Toward Integrating Theory of Mind into Adaptive Decision-Making of Social Robots to Understand Human Intention

O. Can Görür
German-Turkish Advanced Research Center for ICT, and Technische Universität Berlin, Germany
orhan-can.goeruer@gt-arc.com

Benjamin Rosman
Mobile Intelligent Autonomous Systems, CSIR, and the School of Computer Science and Applied Mathematics, University of the Witwatersrand, South Africa

Guy Hoffman
Sibley School of Mechanical and Aerospace Engineering, Cornell University, USA

Şahin Albayrak
Distributed Artificial Intelligence Laboratory (DAI-Labor), Technische Universität Berlin, Germany

Abstract — **We propose an architecture that integrates *Theory of Mind* into a robot's decision-making to infer a human's intention and adapt to it. The architecture implements human-robot collaborative decision-making for a robot incorporating human variability in their emotional and intentional states. This research first implements a mechanism for stochastically estimating a human's belief over the state of the actions that the human could possibly be executing. Then, we integrate this information into a novel stochastic human-robot shared planner that models the human's preferred plan. Our contribution lies in the ability of our model to handle the conditions: 1) when the human's intention is estimated incorrectly and the true intention may be unknown to the robot, and 2) when the human's intention is estimated correctly but the human doesn't want the robot's assistance in the given context. A robot integrating this model into its decision-making process would better understand a human's need for assistance and therefore adapt to behave less intrusively and more reasonably in assisting its human companion.**

I. Introduction

Social robots should understand personal needs and preferences of individuals, and adaptively and non-intrusively assist in order to meet those needs and support the longevity of their usage [1]. However, studies have shown that robots have deficient capacity to adapt to humans' changing affective and motivational states (to empathize), which results in a failure to keep users engaged over repeated and long-term interactions [2]. For the purpose of reasoning over humans' mental states, a *Theory of Mind (ToM)* approach, being the ability to attribute mental states such as intentions, beliefs and desires to others and take them into account, have received significant attention [3]. Previous work on ToM in robotics has mainly focused on visual perspective taking and belief management in understanding the world from the interacting person's point of view [3], [4]. Utilization of this information has been shown to improve human-robot teamwork significantly, leading to more effective and natural collaboration [5], [6].

In human-robot collaboration tasks with a shared goal, it is crucial to infer the human's plan in order for the robot to quickly adapt to observed behaviors. For this purpose, more recent approaches have focused on reverse engineering human ToM, where they show that a human's intents and plans can be inferred by observing the human's actions [7], [8]. However, they are mostly limited to the recognition of human states, and are yet to extend to adaptively making decisions based on these states. It has recently been stated that there is still a gap between the estimation of human mental states, e.g., intentions, beliefs, plans etc., and their explicit use in shared plan execution in human robot collaboration [9]. A new approach has been proposed targeting this gap [9], where the robot estimates its human partner's belief on the state of joint actions of a shared plan during the execution, so as to decipher and adapt to the changing human-robot work division. Although this approach inspires our study, it assumes that the human is committed to the given goal and the belief estimation is a fully observable and deterministic process. In general, most of the available approaches on the modeling of a human plan make two common assumptions:

i) All of the actions a human executes are relevant to a goal or an intention that is known to the experimenter (or the robot) [7]–[11],

ii) Humans always accept a robot's assistance when offered [9], [12].

In reality, a human's various desires along with their dynamic emotional states could result in stochastic intentions, behaviors and expectations over the course of repeated interactions. Therefore, a robot making these assumptions could misinterpret human actions, which may result in unreasonable and intrusive robot behaviors. As an example, a robot may infer that a human needs an object and offer assistance if it catches the human's gaze on that object. However, in reality the human's mental state behind this gaze could be, in contrast to assumption *i*, any intention related to the object other than taking it, or could even be irrelevant to that object. Moreover, even if the human wants the object, they may not want the robot to pick it up, for any of several reasons such as a distrust for the robot and belief that it could damage the object (in contrast to assumption *ii*). The premise of this study is, for the first time to the best of our knowledge, to devise an architecture that is able to reason about these cases.

Our study aims to target this gap in the literature by removing the given *assumptions i* and *ii* in modeling a human's plan over the course of repeated human-robot collaboration. We incorporate human variability in their emotional and intentional states into estimating a human's belief on the state of their own actions and plans towards shared goals. The contribution of this model is the ability to handle the following *two conditions*:

1) When the human's intention is estimated incorrectly and the true intention may be unknown to the robot's knowledge base, and
2) When the human's intention is estimated correctly but the human doesn't want the robot's assistance in the current context, i.e., changing emotional states or the human's task relevant distrust for the robot.

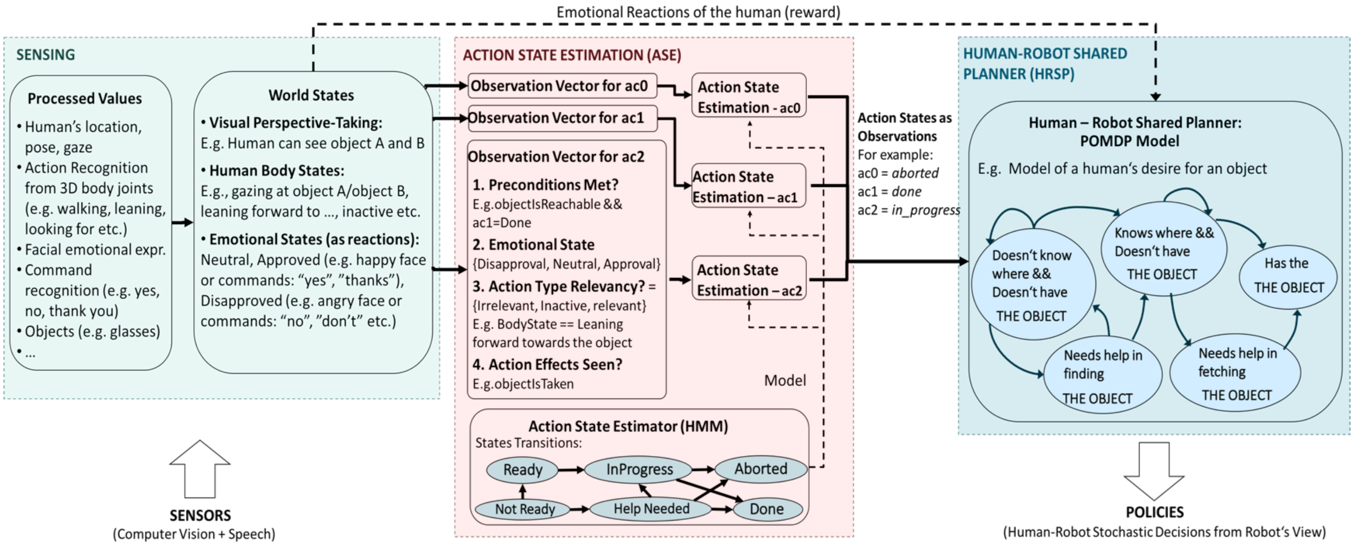

Fig. 1. System architecture of human-robot shared planning

In developing our model, this research first describes a mechanism for stochastically estimating a human belief over the state of all actions that the human could possibly be executing (targeting *condition 1*) while also incorporating human emotional states into the process as reactions to evaluate these estimates (targeting *condition 2*). We then integrate this information into a novel stochastic human-robot shared planner that models the human's preferred plan. In this short paper, we make the limiting assumption that all observations are discrete, and directly available to the system (through a sensing component) in order to focus on the modeling of human mental states and their effects on actions and plans. Finally, this model is integrated into a simulated robot's decision-making process to show its value in the robot's understanding of a human's true need for assistance more accurately, and in so doing adapt to behave less intrusively and more reasonably in assisting its human companion.

II. Methodology

We propose an architecture, shown in Fig.1, consisting of three main building blocks, which are *Sensing, Action State Estimation (ASE)* and the *Human-Robot Shared Planner (HRSP)*. The architecture takes raw sensory data and generates stochastic policies as human-robot shared decisions from the robot's point of view. Although the focus of this work is on the *ASE* and *HRSP* components, which together form the core of our ToM approach, the *Sensing* component is also shown under Fig.1 as it provides the input which drives the remainder of the architecture. Our core intuition in this architecture is that estimating the intention of the human based solely on a single snapshot of activity, that is, taking the human's positional and visual perspective, is insufficient to handle the *two conditions* described in Section I. Loosely inspired by Devin and Alami [9], we hypothesize that one must track the human activity and estimate a belief over the state of the actions the human could possibly be doing. These action states considered in our study are: "ready", "not ready", "in progress", "help needed", "aborted", "done". In addition, we introduce the novel approach of integrating human emotional states into the process, in the form of the reactions of the human to the robot's judgements on each action. By doing so, as the human action progresses and interacts with the robot over the course of a nondeterministic plan, the robot is able to reason about hidden human mental states with more confidence.

The architecture, through the *ASE* component, first estimates the states of possible actions the human could be executing (modeled as a belief distribution in the *HRSP*) using the observations from the world. Then, in the *HRSP,* these are used to further reason about the human's dynamic plans by estimating which action the human actually needs help with, or which action the human is capable of executing towards achieving their intentions. Starting with the *ASE* component, we first construct an observation vector for each action to estimate its state. As shown in Fig.1, this is done in parallel to track each possible action the human could be executing. By doing so, we preserve dynamic and stochastic nature of human intentions behind these observed actions as well as the uncertainties available in the observations, which are modelled in the *HRSP.* As an example of the observation uncertainties, a computer vision system may unable to determine if a human gaze is on the glasses or on the TV remote (defining two different actions in our system) when they lie very close to each other.

Table 1 The observation vector for action state estimation

Feature #	Descriptions	Values
1	If preconditions are met	{0: No,1; Yes}
2	Emotional state of the person as reactions to the robot	{0: Neutral, -1:Disapproval, 1: Approval}
3	If currently recognized action type is relevant	{0: Inactive, -1:Irrelevant, 1:Relevant}
4	If action effects are observed	{0:No, 1:Yes}

The observation vectors, as detailed in Table 1, are obtained from the inputs in the *Sensing* component, and are described by the *World States* which provide the physical world state from the human's perspective. The *World State* values are semantically compared with each of the actions' (i.e., the actions that are known to the robot) relevant descriptions to construct the action specific observation vectors. As an example, the construction of an observation vector of *ac2*, an action that defines the human taking an object, is given in Fig.1. The observation features #1 and #4 correspond to the basic semantic descriptors of an action, which check if preconditions of the action are met and the effects of the action are observed, respectively. The feature #3, on the other hand, compares the detected body states of the human (see the *World States* in Fig.1) with the physical activity type an action requires during its process. For example, if the momentary body state of a human is detected as "gazing at a TV remote", the feature #3 of the action "human taking the glasses" will be "irrelevant". As an example, given the possible action states, which are also demonstrated in *Action State Estimator* in Fig.1, a robot might estimate that an action is "in progress" if feature #1 is "yes" and feature #3 is "relevant".

Since our ultimate goal is to estimate in which actions a human needs help or does not want a robot's help, feature #2 is incorporated so as to evaluate the robot's inference about the human's desire for help with an action. This evaluation is realized from the emotional reactions of the person to the robot's behavior. This feature mainly contributes to the reasoning of the robot on the *two conditions* mentioned, i.e., when the intention is estimated incorrectly, or when the estimation is correct but the human doesn't want help. The robot always assumes feature #2 is "neutral" before interacting with the human. Then, following a possible estimation of an action's state as "help needed" (as shown under *Action State Estimator* in Fig.1), the robot would offer its assistance on that action, as a policy generated from the *HRSP*, and receives feedback from the human. If the human says "no" it is reflected in feature #2 as a "disapproval", which may lead to the action state changing to "aborted" (modeled as "human intention is something different" under the *HRSP,* targeting *condition 1*) or "in progress" (modeled as "human wants to do it on his/her own" under the *HRSP*, targeting *condition 2*). It is worth noting that the estimation of "help needed" for an action does not necessarily result in the robot offering its assistance. As stated, the *HRSP* has a belief distribution on the possible human actions and it plans for being non-intrusive, i.e., negative rewards are acquired from the human disapprovals, therefore the robot needs to be highly certain on the human's need for help before offering its assistance.

We devise a Hidden Markov Model (HMM), shown as *Action State Estimator* in Fig.1, for the action state estimation mainly because it is a stochastic state transition model. An HMM is suitable for this purpose since the action states correspond to the hidden intentions of the human, such that the robot can only observe them indirectly. Moreover, the states possess the Markov property (i.e., the future is independent of the past given the present) and we do not predict the future states of the action. The possible states of an action in our model are: {ready, not ready, in progress, help needed, aborted, done}. We use "aborted" instead of "failed" as it contains the information of either a possible failure or a wrong estimation of the robot from the beginning, i.e., the human was not executing that action at all. Simply put, "aborted" tells us that the human does not want to realize this action.

An example of action state estimation from an observation sequence that has vectors structured as given in Table 1 is shown in Table 2, where the first bit of each observation is feature #1. We note that the observation sequence and the corresponding states given in the table is generated by the early implementation of our *Action State Estimator (HMM)*. In this example, in time step 1 an action is estimated as "not ready" as the preconditions are not met from the human's perspective. However, the observation in time step 2 states that the robot saw a relevant action, which leads to the inference of the action as "in progress". Given that the preconditions are still not met in time step 3 and 4, the state is estimated as "help needed", where the robot offers its assistance for this action. Following this offer, the emotional state of the human, which was neutral in step 4, has changed to "disapproval" in the observations received in time step 5. This leads to the estimation of the state as "aborted", since the observation vector shows the human is not doing something relevant to the action and has already rejected the robot's help. Finally, the robot reasons that this was not the action human was intending to do. Using the model, we generated 100 different observation sequences, such as the one given in Table 2, each having 100 sequences and used them as ground truth data throughout the tests. Playing these back, we then obtained 94% accuracy in estimating the state with the HMM. Although a synthetic experiment, this shows that our initial implementation of *Action State Estimator* is consistent.

Table 2 An observation sequence and estimated states for an action.

Time Step	**1**	**2**	**3**	**4**	**5**
Sequence	0,0,-1,0	0,0,1,0	0,0,1,0	0,0,1,0	0,-1,0,0
States	not_ready	in_progress	in_progress	help_needed	aborted

ASE estimates the state of each action in the human's mind separately, yet does not take into account their correlations for an overall intention. This is done by our adaptive *human-robot shared planner (HRSP).* The states of actions are all estimated in parallel and this information is then integrated into the *HRSP*, as shown in Fig.1. Our novel approach explicitly incorporates these estimated human action states, which also involve emotional states as reactions, into the human-robot shared planner. The planner estimates a human's belief on the state of their plan and stochastically determines the optimal policy for the robot through accomplishing the human goals in the human's preferred manner. This policy generation is through the understanding of a human's true intention and the need for assistance. Both of these factors are assumed in our experiments to be dynamic and randomly changing. Therefore, the planner must additionally be able to model the irrelevant human intentions that are unknown to the model (targeting the *condition 1*) and the human's potential unwillingness to take assistance from the robot (targeting *condition 2*). For these purposes, we propose using a stochastic planner for the *HRSP*, in particular a Partially Observable Markov Decision Process (POMDP), inspired by Baker and Tenenbaum [7]. This has states as the human's beliefs on a plan through a goal, e.g., "needs help in finding the object", "knows where the object is", etc. The actions of the POMDP model are the human and the robot actions and the rewards are introduced in the form of human emotional reactions to the robot (i.e., approval or disapproval). Finally, as a contribution, the POMDP model takes the estimated states of all actions as observations in order to calculate a belief distribution on the state of the plan a human may be realizing (see Fig.1).

An example implementation of this model is included in Fig.1, which shows a model of a human's desire for an object and the process of acquiring it. As shown in the figure, although the states defined refer to a human's mental states regarding the object, the states that are labeled "doesn't know where and doesn't have the object" and "knows where and doesn't have the object" could also model the human's intention for which that object may be irrelevant. In other words, as long as the robot estimates these two states and observes that the human is realizing some action irrelevant to the targeted object (e.g. gazing at another object), the robot may reason that the human's intention is not relevant to the object and unknown to this specific model (targeting *condition 1*). To better describe the role of the action states in the *HRSP,* in Fig.1 we give some example action states from the *ASE* to the *HRSP* compoment. In the example, the state of action *ac2*, which defines the human taking the object, is given as "in progress" whereas *ac1,* which defines the human looking for the object, is given as "done". Given these observations, the belief state of the human in the example model in Fig.1 will more likely be estimated as the state "knows where and doesn't have the object". This means the human knows where the object, but is still trying to acquire it. Afterwards, if the state for *ac2* is estimated as "done", the belief state will make a transition to the state of "has the object" in the model.

Some example policies that are expected from the *HRSP* are: 1) If the human intention is unknown or known but the human rejected the offer for help, the robot stays and expects the human to act; 2) If the human intention is known and the human needs help, the robot acts. Positive rewards for the POMDP model are acquired through receiving positive human reactions (i.e., approvals) to the robot's correct offers for help and when the human reaches his goal. Additionally, we introduce negative rewards for the negative reactions to encourage the planning to be less intrusive, i.e., the robot will not offer help unless it is deemed part of the optimal policy. Therefore, the POMDP solver takes human reactions as rewards and allows the robot to learn on which action the human usually approved the robot's assistance so as to in the future act to assist the human in a personalized manner.

III. Conclusion

We propose an architecture, given in Fig.1, that models the human-robot shared plan as estimated from the perspective of the human, which incorporates the human's variability in their intentional (*condition 1*) and emotional states (*condition 2*). Currently, we make the assumption that the *Sensing* component is available to the system (and provides pre-defined predicates), so as to particularly focus on our proposed *HRSP* component through modeling a human's intention for taking an object and its integration to our *ASE* component. Our first goal is to examine our initial claim, that the estimation of the human's belief over their action states should be treated as observations in modeling human plans instead of using only the human's visual, positional and postural perspective-taking in order to handle the specified *two conditions*. For these purposes, we will develop the *HRSP* model only using human's positional and visual perspective-taking as observations (connecting the *World States* block to the *HRSP* directly, omitting the *ASE*), and only using the action states as observations (the complete architecture in Fig.1) and compare their accuracy in estimating the *two conditions*. In future work, the proposed model will be tested on real human data.

IV. Acknowledgments

This work was supported by the German Federal Ministry of Education and Research under the ISCO project grant (16KIS0580).

References

[1] C. D. Kidd and C. Breazeal, "Robots at home: Understanding long-term human-robot interaction," in *Proc. IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS*, 2008, pp. 3230–3235.

[2] I. Leite, C. Martinho, and A. Paiva, "Social Robots for Long-Term Interaction: A Survey," *International Journal of Social Robotics*, vol. 5, no. 2, pp. 291–308, 2013.

[3] B. Scassellati, "Theory of mind for a humanoid robot," *Autonomous Robots*, vol. 12, no. 1, pp. 13–24, 2002.

[4] M. Berlin, J. Gray, A. L. Thomaz, and C. Breazeal, "Perspective Taking: An Organizing Principle for Learning in Human-Robot Interaction," *Proceedings Of The National Conference On Artificial Intelligence*, vol. 21, no. 2, pp. 1444–1450, 2006.

[5] L. M. Hiatt, A. M. Harrison, and J. G. Trafton, "Accommodating human variability in human-robot teams through theory of mind," in *IJCAI International Joint Conference on Artificial Intelligence*, 2011, pp. 2066–2071.

[6] G. Trafton, L. Hiatt, A. Harrison, F. Tamborello, S. Khemlani, and A. Schultz, "ACT-R/E: An Embodied Cognitive Architecture for Human-Robot Interaction," *Journal of Human-Robot Interaction*, vol. 2, no. 1, pp. 30–55, 2013.

[7] C. L. Baker and J. B. Tenenbaum, "Modeling Human Plan Recognition using Bayesian Theory of Mind," in *Plan, Activity, and Intent Recognition: Theory and Practice*, 2014, pp. 177–204.

[8] S. Holtzen, Y. Zhao, T. Gao, and S. Zhu, "Inferring Human Intent from Video by Sampling Hierarchical Plans," in *International Conference on Intelligent Robots and Systems (IROS'16)*, 2016, no. 1, pp. 1489–1496.

[9] S. Devin and R. Alami, "An Implemented Theory of Mind to Improve Human-Robot Shared Plans Execution," in *11th ACM/IEEE International Conference on Human-Robot Interaction (HRI'16)*, 2016, pp. 319–326.

[10] O. C. Görür and A. M. Erkmen, "Elastic networks in reshaping human intentions by proactive social robot moves," in *The 23rd IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN'14)*, 2014, pp. 1012–1017.

[11] O. C. Görür and A. M. Erkmen, "Intention and Body-Mood Engineering via Proactive Robot Moves in HRI," in *Handbook of Research on Synthesizing Human Emotion in Intelligent Systems and Robotics*, J. Vallverdú, Ed. IGI Global, 2015, pp. 256–284.

[12] S. Penkov, A. Bordallo, and S. Ramamoorthy, "Inverse Eye Tracking for Intention Inference and Symbol Grounding in Human-Robot Collaboration," in *Robotics: Science and Systems (RSS), Workshop on Planning for Human-Robot Interaction*, 2016, pp. 5–7.