# Social Cobots: Anticipatory Decision-Making for Collaborative Robots Incorporating Unexpected Human Behaviors

O. Can Görür
DAI-Labor, Technische Universität Berlin
Berlin, Germany
orhan-can.goeruer@dai-labor.de

Benjamin Rosman
CSIR, and University of the Witwatersrand
Johannesburg, South Africa
brosman@csir.co.za

Fikret Sivrikaya
GT-ARC gemeinnützige GmbH
Berlin, Germany
fikret.sivrikaya@gt-arc.com

Sahin Albayrak
DAI-Labor, Technische Universität Berlin
Berlin, Germany
sahin.albayrak@dai-labor.de

## ABSTRACT

We propose an architecture as a robot's decision-making mechanism to anticipate a human's state of mind, and so plan accordingly during a human-robot collaboration task. At the core of the architecture lies a novel stochastic decision-making mechanism that implements a partially observable Markov decision process anticipating a human's state of mind in two-stages. In the first stage it anticipates the human's task related availability, intent (motivation), and capability during the collaboration. In the second, it further reasons about these states to anticipate the human's true need for help. Our contribution lies in the ability of our model to handle these unexpected conditions: 1) when the human's intention is estimated to be irrelevant to the assigned task and may be unknown to the robot, e.g., motivation is lost, another assignment is received, onset of tiredness, and 2) when the human's intention is relevant but the human doesn't want the robot's assistance in the given context, e.g., because of the human's changing emotional states or the human's task-relevant distrust for the robot. Our results show that integrating this model into a robot's decision-making process increases the efficiency and naturalness of the collaboration.

## KEYWORDS

Human-Robot Collaboration; Anticipatory Decision-Making; Intent Inference

**ACM Reference Format:**
O. Can Görür, Benjamin Rosman, Fikret Sivrikaya, and Sahin Albayrak. 2018. Social Cobots: Anticipatory Decision-Making for Collaborative Robots Incorporating Unexpected Human Behaviors. In *HRI '18: 2018 ACM/IEEE International Conference on Human-Robot Interaction, March 5–8, 2018, Chicago, IL, USA.* ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3171221.3171256


*HRI '18, March 5–8, 2018, Chicago, IL, USA*

ACM ISBN 978-1-4503-4953-6/18/03…$15.00
https://doi.org/10.1145/3171221.3171256

## 1 INTRODUCTION

Robots that are capable of physically interacting with humans are in the process of revolutionizing industry, with these robots envisioned as active companions that are capable of interpreting human behaviors and adapting to them [16]. For an efficient human-robot collaboration (HRC), robots should not only consider the environment and monitor the human partner's actions, but also process those actions to anticipate the human's knowledge, state of mind and contribution to the collaborative task [10, 11, 17]. The use of such anticipated knowledge in a robot's decision-making mechanism is dependent on the adaptation of these collaborative robots (cobots) to various types of humans, their dynamic behaviors and needs. Realizing such adaptivity would result in cobots working more efficiently and seamlessly with their human partners, increasing overall productivity [8, 9, 13, 19, 20].

There is a large body of research focusing on adopting the ability of humans to interpret and predict others' mental states into robotics in order to anticipate human intentions and plans from observed actions, e.g., [1, 12, 24]. A handful of recent studies have extended them to explicitly use these human states in adaptive plan generation and execution, e.g., [6, 13]. These studies are built on the following requirements of collaboration: the agents have a common intention, i.e., commitment to a shared goal along with a common belief about the goal state [5], as well as mutual awareness and mutual support, i.e., the willingness to accept support [10]. In general, the aim of the field has so far been to prove that anticipatory planning increases efficiency in HRC scenarios, but this assumes that these basic requirements hold during the collaboration, and as a result, implicitly makes two common *assumptions*:

*i)* All of the actions a human executes are relevant to a goal or an intention that is known to the cobot [1, 6, 7, 12, 13, 19],
*ii)* Humans always accept the robot's assistance when offered [6, 13, 22].

In reality, a human's dynamic desires and emotional states could result in stochastic intentions, behaviors and expectations over the course of repeated interactions. Therefore, a robot making these assumptions might misinterpret human actions, which may result in unreasonable and intrusive robot behaviors, limiting its reliability and applicability in real-life scenarios. As an example, a robot may infer that a human needs an object if it detects the human's gaze on it. However, the human's mental state behind this gaze could


**Figure 1: Simulation of an HRC at a conveyor belt for the task of product inspection and storing.**

be any other task-irrelevant intention (in contrast to *assumption i*). Moreover, even if the human wants the object, she may not want the robot to pick it up, for any of several reasons such as a distrust in the robot and the belief that it could damage the object or a desire to remain autonomous (in contrast to *assumption ii*). We refer to such human behaviors as "unexpected" in the context of the given *assumptions*. The premise of this study is, for the first time to the best of our knowledge, to devise an architecture that aims to remove these assumptions and model the HRC over the course of repeated interactions to be able to reason about such "unexpected cases" that could affect the performance of the collaboration.

For this purpose, we propose an autonomous robotic architecture, at the core of which lies a novel anticipatory HRC planner: a partially observable Markov decision process (POMDP) with two stages of human state of mind anticipation. In the first stage, the planner incorporates the variability of the human's state of mind during the collaboration, which in our case is the human's task related *availability, intent (motivation)* and *capability*. Then, in the second stage, through these first anticipated states, it tries to estimate if the human needs help and whether the robot should intervene. While doing so, the planner's goal is to increase the efficiency and the reliability of the collaboration, ensuring the safety and the autonomy of the human partner. Our contribution lies in the ability of our model to effectively handle such unexpected conditions:

*(1)* When the human's intention is irrelevant to the assigned task and may be unknown to the robot, e.g. motivation lost, another assignment is received, becoming tired.
*(2)* When the human's intention is relevant to the task, but the human does not want the robot's assistance, e.g. because of the human's changing emotional states or the human's task relevant distrust for the robot.

Our goal is to demonstrate that anticipating and taking into account such human variability increases the overall efficiency (increased success rate over a shorter time) and the naturalness (less warnings received from the human, hence less intrusive robot behaviors) of an HRC. In order to exemplify and validate our approach, we focus on a smart factory environment in this study. We consider an HRC scenario at a conveyor belt for the task of inspecting and storing various 3D-printed products, as illustrated in Figure 1). We model different types of human workers and let the models interact autonomously with two different robot planners, one of which incorporates the *conditions (1)* and *(2)* through this two-stage anticipation (Section 3). Finally, we present our results and analysis of the effects of our robot model on the overall productivity and naturalness of collaboration (Section 4).

## 2 RELATED WORK

Studies in HRC have generated significant results on low-level (functional) planning for robots having safe and productive physical interactions with humans. However, as it has been recently stated by Lasota et al. [17], alongside safe control and motion planning, significant importance should still be given to develop more generalized predictive and anticipatory planning solutions for increased safety and efficiency of the collaboration. Previous work on reasoning over human mental states has mainly focused on visual perspective taking and belief management in understanding the world from the interacting person's point of view [2, 24]. Utilization of this information has been shown to improve human-robot teamwork significantly, leading to more effective and natural collaboration [9, 25]. There is a large body of research focusing on adopting the concept of human Theory of Mind (ToM) into robotics [24]. More recent approaches have focused on reverse engineering ToM, where they show that a human's intentions and plans can be inferred by observing the human's actions [1, 12]. However, these 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 such human mental states and their explicit use in adaptive high-level (shared task-level) plan generation and execution in HRC [6]. Several approaches have been proposed targeting this gap, where a robot estimates its human partner's belief over the state of a shared table cleaning task [6], and a robot adapts to a human's knowledge level in a cooking task [19] during the execution of the tasks to adapt to the changing human-robot work division. Although these approaches inspire our study, they assume that the belief estimation is a fully observable and deterministic process. There have been several studies applying stochastic planning approaches to robot decision-making during social HRI. For example, a robot car uses a POMDP to decipher the intention of the human driver and adapts to it in a driving task [3, 4], a robot is guided by a user while it uses a MOMDP model to anticipate unstable guiding and take some control in a shared autonomy task [21], and a robot with an anticipatory motion planner serves a human based on her desires anticipated from her gaze [13].

Unexpected human actions are partly considered in [15] and [9]; however, the actions are still assumed to be toward fulfilling a task, possibly in a way that differs from the expected plan. Therefore, all these reviewed studies still assume the human is committed to a given goal and focus the collaboration more on the task execution level rather than the higher-level problems like the *two conditions* mentioned in Section 1. In the long-term, humans' diversity and various mental states emerge as significant factors that impact human actions [18]. Particularly in a repeated HRC over some tedious tasks, it is more likely that the human performs behaviors over time that are not even related to the task itself but implicitly affects her performance, e.g., due to fatigue [14]. The robots should

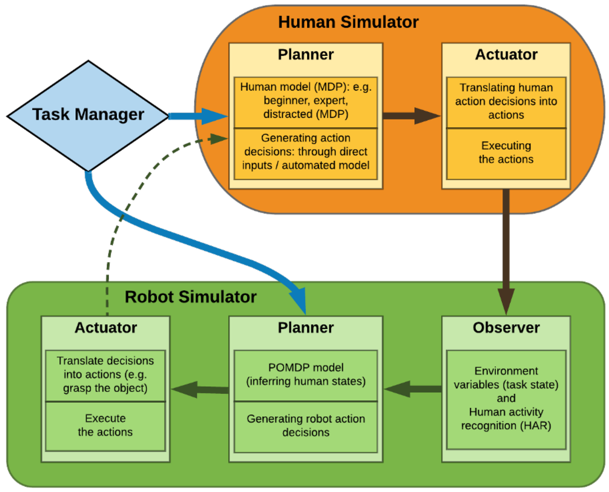


**Figure 2: Overall framework of our autonomous system**

be aware of and adapt to such unexpected behaviors of humans, which to the best of our knowledge, represents a largely unexplored area of research. For this purpose, we propose our POMDP mechanism, inspired by the aforementioned intent detection POMDPs, as a complementary solution to the existing motion planners and as an alternative to the existing high-level decision-makers, that additionally incorporates unexpected human behaviors.

# 3 METHODOLOGY

## 3.1 General Framework

In our general HRC framework, we let both the robot and the human intervene and assist each other when needed, where the robot's ultimate goal is to estimate correctly and non-intrusively when to intervene and assist the human. To facilitate this, we present an architecture consisting of three main building blocks, as depicted in Figure 2: Task Scheduler, Human Simulator, and Robot Simulator. Each HRC scenario starts by receiving a task of inspecting and storing 3D-printed products (see Figure 1) assigned initially either to the human or to the robot. The human takes an action decision, as elaborated in Section 3.3, which is then actuated within the simulation environment in our setup. Although robot perception is not the focus of our contribution, we equip the robot with a human sensing capability, through recognizing distinct human gestures, to achieve a more realistic implementation (Section 4.1). These observations are then fed to our robot planner model to estimate the human's belief and take action decisions accordingly, as explained next.

## 3.2 Robot Models

For the purpose of clearly showing the advantages of using our anticipatory planning approach, we design two robot models where one is intended to handle the *two conditions* stated in Section 1 while the other cannot. We name these the *proactive* and *reactive* models, respectively. Our goal is to examine the additional effects of the stochastic interpretation of a human's need for assistance (*condition (ii)*) and the anticipation of a human's changing *availability, motivation* and *capability* (*condition (i)*) on the overall efficiency and naturalness of the HRC.

**The reactive robot model** is introduced as a base model for the comparative evaluation of our proactive robot. It is a Markov Decision Process (MDP), as shown in Figure 3, with a tuple $\{S, A, T, R, \gamma\}$: $S$ comprises the state of the human collaborator's need for help from the robot's perspective, the global success and failure states that define the result of the task (terminal states), the states of a new task assigned to the human or to the robot (initial states), and a state when the robot receives a warning from the human for any reason; $A$ is the robot actions listed in Figure 3; $T$ is the state transition probabilities; $R$ is the immediate reward the robot receives; $\gamma$ is the discount factor for delayed rewards. Positive rewards are acquired when the global success state is reached (the task has been accomplished by some agent). Negative rewards are assigned when the global fail state is reached, or when warnings are received from the human 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. We solve the MDP model for a robot policy, $\pi$, that is optimal with respect to the robot's expected total reward.

In the reactive model the states are directly observable to the robot through the list of observations listed in Figure 3. How these observations are obtained is given in Section 4. Toward our goal of examining the effect of handling *condition (ii)*, the state of *human needs help* is fully observable in the reactive model, reflecting the assumption mentioned in Section 1, namely that humans accept the robot's help right away. In other words, the MDP treats the human's need for help as a directly observable (deterministic) state. In general, the reactive behavior of the robot is expressed through the robot directly taking over the task no matter what the human's actual internal state is (i.e., ignoring the *two conditions*). The robot assumes the human needs help deterministically when (i) a certain time duration has passed with no success achieved (i.e., leaving enough time for the robot to realize the task before its time limit is reached in a continuous process), (ii) the human is not detected around the work place, or (iii) the human takes a task-related action but no success is detected (e.g., the human failed to grasp and lift in our case).

**The proactive robot model** is a POMDP inspired by available human-intent based POMDP planners, e.g. [4]. The model is a tuple $\{S, A, T, R, \Omega, O, \gamma\}$. The five elements, $S, A, T, R, \gamma$, have the same interpretations as in the reactive model, while $\Omega$ is the set of observations as listed in Figure 3 and $O$ represents the conditional observation probabilities. We also solve the POMDP model for an optimal robot policy, $\pi$. Both the proactive and reactive models share the same immediate reward assignments and receive the same observations from the world. As mentioned, we keep their differences to the level of handling the *two unexpected conditions*. These differences are twofold: (1) although they share the same rules for the detection of a human's need for help (as visualized in Figure 3), the proactive model is not deterministically bound to these rules through its partial observability (as opposed to *assumption (ii)* in Section 1); (2) the proactive model has more belief states than the reactive one, as visualized in Figure 3, to be able to anticipate human's changing availability, motivation and capability

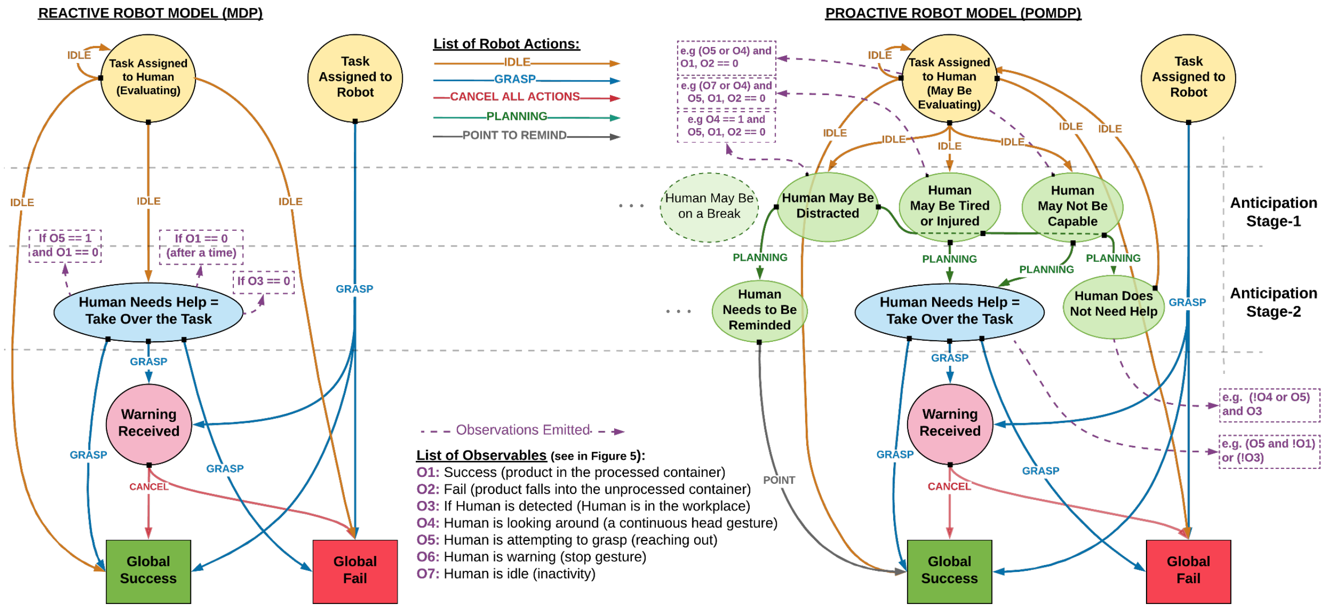


**Figure 3: Reactive and proactive state-action connection models. The observables that are input to both of the systems are listed. The states added to the proactive model which incorporate the further reasoning (anticipation) layer of the robot are shown in light green.**

contrary to *assumption i)*. For the purpose, in our implementation we model the human states of being distracted, tired and not capable of fulfilling a task. In the reactive model, such intermediate states of the human that give unexpected behaviors are encapsulated under *human needs help* whereas those are all handled separately in the proactive case.

In the proactive model, we distribute these states to the two stages of anticipation. The *anticipation stage-1*, as shown in Figure 3, is an additional stage in the proactive model consisting of hidden human states of *availability, motivation* and *capability*. We believe that it is necessary for the robot to be able to account for the so-called unexpected human behaviors given in *condition (1)* in Section 1. Having such a hierarchical anticipation phase also helps the robot to reason about if the human does not need help, see in the *anticipation stage-2* in Figure 3. We put this condition as another state which adds uncertainty to the robot's previous estimation in the anticipation stage-1. This gives the robot the ability to reason that it was actually wrong when the observation gathered from the human changes accordingly. This prevents the direct conclusion of taking over the human task when it should not, or when changes occur during the interaction. As a result, it is expected to decrease the number of warnings the robot receives from the human. For clarity and brevity of presentation, only the most prominent state transitions of the proactive model are visualized in Figure 3. For example, the transitions from anticipation stage-1 to the *Global Success* or *Global Fail* with the *Planning* action is not shown (i.e., the human has already ended the task or it failed). Similarly, we provide only some examples of the observations emitted from the proactive model states in the figure. From these examples, we point out that it is not trivial to estimate the states in the anticipation stage-1 due to their similar observations emitted.

Such a hierarchical approach also provides some additional capabilities to the robot's reasoning that are expected to contribute to the fluency of the collaboration. For example, as depicted in Figure 3, the robot has some additional action decisions such as *planning* and *pointing to remind* as a result of the additional states in the proactive model. In the *planning* action, which is a necessary step for any other action to be taken, the robot starts to plan its motion (e.g., for grasping: find the grasping points and plan for moving the grippers) right after any state in the anticipation stage-1 is estimated. In the cases where the human really needs help, this behavior is expected to save the robot a significant amount of time to execute the action, in our case grasping. In addition, after estimating that the human may be distracted, e.g., detected *looking around* action for a time, the robot may take the decision of pointing toward the object to draw the human's attention to it rather than directly taking it. The human states we include in our proactive model are intended to reflect on the general possible human states in a work environment. As shown in Figure 3, such human states and so robot capabilities can be extended by adding more states to the anticipation stage-1 and some relevant robot decisions to the stage-2.

Some general remarks that apply to the robot models are provided below.

*i)* The models are designed generically to comply with various HRC scenarios. For better clarification we use *grasp* as the action to achieve our specific task but in general any action that a task requires can replace it, making the model adaptable to different use-cases.

*ii)* Both models start with the task assignment states. If the task is assigned to the human, the robot first remains idle and observes the human (see idle action in Figure 3).

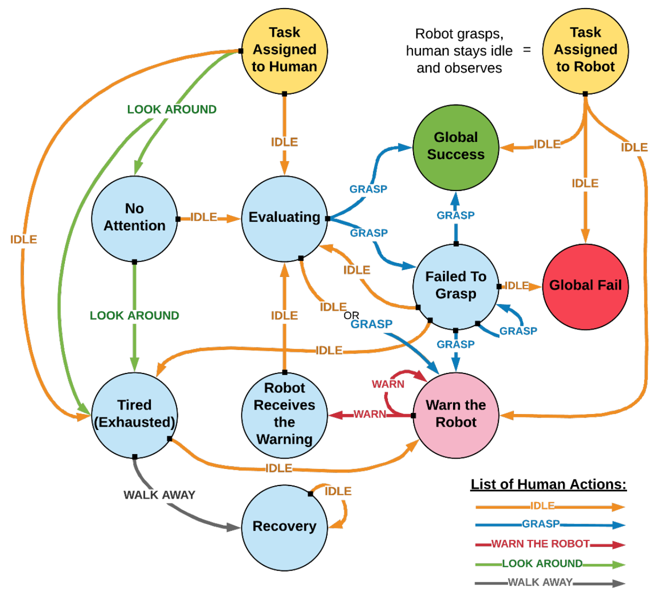


**Figure 4: Human model: human state-action transitions**

*iii)* Since our focus remains on the human states of mind that have a direct impact on the progress of the task, we do not consider the internal states of the robot (e.g., its battery level). We assume that the robot action of grasping always leads to global success, global failure or warnings from the human.

*iv)* The exact reason for a warning is hidden to the robot but it may be due to the human's task related distrust to the robot, the desire to remain autonomous, an incorrect estimation of the robot about the human's need for help, the robot realizing the task incorrectly, etc.

*v)* After a warning is received, the robot cancels its action for the safety and autonomy of the human.

### 3.3 Human Models

As a proof of concept, we devise experiments where we control the degree to which the human displays these unexpected random behaviors. Evaluating these behaviors with the corresponding robot responses would be difficult with real human subjects in experiments (in an uncontrolled environment) as they depend on the personality of the human and are likely to be observed over the course of longer interactions in the workplace [14]. In simulation, through modeled humans, we scale the experiments to emulate many different combinations of such behaviors to rigorously test our robot model's capability to estimate, avoid and respond to such cases (e.g., a human stubbornly rejecting the robot's help, getting tired fast, being distracted easily). We are agnostic to the exact implementation of the human models (i.e., the states, actions, transitions) while our goal remains creating use-cases where a human worker follows the aforementioned *two conditions* and occasionally performs unexpected behaviors.

For our purposes, the representative proof of concept example model is implemented using an MDP as shown in Figure 4. We note again that only the transitions with non-negligible probabilities are shown in the figure for the clarity of presentation. The MDP is a tuple $\{S, A, T, R, \gamma\}$ where $S$ is the human states of mind, $A$ is the human actions, and $T$ is the state transition probabilities and $R$ is the immediate rewards received based on the result of the task and the type of the human to encourage that type of behavior, e.g. a distracted person receives positive rewards in the *global success* and *no attention* states. Through changing $T$ and $R$ and solving for various policies, $\pi$, the model induces random but goal-oriented dynamics for the human collaborator and is used to automate the testing of our robot reasoning under various hard-to-predict conditions.

While this is not necessarily an accurate human model, the states (e.g., no attention, tired) are inspired by the available studies analyzing human workers operating on repeated tedious tasks, e.g. [14]. During the operation the human first selects an action. Then, a state transition occurs randomly (if the task does not succeed or fail), reflecting normal and unexpected human states of mind. In normal behavior the human first evaluates the task, grasps the product, and inspects and places it into the adjacent container (see Figure 5a). The rest of the state transitions reflect the unexpected behaviors through the given examples in Section 1. In the cases when the robot incorrectly estimates the human's need for help (detailed in Section 4.1), the human may transit to the *warn the robot* state, which may resemble the trust in the robot of the human in a given task. The warning state is followed by a human action of gesturing to stop the robot action (see in Figure 5d).

Keeping the same scheme given in Figure 4, different state transition probabilities and rewards lead to different human types. We use four of them, referred to as *i*) normal, *ii*) stubborn, *iii*) distracted and *iv)* tired. Referring to the states shown in Figure 4, the type-*i* model is most likely to succeed after a short period of evaluation of the task. The type-*ii* is more likely to fail the task, in our case the human fails to grasp, e.g., it is too heavy to lift for that person. In this case, the robot should intervene to assist before a global fail occurs. However, the human may want to try once again, iterating between evaluating and failing phases. The type-*iii* human mostly looks around resembling lost attention, or stalls more in the evaluating phase. Finally, type-*iv* is observed also through stalling in the evaluating phase and most likely ending up being exhausted, which then forces the human to leave the workspace for a recovery. The similarities in different types are expected to result in making anticipation difficult for the robot since, for example, staying idle or looking around could indicate the human states of *evaluating*, *no attention* or *exhausted.*

## 4 EVALUATION

### 4.1 Experiments

We implement the proposed architecture in the Robot Operating System (ROS) and simulate our smart factory scenario with a conveyor belt by using the MORSE environment, utilizing the available human and PR2 robot models. All scenarios consist of several sequential task assignments for the purpose of simulating long-term collaboration. Each task of product inspection and storing starts with an initial assignment to either the robot or the human based on the product's weight and fragility. We consider only the cases where the task is assigned to the human, in order to keep our focus on anticipating the human's states and need for assistance. If the

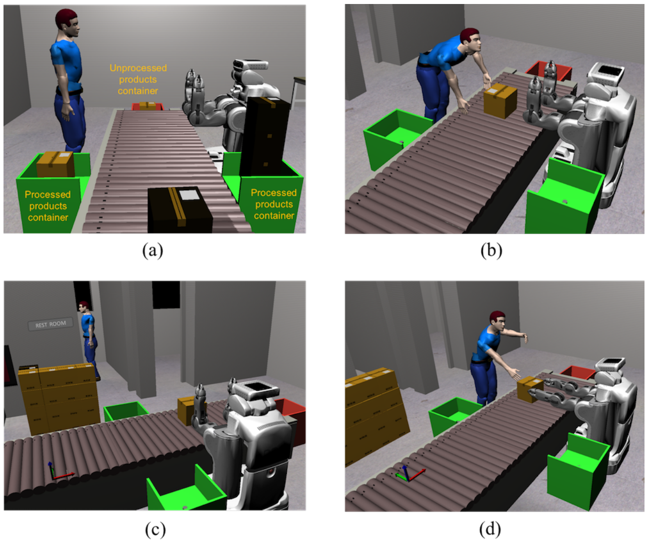


**Figure 5: Our HRC scenario. (a) Idle human and robot, with containers shown; (b) Human grasps the product; (c) Human goes to the rest room (walks away); (d) Robot takes over the task and human gestures to stop the robot (warns the robot).**

product is eventually inspected either by the human or by the robot, it is placed into one of the inspected-product containers (green containers in Figure 5) resulting in the *Global Success* state of all Markov models. Since the collaboration is a continuous process, we set a maximum allowed processing time for each product inspection, $t_{max}$. The conveyor belt waits when a package is between the human-robot team for $t_{max}$ or for the package to be successfully processed, before starting a new task. If the product is not processed before $t_{max}$, it falls into the uninspected-product container (red container in Figure 5a) interpreted as a *Global Fail*.

At this phase of the study both the reactive and proactive robot models are hand-coded addressing the points discussed. In order not to bias the robot models, they are trained agnostic to the state transitions inherent in the human models, i.e., no data generated by the human models is used in training the robot models. During the experiments, we measure the state estimation accuracy using the human models as the ground truth. Our goal at this stage is to match the ground truth and do the tests against rather unexpectedly acting human models unknown to the robots.

The observations the robot receives are the 3D human body joints that are always available directly from the simulated human model and the proximity sensors placed inside the containers to monitor the task status as succeeded, failed or ongoing (listed in Figure 3). A state-of-the-art human activity recognition (HAR) module, inspired by existing studies, e.g., [23], has been implemented to recognize the constrained and distinct simulated human gestures from the body joints available. These are: the human is looking around (from the body pose), attempting to grasp (see in Figure 5b), warning the robot (a special stop gesture shown in Figure 5d), idle (inactive) and walking away (see Figure 5c). The robot then uses these observations in both of the reactive and proactive models to estimate the next state (the human belief). We note that as our focus is in showing the effect of handling such unexpected human behaviors on the collaboration performance, we use the same observation conditions for both models and our insights from the performance comparisons are agnostic to the HAR component that is used.

In each scenario consisting of several tasks, a human is first created as the normal type (referring to the human types introduced in Section 3.3), while the other types (i.e., the policies of the stubborn, distracted and tired MDP models) are then executed on the human during the scenario randomly, but becoming more likely as more tasks are assigned. By doing so, in each scenario the robot models interact with a human with changing levels of stubbornness, tiredness, and distraction. This induces more occurrences of the aforementioned unexpected human behaviors over the course of the collaboration as the number of task repetitions increase. Additionally, the human may warn the robot when the robot estimates the human's need for help incorrectly. It would be correct in the case when the human is tired. In a distracted case the human approves the help proportional to the time of the distraction. Moreover, when the human tries and fails to lift the object, it is more likely to approve the robot's taking over the task unless the human is still trying to grasp. These cases are hard coded and the next state observations are input to the human MDP model executor by the system. The decision of a next state is random if there is no update of the task status or an action from the robot. This allows us to observe various scenarios of the humans randomly transitioning between the states.

To evaluate how our proactive model covering such human *conditions* contributes to the naturalness and the efficiency of the collaboration, we gather the objective measures below during our experiments:

- **Human state distributions:** To show on average which states the human models have selected (based on the state transitions given in Figure 4) and how the state transitions change over time.
- **Estimation accuracy:** To compare how accurately both models estimate the human's true need for help, and to show how our proactive robot model performs in estimating human belief states in general (taking the interacted human models as the ground-truth).
- **Success rate:** The comparison of the success rates of a task with a human alone, a human and a reactive robot collaborating, and a human and a proactive robot collaborating.
- **Rewards gathered:** To show the change of overall rewards gathered over time, and to compare how many warnings the robot receives from the human (also in line with the wrong estimations of the need for help) in reactive and proactive mode. This hints at the naturalness of the collaboration.
- **Time to finish the task:** To compare the average time taken to complete a task with the reactive and the proactive models.

We use an updated version of the DESPOT online POMDP planner [26] to solve for both the POMDP and the MDP models, and execute them in real-time while interfacing the ROS environment through the *Planner* components as demonstrated in Figure 2.

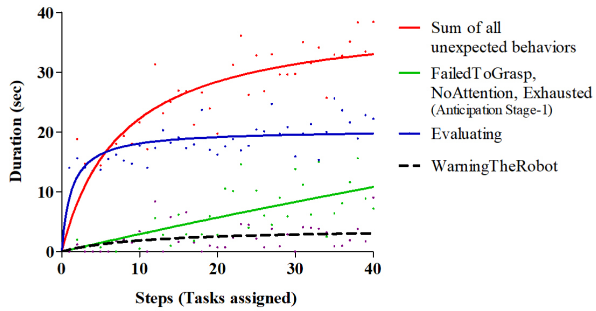


**Figure 6: 2D scatter data of the states of the human models over the course of the tasks assigned (referring to Figure 4). A nonlinear curve fit over the data shows the overall trend of the average time spent in these states.**

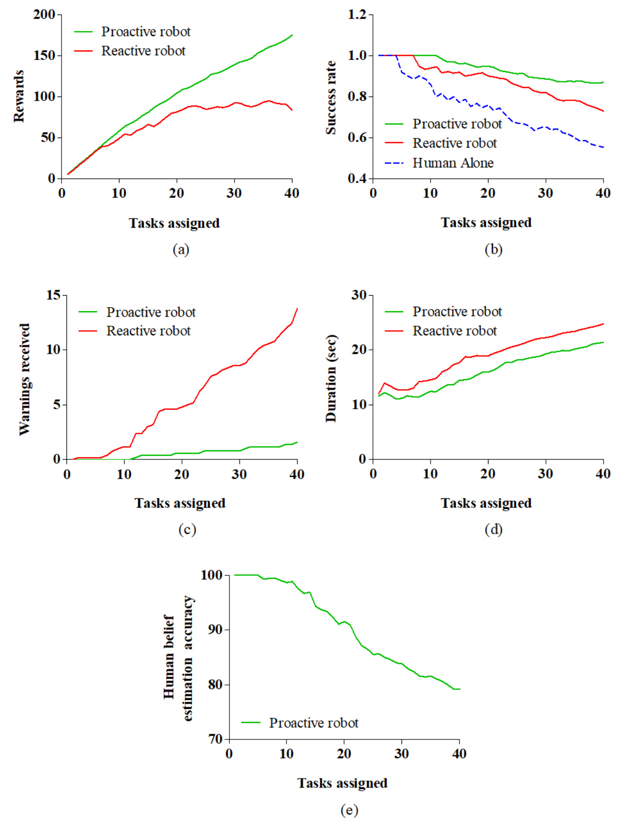


**Figure 7: Comparison of the proactive and reactive robot models over the task assignments: cumulative (a) rewards acquired; (b) moving average of the success rates; (c) number of warnings received; (d) moving average of the task durations in seconds; (e) moving average of human belief estimation accuracy of the proactive model. Each result obtained from proactive and reactive robots are averaged over 5 trials.**

## 4.2 Results

We run 5 different scenarios for each of the reactive and proactive robot models. Each scenario consists of 40 sequential task assignments. As mentioned, the humans exhibit the *two unexpected conditions* listed in Section 1 randomly but with increasing likelihood over repeated task assignments. This is also shown in Figure 6, which plots the average time a human spent in the given states over the simulation steps (i.e., task assignments). In particular, concerning *condition (1)*, as more tasks are assigned the humans stall more in the thinking phase (depicted as *evaluating*), are more distracted (depicted as *no attention*), are more likely to fail to grasp, become tired, or lose motivation. Such states are all summed and depicted as the *sum of all unexpected behaviors* in Figure 6. Concerning *condition (2)*, as more unexpected behaviors are performed the robot is increasingly convinced of the human's need for help and takes over the tasks. This leads to more warnings from the human in the case of wrong estimates (depicted as *warning the robot* in Figure 6).

Table 1 shows the results obtained from the experiments with the reactive and proactive robot models whereas Figure 7 demonstrates the change in the measured values over the simulation time. As is demonstrated in the figure, in the first 10 tasks there is no significant difference between the two robot models. The tasks are mostly succeeded by the humans conveying normal behaviors regardless of the robot models (see Figure 7b); therefore, the robots in these tasks constantly receive the maximum rewards (Figure 7a) and the tasks are completed efficiently (Figure 7d). However, as more tasks are assigned, the average success rate falls and the differences in the performance of the two models are enhanced. This is directly in line with the increased likelihood of observing the unexpected human behaviors over the simulation time. Before beginning the comparison of the two models, we show in Figure 7b that the success rate of such a human model alone is worse in the long-term than a collaboration with any of the robot models. Thereby we underscore the importance of such cobots collaborating with humans in tedious tasks.

The negative rewards are acquired from the task failures and the warnings received from the human, the former having more impact than the latter. The positive rewards are received only when the task succeeds. Therefore, the change in the accumulated rewards is directly in line with the changes in the success rate and the number of warnings received (Figure 7b and 7c, respectively). After the 30$^{\text{th}}$ task in Figure 7b, the average success rate keeps decreasing in the reactive case whereas it stabilizes in the proactive model to a rate of around 87% (see Table 1). Additionally, in the reactive case the robot continues to receive warnings during the collaboration while in the proactive model this amount is kept low. These two together contribute to a gradual decline in the increase of the accumulated reward in the reactive case (see Figure7a). On the other hand, as shown in the same graph, the accumulated reward for the proactive case is affected much less by the changing human conditions.

The proactive model copes well with the unexpected human behaviors mainly due to the model's significantly better performance in estimating the human's true need for help, as indicated by the anticipation stage-2 estimation accuracies in Table1. Recall that the robot's anticipation of the human beliefs is divided into two stages, where stage-1 involves reasoning about conditions such

**Table 1: Proactive Robot vs. Reactive Robot Final Results**

| Type | Reactive | Proactive |
|---|---|---|
| **Rewards (Avg.)** | 2.10 | 4.37 |
| **Number of Warnings per Scenario (Avg.)** | 13.8 | 1.6 |
| **Anticipation Stage-1 Estimation Accuracy** | Not Applicable | 79.26% |
| **Anticipation Stage-2 Estimation Accuracy** | 44.69% | 71.37% |
| **Success Rate** | 0.73 | 0.87 |
| **Task Duration** | 24.77*secs* | 21.40*secs* |

as the human being tired, distracted or incapable, and stage-2 is to further estimate the human's need for help possibly as a result of the stage-1 anticipation. In the reactive model the robot takes over the task through the given deterministic conditions without anticipating the stage-1 conditions (see Section 3.2). Therefore, it is often either too late to judge the human's need for help or the robot's interference ends up with the human warning the robot, which results in the poor stage-2 estimation accuracy of the reactive model (44.69%). As mentioned in Section 4.1, if the robot intervenes when the human is still trying to grasp, is only distracted for a short time, or is still evaluating, then the human stochastically rejects the robot's offer for help in line with *condition (2)*. Since these conditions are anticipated by the proactive model in the first phase, the proactive robot stochastically ends up estimating that the human may not yet need help or that the human needs to be reminded when she is distracted, rather than directly taking over the task based on the continuous observations received.

The average estimation accuracy of the proactive robot in the anticipation stage-1 is about 79.26%, which has a direct influence on the stage-2 estimation that is about 71.37%. Although we use a simulation environment and the human has a finite set of states and actions, we observe that it is still nontrivial for the robot to estimate the states given under stage-1. As we demonstrate in Figure 7e, the average human belief estimation accuracy of the proactive model decreases over the task steps, being subjected to the higher frequency of unexpected human behaviors. The model often confuses the states of *human may be distracted* with *human may be tired*, as expected, which is due to these two conditions causing similar human actions, i.e., looking around and staying idle longer. Such wrong estimations lead to task failures since, for example, the robot points out the object to remind rather than offering help when the human is really tired.

We also measure the average task durations in both models (whether it succeeded or failed). As shown in Table 1 the proactive model improves the time a task takes over the reactive case by about 14%, i.e., 3.37 seconds. This can be thought of a significant change especially in an industry with mass production. The improvement is again driven by the hierarchical anticipation in the proactive robots. First of all, a successful estimation of a state in the anticipation stage-1 allows the proactive robot to estimate the human's need for help faster, whereas in the reactive case this needs to wait until observing the human failing to grasp and giving up. Secondly, the robot starts the *planning* in the stage-1 (see Figure 3) while it still observes for stage-2, the estimation of which then leads to a direct execution of the *grasp*. In the reactive case, however, even though the robot's taking over the task is approved by the human, the robot still needs to plan for its grasping motion execution.

Finally, our basic HAR system works with approximately 10% noise, which emerges from the misclassification of human actions, especially for warning the robot and grasping gestures producing similar observations (see Figure 5). To understand the effect of this noise on the human belief estimation accuracy, we also tested providing the actual human actions as direct inputs to the robot, thereby taking HAR out of the loop. In this case the estimation accuracy of the proactive model increased, as expected, to around 85%. For an actual deployment in a noisy real-life environment, where the observations of the hidden human states are not limited to a small set of human actions as in our simulation, belief estimation accuracies are likely to be lower, mainly due to lower accuracies in the HAR system.

In conclusion, we show that the common human conditions of *(1)* intention being irrelevant to the assigned task (e.g., due to motivation loss or tiredness), and *(2)* disapproval of the robot's help (e.g., no trust in the task, not needing/desiring the robot's help) lead to a drastic decrease in the overall success rate of the collaborative tasks. Since these conditions are expected in long-term HRC scenarios, we show in simulation that cobots that can anticipate and handle these conditions yield more efficient (i.e., increased success rate, lower task duration) and more natural (i.e., less intrusive) collaboration when the collaborating human demonstrates these behaviors. We show the feasibility and effectiveness of this concept through our two-stage anticipatory decision-maker, hierarchically and stochastically reasoning about such human behaviors.

## 5 CONCLUSION

This study examines the effects of a robot's anticipation and response to the unexpected human behaviors that could be observed in a long-term collaboration. In this study, as a proof of concept of our novel anticipatory modeling scheme, we use simulated human models in our experiments, which provide a more robust setup to test such various unexpected conditions rather than expecting real human feedback in the long-term. We are aware of the possible biases in the experiments which could be introduced by the simulated humans due to their limited reflections (limited actions) of the hidden human states and the fact that they are hand-coded. We stress that these are not necessarily accurate models; however, the abstracted states in our design are expected to be observed in a real human. Our model has generated promising results, encouraging us to move towards validating these results with real human experiments through a similar model design but training with real human data. This will also provide us with the chance to consider real humans' trust building in the robot over the course of the collaboration, which in our current setup we assume random, thereby creating a more difficult experimental case.


## ACKNOWLEDGMENTS

This work was supported in part by the German Federal Ministry of Education and Research (BMBF) under Grant 01IS16045. We would like to thank Guy Hoffman for his valuable feedback.

# REFERENCES


[1] Chris L. Baker and Joshua B. Tenenbaum. 2014. Modeling Human Plan Recognition using Bayesian Theory of Mind. In *Plan, Activity, and Intent Recognition: Theory and Practice*. 177–204.

[2] Matt Berlin, Jesse Gray, Andrea L. Thomaz, and Cynthia Breazeal. 2006. Perspective Taking: An Organizing Principle for Learning in Human-Robot Interaction. In *Proceedings of the 21st National Conference On Artificial Intelligence (AAAI)*. 1444–1450.

[3] Frank Broz, Illah Nourbakhsh, and Reid Simmons. 2011. Designing POMDP models of socially situated tasks. In *International Symposium on Robot and Human Interactive Communication, 2011 RO-MAN*. 39–46.

[4] Frank Broz, Illah Nourbakhsh, and Reid Simmons. 2013. Planning for Human-Robot Interaction in Socially Situated Tasks: The Impact of Representing Time and Intention. *International Journal of Social Robotics* 5, 2 (2013), 193–214.

[5] Philip R. Cohen and Hector Levesque. 1991. Teamwork. *Nous* 24, 4 (1991), 485–512.

[6] Sandra Devin and Rachid Alami. 2016. An Implemented Theory of Mind to Improve Human-Robot Shared Plans Execution. In *Proceedings of the 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI'16)*. 319–326.

[7] O. Can Görür and Aydan M. Erkmen. 2015. Intention and Body-Mood Engineering via Proactive Robot Moves in HRI. In *Handbook of Research on Synthesizing Human Emotion in Intelligent Systems and Robotics*, Jordi Vallverdú (Ed.). IGI Global, 256–284.

[8] O. Can Görür, Benjamin Rosman, Guy Hoffman, and Şahin Albayrak. 2017. Toward Integrating Theory of Mind into Adaptive Decision-Making of Social Robots to Understand Human Intention. In *Workshop on Intentions in HRI at ACM/IEEE International Conference on Human-Robot Interaction (HRI'17)*.

[9] Laura M. Hiatt, Anthony M. Harrison, and J. Gregory Trafton. 2011. Accommodating human variability in human-robot teams through theory of mind. In *Proceedings of the 22nd International Joint Conference on Artificial Intelligence (IJCAI'11)*. 2066–2071.

[10] Guy Hoffman and Cynthia Breazeal. 2004. Collaboration in Human-Robot Teams. In *AIAA 1st Intelligent Systems Technical Conference*. Chicago, IL, USA.

[11] Guy Hoffman and Cynthia Breazeal. 2007. Effects of Anticipatory Action on Human-robot Teamwork Efficiency, Fluency, and Perception of Team. In *Proceedings of the ACM/IEEE International Conference on Human-robot Interaction (HRI '07)*. 1–8.

[12] Steven Holtzen, Yibiao Zhao, Tao Gao, and Song-chun Zhu. 2016. Inferring Human Intent from Video by Sampling Hierarchical Plans. In *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS'16)*. 1489–1496.

[13] Chien Ming Huang and Bilge Mutlu. 2016. Anticipatory robot control for efficient human-robot collaboration. In *Proceedings of the 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI'16)*. 83–90.

[14] Qiang Ji, Peilin Lan, and Carl Looney. 2006. A probabilistic framework for modeling and real-time monitoring human fatigue. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans* 36 (2006), 862–875.

[15] Hema S. Koppula, Ashesh Jain, and Ashutosh Saxena. 2016. *Anticipatory Planning for Human-Robot Teams*. Springer International Publishing, Cham, 453–470.

[16] KUKA. [n. d.]. *Hello Industrie 4.0 Glossary*. Technical Report. Kuka AG. https://www.kuka.com/-/media/kuka-corporate/documents/press/industry-4-0-glossary.pdf

[17] Przemyslaw A. Lasota, Terrence Fong, and Julie A. Shah. 2017. A Survey of Methods for Safe Human-Robot Interaction. *Foundations and Trends in Robotics* 5, 4 (2017), 261–349.

[18] Iolanda Leite, Carlos Martinho, and Ana Paiva. 2013. Social Robots for Long-Term Interaction: A Survey. *International Journal of Social Robotics* 5, 2 (2013), 291–308.

[19] Grégoire Milliez, Raphaël Lallement, Michelangelo Fiore, and Rachid Alami. 2016. Using human knowledge awareness to adapt collaborative plan generation, explanation and monitoring. In *Proceedings of the 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI'16)*. 43–50.

[20] Stefanos Nikolaidis, Ramya Ramakrishnan, Keren Gu, and Julie Shah. 2015. Efficient Model Learning from Joint-Action Demonstrations for Human-Robot Collaborative Tasks. In *Proceedings of the 10th ACM/IEEE International Conference on Human-Robot Interaction (HRI'15)*. 189–196.

[21] Stefanos Nikolaidis, Yu Xiang Zhu, David Hsu, and Siddhartha Srinivasa. 2017. Human-Robot Mutual Adaptation in Shared Autonomy. In *Proceedings of the 12th ACM/IEEE International Conference on Human-Robot Interaction (HRI '17)*. 294–302.

[22] Svetlin Penkov, Alejandro Bordallo, and Subramanian Ramamoorthy. 2016. 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*.

[23] Alina Roitberg, Alexander Perzylo, Nikhil Somani, Manuel Giuliani, Markus Rickert, and Alois Knoll. 2014. Human activity recognition in the context of industrial human-robot interaction. In *2014 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference, APSIPA 2014*.

[24] Brian Scassellati. 2002. Theory of mind for a humanoid robot. *Autonomous Robots* 12, 1 (2002), 13–24.

[25] Greg Trafton, Laura Hiatt, Anthony Harrison, Frank Tamborello, Sangeet Khemlani, and Alan Schultz. 2013. ACT-R/E: An Embodied Cognitive Architecture for Human-Robot Interaction. *Journal of Human-Robot Interaction* 2, 1 (2013), 30–55.

[26] Nan Ye, Adhiraj Somani, David Hsu, and Wee Sun Lee. 2017. DESPOT: Online POMDP planning with regularization. *Journal of Artificial Intelligence Research* 58 (2017), 231–266.