

A Model for Verbal and Non-Verbal Human-Robot Collaboration*

Laëtitia Matignon and Abir-Beatrice Karami and Abdel-illah Mouaddib

GREYC, CNRS UMR 6072

University of Caen Basse-Normandie, France

laetitia.matignon,abir.karami,abdel-illah.mouaddib@info.unicaen.fr

Abstract

We are motivated by building a system for an autonomous robot companion that collaborates with a human partner for achieving a common mission. The objective of the robot is to infer the human's preferences upon the tasks of the mission so as to collaborate with the human by achieving human's non-favorite tasks. Inspired by recent researches about the recognition of human's intention, we propose a unified model that allows the robot to switch accurately between verbal and non-verbal interactions. Our system unifies an epistemic partially observable Markov decision process (POMDP) that is a human-robot spoken dialog system aiming at disambiguating the human's preferences and an intuitive human-robot collaboration consisting in inferring human's intention based on the observed human actions. The beliefs over human's preferences computed during the dialog are then reinforced in the course of the task execution by the intuitive interaction. Our unified model helps the robot inferring the human's preferences and deciding which tasks to perform to effectively satisfy these preferences. The robot is also able to adjust its plan rapidly in case of sudden changes in the human's preferences and to switch between both kind of interactions. Experimental results on a scenario inspired from robocup@home outline various specific behaviors of the robot during the collaborative mission.

Introduction

Robots will likely become increasingly familiar companions in homes and the development of services and assistive robot technology is essential for future personal domestic applications and for providing assistance to an increasing elderly population. This is indicated by the large number of this year participants to the RoboCup@Home league and by several recent successful approaches for assistive human-robot interaction, such as robot systems that interact with the elderly in a nursing home (Pineau et al. 2003) or a wheelchair adaptive assistance (Taha, Miró, and Dissanayake 2008).

The increasing capabilities of these household robots require more and more sophisticated and manifold methods of interaction with humans. These forms of Human-Robot-Interaction vary between applications with different levels

of relationships between the robot and the human. Hoey et al. (2010) presented an automated system to help people with dementia in completing daily life actions like washing hands. The system as an assistant (advisor) recommends to the human what he should do next in case it observes a confusion or a mistake in the sequence of actions. In (Mouaddib et al. 2010), a decision-theoretic approach for adjustable autonomy is proposed where the control of decision making can switch between the human and the robot when needed.

Others focused on applications where a robot collaborates with a human partner to accomplish a common mission, as a collaborative button-pressing task (Breazeal and Hoffman 2004) or assembling a piece of furniture (Dominey, Mallet, and Yoshida 2009). In this form of interaction, commonly known as Human-Robot Collaboration (HRC), the concepts of master-slave or assistant relationship exist no more. The robot and the human act jointly as equals to accomplish their mission and must share the tasks of the mission. The robot companion in this type of collaboration is designed to best satisfy the mission success and both the human's and the robot's preferences.

Assisting a human for a collaborative mission requires the robot to be aware of its partner's preferences upon the tasks so as to effectively satisfy human's desires during the mission. Given that the human's preferences are a state of mind, a robot during its collaboration with a human should be able to infer its partner's preferences, or at least the probable human's preferences. Using those information, the robot should make better decisions for the collaborative mission, that means performing the human's non-favorite tasks. The robot should also adjust its plan rapidly in case of a sudden change in the human's intentions during the mission.

Throughout the rest of this paper, we present a system that infers the human's preferences upon the tasks of the mission and that decides which tasks to perform to effectively satisfy the partners preferences during a common mission. A natural approach to this problem could be to use partially observable Markov decision processes (POMDPs) that can model situations in which an agent acts without complete knowledge of the current state of its environment. Using POMDPs allows the robot to compute beliefs about the preferences of the human based on observations and to act accordingly. They also provide policies that are adjustable to all possible changes in the human's preferences. The paper

*This work was supported by the AMORCES NRA (French National Research Agency) project (ANR-07-ROBO-0004). Copyright © 2010, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

is organized as follows: we briefly introduce the POMDP framework followed by a survey on related work about the recognition of human’s intention for HRC. Then we present how we used POMDPs to formalize our unified model for verbal and non-verbal HRC. Finally, we propose a human-robot interaction scenario inspired from RoboCup@Home applications and some results we obtained using an approximate POMDP solver that is appropriate to the problem.

Background on Markov Decision Processes

A **POMDP** (Cassandra, Kaelbling, and Littman 1994) is a partially observable Markov Decision Process represented by a tuple $\langle S, A, T, Z, O, R, b_0 \rangle$ where: S is a finite set of states, A is a finite set of actions, Z is a finite set of observations, $T : S \times A \times S \rightarrow [0, 1]$ is a state transition probability distribution, $O : S \times A \times S \times Z \rightarrow [0, 1]$ is a discrete probability distribution over Z , $R : S \times A \rightarrow \mathfrak{R}$ is the reward function and $b_0(s)$ is a probability of being in state s at time $t = 0$.

Given that the state is not directly observable, the system instead maintains a belief distribution over S . $b_t(s)$ is the probability that the system is in state s at time t , given the history of all observations/actions the agent received/affected and the initial belief state b_0 , $b_t(s) = Pr(s_t = s | z_t, a_{t-1}, z_{t-1}, \dots, a_0, b_0)$. Knowing the last action a_{t-1} and last observation z_t , the agent calculates a new belief state at each time step t by applying the following belief update function:

$$\begin{aligned} b_t^{az}(s') &= \frac{Pr(z_t | s', a_{t-1}, b_{t-1}) Pr(s' | a_{t-1}, b_{t-1})}{Pr(z_t | a_{t-1}, b_{t-1})} \\ &= \frac{\sum_{s \in S} O(s, a_{t-1}, s', z_t) T(s, a_{t-1}, s') b_{t-1}(s)}{Pr(z_t | a_{t-1}, b_{t-1})} \end{aligned} \quad (1)$$

where $Pr(z_t | a_{t-1}, b_{t-1})$ acts as a normalizing constant.

The goal of a POMDP planning is to find a sequence of actions maximizing the expected sum of rewards. Such a plan is called a policy π . An optimal policy specifies for each b the optimal action to execute at the current step assuming the agent will also act optimally at future time steps. The value of an optimal policy π^* is defined by the optimal value function V^* that satisfies the Bellman optimality equation:

$$V^*(b) = \max_{a \in A} \left[\sum_{s \in S} R(s, a) b(s) + \gamma \sum_{z \in Z} Pr(z | b, a) V^*(b^{az}) \right]$$

Related Work

Several successful approaches are interested in inferring the human’s intention to enhance the collaboration between the robot and the human and thus effectively satisfy human’s desires during the mission. They vary between applications with implicit or explicit communication for the collaboration. Implicit communication is based on a non-verbal intuitive HRC and explicit communication concerns spoken dialog systems.

Intuitive Human-Robot Collaboration

Recent researches are motivated by building a system for a collaborative robot assistant sharing a mission with a human partner without explicit communication nor a shared plan. They are based on an intuitive HRC inspired by the behavior of people in human-human collaboration. In this way, the robot is able to read the human’s non-verbal cues to infer his intention. The key to the implicit non-verbal interaction is to assume that human actions are directly caused by his intentions, and thus the human’s intention is estimated mainly with the observation of the human actions. A probabilistic approach is used to compute the probable human’s intentions from the observed human actions.

Schmid, Weede, and Wörn (2007) proposed to estimate the human’s intention by observing his actions and using Hybrid Dynamic Bayesian Networks. An appropriate robot task is selected without requiring an explicit verbal communication. They also include pro-active actions, whose objective is to trigger a clarifying reaction from the human so as to remove any uncertainty about the human’s intention.

Other approaches use a POMDP model to predict the human’s intention for an elevator riding task (Broz, Nourbakhsh, and Simmons 2008), wheelchair navigation (Taha, Miró, and Dissanayake 2008) or cleaning mission (Karami, Jeanpierre, and Mouaddib 2010). Notably, Karami et al. (2010) developed a model that helps the robot making the best action towards the mission achievement considering its belief over the human’s intentions and the priority of human comfortability. In order to link the human actions to the possible human’s intentions, they calculate for each possible human’s action, a value that corresponds to the worth of this action toward a possible intended task. To do that, the authors compute human MDP policies, each of them includes one agent (the human) and one task (his intention). Each human MDP policy calculates an action value function that gives a value for each pair (state, human action). Having those values for all possible intentions, the robot then calculates the probability of observing each human action given his intention and uses it in the POMDP model to build a belief over all possible human’s intentions given the observed human actions.

Spoken Dialog Systems

Others are interested in building spoken dialog systems that help humans achieve their goals via speech communication. In those systems, the dialog agent should maintain an efficient and natural conversation with the human. The objective of the agent is to interpret the dialog accurately so as to discover the human’s intention. With this in view, it must decide which sequence of actions to follow to gather accurate information about the human’s intention and then, which final decision to take to match this intention. Actions during the dialog might include asking a question, confirming a human’s intention, or querying a database.

The dialog management is complex for several reasons. First the system observes the human utterances during the dialog via automated speech recognition and language parsing, which are imperfect technologies corrupting the observations. Second, each human utterance (even if it could be

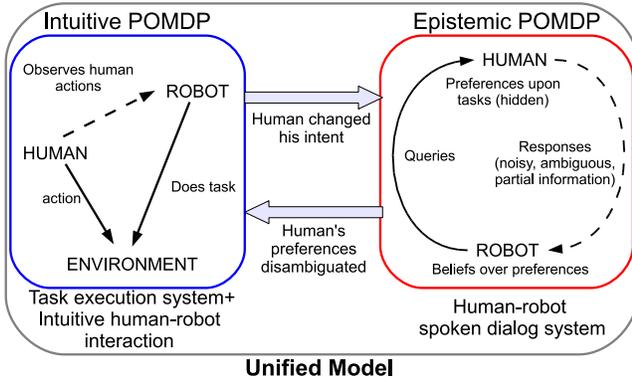


Figure 1: The architecture of our unified model for HRC.

observed accurately) provides incomplete information about the human’s intention, so the system must assemble evidence over time. Thirdly, because the human might change his intention at any point during the dialog, inconsistent evidence could either be due to speech recognition error or due to a modified intention. Thus the challenge for the dialog agent would be interpreting conflicting evidence in the human utterances to estimate his intention.

Finally, the agent must make trade-offs between the cost of gathering additional information (increasing its certainty of the human’s intention, but prolonging the conversation) and the cost of making a final decision that might not match the human’s intention. That is, the system must perform planning to decide which sequence of actions to take to best disambiguate the human’s intentions and achieve the human’s goal. For all of these reasons, the spoken dialog system problem can be regarded as planning under uncertainty. Many researchers have found POMDP frameworks suitable for designing a robust dialog agent in spoken dialog systems. These researches range from robot system that interacts with the elderly in a nursing home (Pineau et al. 2003), automated system that assists people with dementia (Hoey et al. 2010), flight agent assisting the caller to book a flight ticket (Williams 2006; Young et al. 2010) or a wheelchair directed by her patient (Doshi and Roy 2008).

Unified Model for Human-Robot Collaboration

Existing approaches to infer the human’s intention so as to enhance the collaboration between the robot and its human partner use either intuitive human-robot interaction, or explicit verbal interaction. We propose to unify these both kinds of interaction in a unified model for HRC. In the framework of a collaborative mission, our unified model should give the robot the opportunity to switch accurately between inferring the human’s preferences using queries and inferring the possible change in human’s intentions thanks to non-verbal cues. Then it decides which tasks to perform to effectively satisfy the human’s preferences.

The architecture of our unified model is shown in Figure

1. It is made up of a human-robot spoken dialog system regarding as an epistemic interaction. Indeed, human’s preferences are not observable. Robot’s actions during this interaction are queries asked to the human with potentially noisy or ambiguous answers. The robot must choose queries that supply information about the preferences such that the series of questions aim at disambiguating the human’s preferences. Using information obtained during the dialog, i.e. observed responses of the human, the robot builds a belief over its partner’s preferences despite uncertainty in the observed responses. Once sufficiently certain and based on this belief, the robot should be able to decide which tasks to perform to effectively satisfy the human’s preferences and then switch to the task execution system to apply those decisions.

However, human’s preferences may change over the course of the mission and one challenging issue is to detect this change. This can be done by querying the human and assembling conflicting evidence in the human utterances. But the robot should avoid to constantly ask queries since too much questions could annoy the human; the dialog must only be used for gathering information about the human’s preferences. We propose to use intuitive human-robot collaboration during the task execution to detect the change of preferences. Thus, the robot will avoid a systematic return to the dialog to check this change. Once in the task execution system, the beliefs over human’s preferences (built during the dialog) should be reinforced by an intuitive human-robot interaction based on observed human actions. Indeed, the observed human actions may bring the robot information concerning the current human’s intention (Karami, Jeanpierre, and Mouaddib 2010).

Thus our unified model allows the robot to switch from an explicit verbal interaction, that aims at disambiguating the human’s preferences, to the task execution system where the robot achieves the tasks of the mission which are assigned to it and at the same time, checks if the human has changed his preferences. In case of a change in the human’s preferences, the robot returns to the dialog system to infer the new preferences of the human; otherwise, it continues to execute tasks according to its belief over the human’s preferences. Therefore the unified model provides an accurate switch over both kinds of verbal and non-verbal interaction that will be illustrated in our results.

The Unified POMDP Model

We now present a specific POMDP detailed model that captures the desiderata described above relating to our unified verbal and non-verbal HRC. Our scenario, inspired from RoboCup@Home applications, consists of a robot and a human in a house. They both share a mission M that includes a list of N tasks $\{t_1, t_2, \dots, t_N\}$. Each of the tasks is matching a specific kind of housework as gardening, cooking or cleaning. The human has preferences upon the tasks modeled as his internal state s_h which may change over the course of the mission. Preferences are, for each task t_i of the mission:

- $s_h(t_i) = 0$ if the human would rather do the task t_i ;
- $s_h(t_i) = 1$ if the human would rather the robot did the task t_i ;

- $s_h(t_i) = 2$ if the human has not yet decided his preference upon the task t_i .

Next, we formalize our unified model using the POMDP framework. We will assume that the state, action and observation sets are all discrete and finite.

States

Our state space brings together the set of human’s preferences upon the tasks of the mission (non-observable) and the status of each tasks (*done* or *not yet done*) that is observable. The human’s preferences can be 0, 1 or 2 for any tasks of the mission that has the status *not done*. The state s is then characterized by a function that associates, at each task $(t_i)_{i \in [1, N]}$ of the mission, either the human’s preference: $s(t_i) \in \{0, 1, 2\}$; or the status if t_i is *done*: $s(t_i) = \text{done}$.

For instance, a POMDP state for $N = 5$ tasks, can be $s = \langle 1, \text{done}, 2, 0, \text{done} \rangle$ that means t_2 and t_5 are done, the human would rather do t_4 , would rather the robot did t_1 and has not yet decided his preference upon the task t_3 .

Actions

Possible actions for the robot include queries to the human asked during the dialog. The robot can choose from three kinds of queries: it can choose to ask a general question such as “Which task should I do?”, to confirm a preference upon a specific task t_i such as “Should I do the task t_i ?” and to greet the human and ask “How can I help you?”. The robot can also choose to achieve a task of the mission. We assume the robot has a list of predefined policies to accomplish correctly each task of the mission. The action $do(t_i)$ then leads the robot to follow the corresponding policy. The robot may also choose to wait, for instance because remaining tasks are preferred by the human. The robot action set is: $A = \{\text{wait}, do(t_1), \dots, do(t_N), \text{confirm}(t_1), \dots, \text{confirm}(t_N), \text{ask}, \text{greet}\}$.

Observations

The observation set includes different ways of partially or fully communicating the human’s preferences. In reply to a general question or a greeting, observations consist of N observations $\{\text{prefdo}(t_1), \dots, \text{prefdo}(t_N)\}$ associated with each of the N tasks plus the $\text{prefdo}(\emptyset)$ observation. Observations *yes* and *no* stand for positive and negative confirmations in response to confirm queries. Observations *not yet* stands for a not yet decided response. The robot may also observe *nothing*. Finally, the robot may observe $hdid(t_i)$ when the human has just achieved the task t_i . The robot observation set is: $Z = \{hdid(t_1), \dots, hdid(t_N), \text{nothing}, \text{prefdo}(t_1), \dots, \text{prefdo}(t_N), \text{prefdo}(\emptyset), \text{yes}, \text{no}, \text{not yet}\}$.

Transition Function

The transition function $T(s, a, s')$ shows what the next state s' is likely to be given the current state s and the action a . The effects of the robot action a on a state s are relatively clear: when the robot does a task, the task status changes to *done*. Other actions like queries or wait action do not modify the state.

However, the transition from state s to s' is not only defined by the robot action, but also by the human actions and intentions. Indeed, as we assume the human has a small probability to change their preferences mid-dialog, the human’s preferences might change independently of the robot action. As well, the task status that can change to *done* when the human did a task. We suppose that: the human does only tasks he would rather do, i.e. tasks whose preferences are 0 or 2; the human will keep same preferences with a probability pKI ; the human might do a task he is intended with a probability $pHDo$; otherwise the human changes his preferences upon tasks that are *not yet done* to another preferences chosen uniformly¹. Thus:

- $T(s, a, s' = s) = pKI$
- $T(s, a, s' \in \text{hDo}(s)) = \frac{pHDo}{\text{size}(\text{hDo}(s))}$
- $T(s, a, s' \in \text{hI}(s)) = \frac{1-pKI-pHDo}{\text{size}(\text{hI}(s))}$ if $\text{size}(\text{hDo}(s)) \neq 0$
- $T(s, a, s' \in \text{hI}(s)) = \frac{1-pKI}{\text{size}(\text{hI}(s))}$ if $\text{size}(\text{hDo}(s)) = 0$

where $\text{hDo}(s)$ is the set of all reachable states from s when the human does one task that he prefers; and $\text{hI}(s)$ is the set of all possible permutations of preferences upon *not yet done* tasks in s . We obtain the same probabilities when $a = do(t_i)_{i \in [1, N]}$ except that the status of the task t_i in s' is *done*.

Observation Function

Based on the most recent action a and the current and future state (s, s') of the system, the robot has a model of which observation z it may receive. First the observation function $O(s, a, s', z)$ gives in a deterministic way the observation $z = hdid(t_i)_{i \in [1, N]}$ when the human just did a task t_i . The robot also observes *nothing* when it waits or does a task: $O(s, a \in \{\text{wait}, do(t_i)_{i \in [1, N]}\}, s', z = \text{nothing}) = 1$.

The observation function $O(s, a, s', z)$ also encodes both the words the human is likely to use to reply to the queries and the speech recognition errors that are likely to occur. We suppose speech recognition errors are different according to the kind of queries. If the robot made a general query or a greeting, then it observes the right answer with a probability $pAsk$. Thus, when $a \in \{\text{ask}, \text{greet}\}$:

- $O(s, a, s', z = \text{prefdo}(\emptyset)) = pAsk$ if $\text{nbRD}_O(s') = 0$ and $\text{nbNYet}(s') = 0$
- $O(s, a, s', z = \text{not yet}) = pAsk$ if $\text{nbRD}_O(s') = 0$ and $\text{nbNYet}(s') \neq 0$
- $O(s, a, s'(t_i) = 1, z = \text{prefdo}(t_i)) = \frac{pAsk}{\text{nbRD}_O(s')}$

where $\text{nbRD}_O(s')$ is the number of tasks in s' that the robot can do according to human’s preferences and $\text{nbNYet}(s')$ the number of tasks in s' upon which the human is not decided yet. If the robot made a confirm query, then it observes the right answer with a probability $pConf$. Thus, when $a = \text{confirm}(t_i)$:

- $O(s, a, s'(t_i) = 1, z = \text{yes}) = pConf$

¹Once the human has decided his preference upon t_i ($s(t_i) \in \{0, 1\}$), he cannot return to a not yet decided preference ($s(t_i) = 2$), yet $s(t_i)$ can switch between 0 or 1.

state s	action a	$R(s, a)$
First step of dialog	<i>greet</i>	-1
First step of dialog	\neq <i>greet</i>	-1000
Not first step of dialog	<i>greet</i>	-1000
$s(t_i) \in \{0, 2, 3\}$	<i>do</i> (t_i)	-1000
$s(t_i) = 1$ and $\exists j s(t_j) = 2$	<i>do</i> (t_i)	-1000
$s(t_i) = 1$ and $\forall j s(t_j) \neq 2$	<i>do</i> (t_i)	300
$s(t_i) \in \{1, 2\}$	<i>confirm</i> (t_i)	-50
$s(t_i) \in \{0, 3\}$	<i>confirm</i> (t_i)	-10
Not first step of dialog	<i>ask</i>	-20
$\exists j s(t_j) \in \{1, 2\}$	<i>wait</i>	-30
$\forall j s(t_j) \in \{0, 3\}$	<i>wait</i>	30

Table 1: Reward function for our POMDP model.

- $O(s, a, s'(t_i) \in \{0, 3\}, z = no) = pConf$
- $O(s, a, s'(t_i) = 2, z = not\ yet) = pConf$

In both queries *ask* and *confirm*, the robot observes in addition to the right answer an arbitrary response uniformly at random from the remaining possible replies *prefdo*(t_i), *prefdo*(\emptyset), *yes*, *no*, *not yet*, *nothing*.

Reward Function

The reward function specifies what the “right” actions are in different states. This function is given in Table 1. The robot must greet the human only at the first step of the dialog. It receives a higher penalty for an incorrect confirmation than for a correct one. The reward function also specifies how much the human is willing to tolerate *ask* versus *confirm* queries thanks to the choice of the reward for a general query. The robot is penalized for doing a task already done or a human’s favorite task. It is also penalized if it does a task although there remains undecided tasks or if it waits although it would better do a task or query the human. Finally, the robot gets a high reward when it does a human’s non-favorite task and all the human preferences upon the tasks have been given. It also gets a high reward when it waits while all remaining tasks are human’s favorite tasks.

Belief State and Human’s Intent Change

Initially, the belief state is uniform among all states where all the tasks are *not yet done*. The belief of the robot over the state at time t is $b_t(s)$ and is a probability distribution over the human’s preferences upon remaining tasks of the mission. It is updated given eq. 1.

When the robot is in the task execution system, it can detect a change in the human’s preferences thanks to the observation of the human actions. Indeed, we use an intuitive HRC method that builds a belief over all possible human’s intentions given the observed human actions (Karami, Jeanpierre, and Mouaddib 2010). If the beliefs over human’s intentions computed with the intuitive method do no match the beliefs calculated during the dialog, then the robot has detected a change in the human’s intentions and should switch to the spoken dialog system. To do that, the belief state is reinitialized uniformly at random.

Experimental Results

The POMDP Solver

Finding an optimal policy over the entire belief space for a finite horizon POMDP is PSPACE-complete, that is why a lot of approximate solvers were presented for solving POMDPs. We have chosen an approximate POMDP solver to calculate a policy that profits from the topological structure in our scenario. Indeed, given that the action *do a task* (by robot/human) is definitive, the sets of *done/not yet done* tasks are topological sets that can be followed. For example, the moment the task t_1 is done by the human or the robot, the system will move to another sub-belief-state space that represents all possibilities with the task t_1 *done*. After this point, the sub-belief-state space that represents all possibilities with the task t_1 *not yet done* will no more be accessible. The approximate and topological chosen solver is Topological Order Planner (TOP) (Dibangoye et al. 2009): during its policy calculation, it creates layers with pairs of (states, belief states) and the possible paths between layers and between pairs of the same layer. Problems that create higher number of layers are more interested to be solved with TOP.

Parts of dialogs

We chose a scenario with a mission composed of 5 tasks ($|S| = 1025, |A| = 13, |O| = 15$). TOP created 244 layers and 1200 minutes were required to calculate the policy including the TOP solver preparations for creating layers and paths values. The model parameters used were $pKI = 0.9, pHDo = 0.05, pAsk = 0.7$ and $pConf = 0.9$. Table 2 shows a dialog between the human and the robot during the achievement of the mission and following the computed policy. Those parts outline various specific behaviors of the robot during the collaboration. We note in step 4 the reinforcement of the preference upon the task 0 which is due to the increased probability of the change of the human’s intention since the last confirmation (step 1). In step 10 and 13, we notice that the robot has switched to the dialog system for two different reasons. At step 10, after finishing all the non-preferred tasks, the robot checks the possible change in the human’s preferences that might have occurred during the execution period. At step 13, the robot receives an observation from the intuitive system that declares an observed human’s intent change. For this, it reinitializes all the remaining tasks to an equal probability and switches to the dialog system to infer the new preferences.

Conclusion

In this paper, we have presented a unified model that allows an autonomous robot companion that collaborates with a human partner to infer the human’s preferences and to switch accurately between verbal (epistemic system) and non-verbal (intuitive system) interactions. We presented an example of the POMDP policy for 5 tasks that shows how the robot optimally switches between the epistemic and the intuitive systems, in addition to its good choice of type of queries at each dialog interaction. We aim to improve these results by exploring how robust our solution is to varying levels of model uncertainty and by trying missions

	POMDP HIDDEN STATE	ROBOT ACTION	HUMAN ACTION/REPLY
Greeting at the beginning of the dialog:			
1	< 1, 2, 0, 2, 0 >	ROBOT: Hello, how can I help you ?	HUMAN: I'd rather you do the task 0.
2	< 1, 2, 0, 2, 0 >	ROBOT: Should I do the task 2?	<i>Human does task 2.</i>
The robot confirms all remaining (not yet done) and ambiguous tasks:			
3	< 1, 0, done, 1, 0 >	ROBOT: Should I do the task 4?	HUMAN: No.
4	< 1, 0, done, 1, 0 >	ROBOT: Should I do the task 0?	HUMAN: Yes.
5	< 1, 0, done, 1, 0 >	ROBOT: Should I do the task 3?	HUMAN: Yes.
6	< 1, 0, done, 1, 0 >	ROBOT: Should I do the task 1?	HUMAN: No.
Switch from the epistemic interaction to the intuitive interaction to do the human's non-preferred tasks:			
7	< 1, 0, done, 1, 0 >	<i>Robot does task 0.</i>	
8	< done, 0, done, 1, 0 >	<i>Robot does task 3.</i>	
9	< done, 0, done, done, 0 >	<i>Robot waits.</i>	
Switch from the intuitive interaction to the epistemic interaction to check if the human has changed his preferences:			
10	< done, 0, done, done, 0 >	ROBOT: Which tasks should I do?	HUMAN: I'd rather you do nothing.
Switch from the epistemic interaction to the intuitive interaction:			
11	< done, 0, done, done, 0 >	<i>Robot waits.</i>	
12	< done, 1, done, done, 1 >	<i>Robot waits.</i>	INTUITIVE SYSTEM: <i>Change in human's intentions.</i>
Switch from the intuitive interaction to the epistemic interaction responding to the intuitive system observation:			
13	< done, 1, done, done, 1 >	ROBOT: Should I do the task 4?	HUMAN: Yes.
14	< done, 1, done, done, 1 >	ROBOT: Should I do the task 1?	HUMAN: Yes.
Switch from the epistemic interaction to the intuitive interaction to do the human's non-preferred tasks:			
15	< done, 1, done, done, 1 >	<i>Robot does task 1.</i>	
16	< done, done, done, done, 1 >	<i>Robot does task 4.</i>	
17	< done, done, done, done, done >		

Table 2: A dialog example between the human and the robot during the achievement of the mission.

with more tasks; indeed, POMDP policies for such missions could be calculated but with time periods that exceed a day. We also plan in future work to structure the problem such as the mission is divided into a group of tasks, each of them includes a number of similar kind of tasks, hoping by this to solve problems with much higher number of tasks. We also plan to enhance our unified model by adding the possibility that the robot proposes to do accurate tasks instead of waiting even if the human's preferences were to do all the remaining tasks, in cases where it believes that it is the best for the mission achievement.

References

- Breazeal, C., and Hoffman, G. 2004. Robots that work in collaboration with people. In *Proceedings of AAAI Fall Symposium on The Intersection of Cognitive Science and Robotics: From Interfaces to Intelligence*.
- Broz, F.; Nourbakhsh, I. R.; and Simmons, R. G. 2008. Planning for human-robot interaction using time-state aggregated pomdps. In *AAAI'08*, 1339–1344.
- Cassandra, A. R.; Kaelbling, L. P.; and Littman, M. L. 1994. Acting optimally in partially observable stochastic domains. In *AAAI'94*, 1023–1028.
- Dibangoye, J. S.; Shani, G.; Chaib-draa, B.; and Mouaddib, A.-I. 2009. Topological order planner for pomdps. In *IJCAI '09*, 1684–1689.
- Dominey, P. F.; Mallet, A.; and Yoshida, E. 2009. Real-time spoken-language programming for cooperative interaction with a humanoid apprentice. *I. J. Humanoid Robotics* 6(2):147–171.
- Doshi, F., and Roy, N. 2008. The permutable pomdp: fast

solutions to pomdps for preference elicitation. In *AAMAS '08*, 493–500.

Hoey, J.; Poupart, P.; von Bertoldi, A.; Craig, T.; Boutilier, C.; and Mihailidis, A. 2010. Automated handwashing assistance for persons with dementia using video and a partially observable markov decision process. *Computer Vision and Image Understanding* 114(5):503–519.

Karami, A.-B.; Jeanpierre, L.; and Mouaddib, A.-I. 2010. Human-robot collaboration for a shared mission. In *HRI '10*, 155–156.

Mouaddib, A.; Zilberstein, S.; Beynier, A.; and Jeanpierre, L. 2010. A decision-theoretic approach for adjustable autonomy. In *ECAI '10*.

Pineau, J.; Montemerlo, M.; Pollack, M.; Roy, N.; and Thrun, S. 2003. Towards robotic assistants in nursing homes: Challenges and results. *Robotics and Autonomous Systems* 42(3-4):271–281.

Schmid, A. J.; Weede, O.; and Wörn, H. 2007. Proactive robot task selection given a human intention estimate. In *RO-MAN '07*, 726–731.

Taha, T.; Miró, J. V.; and Dissanayake, G. 2008. Pomdp-based long-term user intention prediction for wheelchair navigation. In *ICRA '08*, 3920–3925. IEEE.

Williams, J. D. 2006. Scaling pomdps for dialog management with composite summary point-based value iteration. In *AAAI Workshop for Spoken Dialogue Systems*.

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