is all the more compelling when we consider that our goal is not to replace human caregivers completely, but rather to relieve the burden of caregiving by increasing the odds of independent task completion or simply making greater progress in the task before requiring assistance.

Furthermore, the qualitative comments made by the participants provided insight into the areas that need further work to improve the appropriateness of the system's guidance. For example, qualitative comments made by the participants included the sequencing of activity steps. Many participants felt that turning the water on before asking a patient to use the

soap or turning the water off before asking a patient to dry his/her hands provides a natural cue as to what the next step in the activity is and helps to avoid possible confusion. Suggestions were also made on ways the wording of the prompts could be made to be more effective. These comments give insight into plausible reasons why MDP guidance rated lower than human guidance and can be incorporated to improve future versions of the guidance system, as discussed in Section V.B. Note that many of these suggestions do not expose fundamental limitations of the MDP framework *per se*, but simply reflect specific modeling and prompting choices that can easily be incorporated without significant changes to the model itself.

Some of the respondents did not feel the scenarios were believable and/or recognised the patient in the videos as an actor. One possible reason for this could be the lack of physical guidance or gesturing by the caregiver, which is commonly used to get attention when guiding people with dementia. When the respondents who thought the scenarios were unbelievable and/or identified the patient as an actor were removed, there was no change in the data trends and differences between human and MDP guidance were still significant ($P_\alpha < 0.001$).

The MDP system also handled user regression in the task reasonably effectively (criterion 5 from Section II.A), as exemplified by the MDP scenario depicted in Fig. 4. This ability was given positive feedback by the study participants. Not only did the participants who gave feedback approve of the prompting action when regression was encountered, they also approved of the fact there was no reprimanding of the subject. For example, one reviewer commented, "Good - caregiver did not correct John, but went back to rinsing prompt."

User progress through the task was successfully identified by the planning system in the efficacy trials. This was done by matching the real-time *PS* values calculated by the MDP model to the observed status of the user and the handwashing environment, as discussed in Section III.B. The value of *PS* matched the conditions of the trial in every observed instance, satisfying criterion 3).

Some of the comments made by the participants cannot be addressed through modifications to the current planning system, such as checking the temperature of the water and the ability for the system to respond to user queries. A concern such as water temperature could be addressed through adaptation of the environment, such as the installation of a water temperature regulation device on the faucet. It is unlikely that a system could have the capability of carrying on a conversation with the user using a pure MDP model. However, it would be valuable to be able to detect and react to certain key phrases from a user (e.g. "help").

*B. Future Work*

The model for the planning system that is described in this paper was based solely on a researcher's estimate of user behaviour, which was gained from the observation of

handwashing in a clinical setting. The authors are in the process of inducing model parameters from clinical data. It is anticipated that the incorporation of clinical trial data will result in a more accurate representation of user dynamics, therefore improving the performance of future versions of the planning system.

Many of the improvements suggested by participants regarding the guidance strategies employed in both the human- and MDP-guided scenarios can be incorporated into the MDP model. These recommendations include guiding the patient with his/her preferred ordering of steps, allowing the patient more time to complete a step, and using a level of detail in the prompts that the patient responds to.

As identified by the participants in the efficacy study, people have different preferences and abilities. A person is far more likely to comply with ADL guidance if s/he is prompted in a way that s/he is familiar with and understands. Therefore, it will be important to include information about user types or preferences in the state space for future versions of the planning module. For example, we can roughly characterize dementia patients into levels of responsiveness and attention, aspects that could be added to the MDP model.

However, in most cases, information about the user types will not be readily available on a day-to-day basis: these variables are not directly observable, can only be assessed probabilistically by a caregiver initially, and may evolve over time. One approach to this problem is the use of partially observable Markov decision processes (POMDPs) [21]. A POMDP provides a model of sequential decision making in which actions are conditioned on probabilistic estimates of the underlying state of the world (as opposed to MDPs, which always assume full knowledge of the true state). This allow one to represent inherent uncertainties in the state of the world, based on noisy observations (e.g. the noisy output of a vision system) and hidden variables (e.g. hidden user characteristics such as level of responsiveness) whose values can be inferred (probabilistically) through observed behaviour. The POMDP model reflects how user traits would impact state transitions, resulting in a more accurate representation of the washroom environment and better user guidance. It is also important that the system be able to autonomously converge on the prompting strategies that work best with a specific user, allowing the device to customize prompting strategies autonomously over time. The POMDP model allows for this by maintaining a probabilistic estimate of such characteristics that is updated with each observation of user behaviour. Apart from tailoring its prompting strategy to specific users, this also enables the system to adjust to changes in user abilities, especially important for users with progressive forms of dementia, such as Alzheimer's disease. Research into the application of POMDPs to ADL guidance is currently underway [22].

So far, we have used handwashing as an initial test-bed for planning and guidance concepts. However, if a system is to be implemented that successfully provides simultaneous relief to the caregiver and independence to the care recipient, it must be

able to address more than one ADL. ADLs such as dressing, cooking, and toileting, represent a greater guidance challenge, as they are more complex and the consequences of misguidance are higher than those in handwashing. Since assistance with these ADLs is cited as the most problematic for caregivers (toileting especially), automated guidance of these activities would likely provide the most relief to a caregiver. The MDP model presented in this paper could conceivably be altered to suit other ADLs. Although research would have to be conducted to determine the compositions of different activities, ADLs such as toileting and dressing represent sequential decision making problems that are similar in structure to handwashing. Thus a similarly structured MDP could likely handle a different ADL, and indeed, many of the variables would likely be the same, such as *NW*, *LP*, and *PL*.

The most crucial immediate task is to merge the sensing, planning, and prompting systems to create an autonomous guidance system. Experiments with this integrated guidance system are necessary to further improve the design of the planning system, such as fine-tuning the values of the reward function, as well as the guidance system as a whole. Trials with a clinical population must be conducted to truly evaluate the performance and validity of the planning system. These should be conducted after merging the systems into an autonomous system and after possible improvements identified by the efficacy study are incorporated, such as the integration of clinical data and the re-ordering of step sequences.

### C. Implications of Research

While prior research has been conducted to examine scheduling [10] and guidance through handwashing [14], this research is the first to show that MDPs offer a plausible approach to the modeling of a planning system to guide users with dementia through ADLs. The benefits of using an MDP system include the model's implicit ability to handle stochastic dynamics, the ability to capture tradeoffs between conflicting objectives (both short- and long-term), the potential generalizability to other ADLs, and its role as a precursor to more robust models and techniques for policy construction, like POMDPs.

Although no individuals with dementia were implicated at this stage, the efficacy study not only provided a preliminary assessment of planning device performance through feedback from professionals, but served to identify other areas of research that should be investigated to support the development of the field of assistive technology. The data from this study is valuable for validating and improving the system before use with a clinical population with dementia. Studies with a clinical population should examine the effects of: 1) the administration / misappropriation of positive reinforcement on activity completion, 2) the incorporation of gesturing through the use of visual prompts demonstrating a step on a screen, 3) the exact wording of prompts on care recipient responsiveness, and 4) the implications of permitting multiple regressions in the activity. Such studies would be insightful in attempting to balance aspects such as user compliance, user irritability, and activity completion as a result of device design and implementation.

This paper adds to the mounting evidence that assistive information technology has promising potential to substantially augment caregiving duties. Although the system presented here would not guide a person with dementia in the same manner as a professional caregiver, it is plausible that any deficiencies in guidance efficiency would be compensated by increases in independence for both the user and his/her caregiver. As such, technologies like the one described in this paper are leading the way to achieving aging-in-place.

### REFERENCES

[1]    A. Mihailidis, B. Carmichael, and J. Boger, "The use of computer vision in an intelligent environment to support aging-in-place, safety, and independence in the home," *IEEE Trans. Inform. Technol. Biomed.*, vol. 8, pp. 238-247, Sept. 2004.

[2]    J. Bates, J. Boote, and C. Beverly, "Psychosocial interventions for people with a dementing illness: A systematic review," *J. Adv. Nurs.*, vol. 45, no. 6, pp. 644-658, 2004.

[3]    J. Lindsay, E. Sykes, I. McDowell, R. Verreault, and D. Laurin, "More than the epidemiology of Alzheimer's disease: Contributions of the Canadian study of health and aging," *Can. J. Psychiatry*, vol. 49, no. 2, pp. 83-91, 2004.

[4]    E. Mynatt, I. Essa, and W. Rogers, "Increasing the opportunities for aging in place," in *ACM Conf. Universal Usability (CUU)*, Arlington, VA, 2000, pp. 65-71.

[5]    A. Mihailidis and G. Fernie, "Context-aware devices for older adults with dementia," *Gerontechnology*, vol. 2, pp. 173-189, 2002.

[6]    M. Bourgeois, "Enhancing conversation skills in patients with Alzheimer's disease using a prosthetic memory aid," *J. Appl. Behav. Anal.*, vol. 23, no. 1, pp. 29-42, 1990.

[7]    M. Hoffman, C. Hock, A. Kühler, and F. Müller-Spahn, "Interactive computer-based cognitive training in patients with Alzheimer's disease," *J. Psychiatr. Res.*, vol. 30, no. 6, pp. 493-501, 1996.

[8]    M. Oriani, E. Moniz-Cook, G. Binetti, G. Zanieri, G. Frisono, C. Geroldi, L. De Vreese, and O. Zanetti, "An electronic memory aid to support prospective memory in patients in the early stages of Alzheimer's disease: A pilot study," *Aging Ment. Health.*, vol. 7, no. 1, pp. 22-27, 2003.

[9]    R. Levinson, "The planning and execution assistant and training system," *J. Head Trauma Rehabil.*, vol. 12, no. 2, pp. 769-775, 1997.

[10]   M. Pollack, L. Brown, D. Colby, C. McCarthy, C. Orosz, B. Peintner, S. Ramakrishnan, and I. Tsamardinos, "Autominder: An intelligent cognitive orthotic system for people with memory impairment," *Robot. Auton. Syst.*, vol. 44, pp. 273-282, 2003.

[11]   J. Pineau, M. Montemerlo, M. Pollack, N. Roy, and S. Thrun, "Towards robotic assistants in nursing homes: challenges and results," *Rob. Auton. Syst.*, vol. 42, pp. 271-281, 2003.

[12]   M. Puterman, *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. New York, NY: John Wiley & Sons, Inc., 1994.

[13]   C. Boutilier, T. Dean, and S. Hanks, "Decision-Theoretic Planning: Structural Assumptions and Computational Leverage," *J. of Artificial Intelligence Research,* 11:1-94 (1999).

[14]   T. Dean and K. Kanazawa, "A model for reasoning about persistence and causation", *Computational Intelligence* 5(3):142-150 (1989).

[15]   J. Hoey, R. St-Aubin, A. Hu, and C. Boutillier, "SPUDD: Stochastic Planning using Decision Diagrams," in *Proceedings of the Fifteenth International Conference on Uncertainty in Artificial Intelligence*, Stockholm, 1999, pp. 279-288.

[16]   J. Boger, "Design of a Markov decision process system to guide people with moderate-to-severe dementia through the activity of handwashing," M.A.Sc. thesis, Institute of Biomaterials and Biomedical Eng., Univ. of Toronto, Toronto, Canada, 2004.

[17]   A. Mihailidis, J. Barbenel, and G. Fernie, "The efficacy of an intelligent cognitive orthosis to facilitate handwashing by persons with moderate-

to-severe dementia," *Neuropsychol. Rehabil.,* vol. 14, pp. 135-171, 2004.

[18] F. Marquis-Faulkes, S. McKenna, P. Gregor, and A. Newell, "Scenario-based Drama as a Tool for Investigating User Requirements with Application to Home Monitoring for Elderly People," presented at the HCI International, Crete, Greece, 2003.

[19] A. Mihailidis, "The Development of an Intelligent Cognitive Orthosis to Facilitate Handwashing for Persons with Moderate to Severe Dementia," in *Bioengineering Unit.* Glasgow, Scotland: University of Strathclyde, 2002.

[20] E. Babbie and L. Benaquisto, *Fundamentals of Social* Research. Scarborough, ON: Nelson Thomson Learning, 2002, pp. 152-153.

[21] G. Monahan, "A survey of partially observable Markov decision processes: theory models and algorithms," *Mgmt. Science* 28:1-16, 1982.

[22] J. Boger, P. Poupart, J. Hoey, C. Boutilier, G. Fernie, and A. Mihailidis, "A decision-theoretic approach to task assistance for persons with dementia," to appear*, Proceedings of the Nineteenth International Joint Conference on Artificial Intelligence (IJCAI-05),* Edinburgh, 2005.

**Jennifer N. Boger** received the M.A.Sc. in biomedical engineering (2004) from the University of Toronto in Toronto, Canada and the B.A.Sc. in biological engineering (2002) from the University of Guelph in Guelph, Canada.

She is a senior researcher in the Department of Occupational Therapy, University of Toronto in Toronto, Canada. Her research focuses on the use of artificial intelligence in systems to guide people with dementia through activities of daily living.

**Pascal Poupart** received the Ph.D. in computer science (2005) from the University of Toronto in Toronto, Canada, the M.Sc. in computer science (2000) from the University of British Columbia in Vancouver, Canada and the B.Sc. in mathematics and computer science (1998) form McGill University in Montreal, Canada.

He is an Assistant Professor in the School of Computer Science at the University of Waterloo in Waterloo, Canada. His research focuses on the development of decision theoretic planning and machine learning techniques for autonomous and adaptive systems in various domains including elder care and health informatics.

**Craig Boutilier** received the Ph.D. in computer science (1992) from the University of Toronto in Toronto, Canada.

After eight years at the University of British Columbia, he returned to the Department of Computer Science, University of Toronto, where he is currently Professor and Chair. He has served as consulting professor at Stanford University and is currently a member of the Advisory Board of CombineNet, Inc. His research interests span a wide range of topics, focused on computational and human-interaction issues in decision making uncertainty. Specific interests include Markov decision processes, preference elicitation, game theory and multiagent decision processes, economic models and mechanism design, reinforcement learning, and probabilistic inference.

**Geoff Fernie** received the Ph.D. in mechanical engineering (1973) from the University of Strathclyde in Glasgow, Scotland.

He is Vice President, Research at Toronto Rehabilitation Institute and Full Professor in the Department of Surgery at the University of Toronto. Cross appointments include the Institute of Biomaterials and Biomedical Engineering, the Graduate Department of Rehabilitation Science and the Departments of Mechanical and Industrial Engineering, Physical Therapy and Occupational Therapy at the University of Toronto.

Dr. Fernie was the recipient of the 2002 Jonas Salk Award and the 2003 MEDEC Award in recognition of his contribution to the quality of life of people with disabilities through the development of innovative technologies.

**Alex Mihailidis** received the Ph.D. in biomedical engineering (2002) from the University of Strathclyde in Glasgow, Scotland, and the M.A.Sc. in biomedical engineering (1998) and the B.A.Sc. in mechanical engineering (1996) both from the University of Toronto in Toronto, Canada.

He is an Assistant Professor in the Department of Occupational Therapy and the Institute of Biomaterials and Biomedical Engineering at the University of Toronto in Toronto, Canada. His research focuses on the development of intelligent assistive technology and environments for older adults with cognitive and mobility impairments.