



HAL
open science

An Epistemic Human-Aware Task Planner which Anticipates Human Beliefs and Decisions

Shashank Shekhar, Anthony Favier, Rachid Alami

► To cite this version:

Shashank Shekhar, Anthony Favier, Rachid Alami. An Epistemic Human-Aware Task Planner which Anticipates Human Beliefs and Decisions. 16th International Conference on Social Robotics (ICSR 2024), Oct 2024, Odense, Denmark. ⟨hal-04942435⟩

HAL Id: hal-04942435

<https://hal.science/hal-04942435v1>

Submitted on 12 Feb 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons CC BY 4.0 - Attribution - International License

An Epistemic Human-Aware Task Planner which Anticipates Human Beliefs and Decisions

Shashank Shekhar¹, Anthony Favier^{1,2}, and Rachid Alami^{1,2}

¹ LAAS-CNRS, Université de Toulouse, CNRS, INSA, UPS, Toulouse, France

² Artificial and Natural Intelligence Toulouse Institute (ANITI)
{sshekhar,afavier,alami}@laas.fr

Abstract. We present a substantial extension of our Human-Aware Task Planning framework, tailored for scenarios with intermittent shared execution experiences and significant belief divergence between humans and robots, particularly due to the uncontrollable nature of humans. Our objective is to build a robot policy that accounts for uncontrollable human behaviors, thus enabling the anticipation of possible advancements achieved by the robot when the experience is not shared, e.g., when humans are briefly absent from the shared environment to complete a sub-task. But, this anticipation is considered from the perspective of humans who have access to an *estimated* robot’s model. To this end, we propose a novel planning framework and build a solver based on AND/OR search, which integrates knowledge reasoning, including situation assessment by perspective taking. Our approach dynamically models and manages the expansion and contraction of potential advances while precisely keeping track of when (and when not) agents share the task execution experience. It systematically *assesses* the situation and ignores worlds that it has reason to think are impossible for humans. Overall, our new solver can estimate the distinct beliefs of the human and the robot along potential courses of action, enabling the synthesis of plans where the robot selects the right moment for *communication*, i.e. telling or replying to an inquiry, or defers ontic actions until the execution experiences can be shared. Preliminary experiments in two domains — one novel and one adapted — demonstrate the framework’s effectiveness.

1 Introduction

The increasing number of robot-assisted applications has led to a heightened focus on human-robot collaboration (HRC) research [2, 30]. Collaborative robots have proven beneficial in real-world scenarios such as construction engineering [23], workshops [8], and nursing care [24].

Studies in psychology and cognitive science, particularly within the domain of joint actions, suggest that humans consider each other’s actions and beliefs, indicating that they model each other’s tasks when planning [28, 19, 29]. However, they often hold inaccurate models of their partners’ beliefs and capabilities. This phenomenon is particularly evident in joint action scenarios, where partners

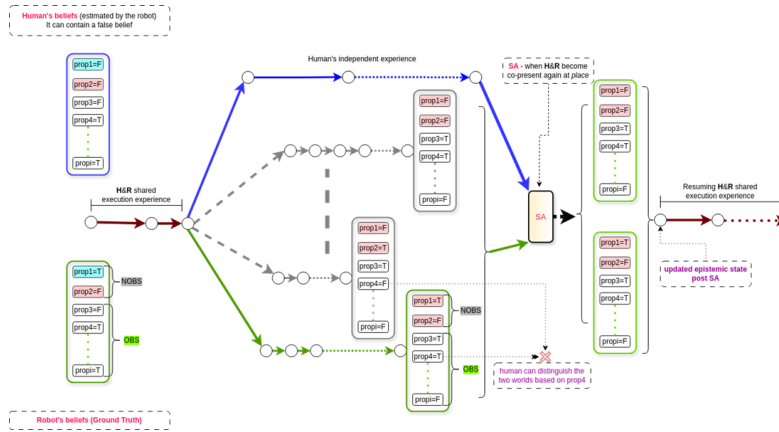


Fig. 1. Our planning framework is endowed with the ability to make the difference between $H\&R$ shared and individual execution experiences in the planned activities. It can anticipate potential belief divergence between $H\&R$ and also estimate the updated beliefs of H when they meet again (situation assessment (SA)) based on a distinction between observable and non-observable facts. This will be used to plan communicative actions or adapt the R 's plan to ensure the shared experience of some actions. In this diagram, we roughly depict what happens when $H\&R$ no longer share the execution experience, H has independent experience (blue), while R progresses towards the goal (green), with anticipated courses of action that the robot can choose along with the green trace but from the H 's perspective. Upon co-presence at place, SA eliminates impossible worlds, e.g., those with state property $prop_4=F$ (since it is observable), aiding H to ignore wrongly estimated worlds.

share a common goal. Individuals tend to expect and estimate their partners' actions based on their own mental models, which may not always be accurate.

When partners are temporarily absent, they rely on these mental models to estimate what the other might have accomplished during the separation. These expectations can be skewed by an incorrect or incomplete understanding of their collaborator's beliefs, intentions, and actions. This cognitive bias shows the complexities of effective collaboration between robots and humans. And, highlights the need for strong frameworks to manage these differences.

Inspired by these insights, we take the first step in this paper towards building robot policies that address the issue of mental models that can be inaccurate. Our proposed strategy integrates tools developed for epistemic planning [3], Dynamic Epistemic Logic (DEL) [4], and human-aware planning [1, 7, 5, 16, 35].

To this end, we propose a novel *epistemic human-aware task planning* framework. It substantially extends our past works and enables the robot to estimate, anticipate, and adapt to scenarios in which uncontrollable human and robot partners have disrupted shared execution experiences. Specifically, it considers the human's perspective and estimation regarding the potential advances achieved by the robot, even when the exact progress is not directly experienced by the humans, who may hold an *incorrect* robot model.

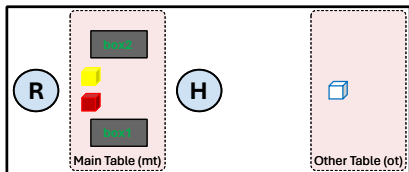


Fig. 2. Three cubes c_r (red), c_y (yellow), and c_w (white) are shown. c_r and c_y are placed on mt (main table), and c_w is on ot (other table). There are two boxes, box_1 and box_2 , placed on mt , which can be either transparent or opaque. The shared task is to organize the cubes in a way that cubes from one table are placed in one box. The choice of which box is flexible as long as each table’s cubes end up in separate boxes.

Our framework employs tools developed in the literature, including those for DEL-based epistemic planning. However, as we will see, it offers more flexibility because, unlike the majority of epistemic planners, we do not need to script in the input all the effects on the beliefs of all agents in action models.

Figure 1 provides a rough illustration of a single plan trace, showing what happens when agents share execution experience and when they do not in the process of achieving the shared task.

In addition, we build an AND/OR search-based offline planner that facilitates Theory of Mind (ToM) by integrating knowledge reasoning and incorporating situation assessment. It dynamically manages the evolution or contraction of estimated possible worlds from the human’s point of view. This helps the planner to prepare itself with a set of worlds that humans would consider possible.

Our *contributions* specific to this paper compared to our past works are:

- We introduce the notion of the human mental model during the deliberation process. We will see how non-trivial changes are required in the existing framework and algorithm to tackle the problem motivated in this paper.
- Non-controllability is not unique here, but we extend [5, 16, 31] to address events of (non-)shared execution experiences and managing beliefs. We present enriched models for co-presence, observability, and situation assessment.

Thanks to the new framework and our planner, it enables the robot to take proactive steps, such as anticipating humans to be *inquiring* about an unknown variable’s value, *communicating* relevant information without being annoying (*e.g.*, not verbalizing a fact already known to them), or deferring *executing* an action until **H&R** reunite, allowing **H** to further narrow down the possibilities.

The paper is structured as follows. A case study is presented, followed by background information on necessary tools. Next, we describe our proposed framework, followed by the AND/OR search-based algorithm. The subsequent section discusses related work, followed by preliminary experiments showing the effectiveness of the framework in diverse scenarios. Finally, we conclude our work.

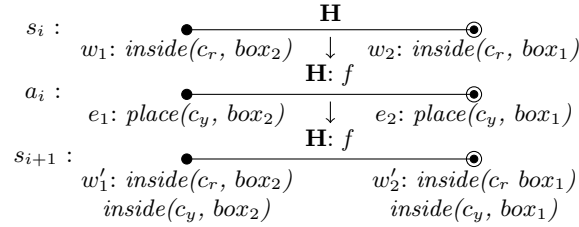


Fig. 3. We represent a state (s_i), action (a_i), and how applying a_i in s_i leads to next state ($s_{i+1} = s_i \otimes a_i$). f is a formula that captures if $\mathbf{H}\mathcal{E}\mathbf{R}$ were co-present when the events took place. Common facts for both worlds, such as $\text{opaque}(\text{box}_1)$, are not shown. Also, each world is fully defined, with either an atom or its negation holding true.

2 The Cube Organization Case Study

Take the case illustrated in Figure 2, in which the job of arranging cubes in boxes is shared by a human and a robot. The task is that the cubes from different tables must be arranged in different boxes.

Say only \mathbf{H} is capable of moving around and exhibits unpredictable behavior (*nondeterminism*), such as moving to the other table (*ot*) to retrieve cubes, while \mathbf{R} may continue to act. From the \mathbf{H} 's perspective, \mathbf{R} may move some or all of the cubes from the main table (*mt*) and place them into one of the boxes, or it may choose to take no action at all. Upon returning to the main table *mt*, \mathbf{H} may discover that some, none, or all of the cubes originally on *mt* are missing, indicating that they have been placed in one of the boxes.

If \mathbf{R} places some cubes from *mt* into one of the boxes, \mathbf{H} will only learn about this decision upon encountering transparent boxes. But when opaque, \mathbf{R} has several options: it can communicate, wait for \mathbf{H} to inquire, or select a remaining cube of *mt* to place in the correct box when \mathbf{H} and \mathbf{R} are co-present.

Planning is done from the robot's point of view by taking into account \mathbf{R} 's and \mathbf{H} 's task models, including their individual beliefs. The human collaborator has an approximation of the robot's model, which enables them to anticipate the robot's action. We later provide more details on these models and about their accuracy and falsity. These models are contained within the robot and are used in planning such that human behavior can only be estimated.

3 Background

Dynamic Epistemic Logic (DEL): We focus on epistemic languages ($\mathcal{L}_{\mathcal{K}}$), a state (s — comprising a set of worlds w_i), an action (a — comprising a set of events e_i), and state transitions (via the *cross product* \otimes operator) as derived from the literature [3, 4], with necessary trivial adjustments for our needs. For other basic concepts like *indistinguishability* and *equivalence* relations, *perspective shift*, and *truth* of epistemic formulas, readers are referred to the literature.

Here, we highlight only required DEL concepts to build it through examples based on our use case study. Recall the requirements for the task.

Example 1. Say the task is in the state s_i (Fig. 3), in which c_r is inside box_1 and both the boxes are opaque, and the robot holding c_y and the human comes back with c_w , and assesses the situation. We assume that the human can see the robot holding c_y . The epistemic state s_i such that $s_i \models K_{\mathbf{R}}inside(c_r, box_1)$, but concerning the human partner, $s_i \models \neg K_{\mathbf{H}}inside(c_r, box_1) \wedge \neg K_{\mathbf{H}}inside(c_r, box_2)$. Here, $K_i p$ represents agent i knows that the literal p is true.

Example 2. The next state s_{i+1} is such that, the epistemic action the robot will execute in epistemic state s_i is a_i that is placing c_y in the *correct* box. We describe how the next epistemic state s_{i+1} looks like *when* and *when not* \mathbf{H} & \mathbf{R} are *co-present* (i.e., whether they share this experience) during execution: An indistinguishability relation is only for \mathbf{H} when the formula f , e.g., $at(\mathbf{R}, place) \ \& \ not(at(\mathbf{H}, place))$, holds. \mathbf{R} always knows that the designated world is w_2 . That means if the human is co-present, they will know that the real world is w_2 .

Human-Aware Task Planning: We briefly discuss the human-aware task planning paradigm here. HATP/EHDA [5] comprises a dual Hierarchical Task Network (HTN) based task specification model. It is a recently proposed planner that estimates and emulates human decisions and actions for HRC. It solves problems in a turn-taking fashion, as formalized in our previous work [16, 15]. The following language adheres to this framework for easier understanding.

Consider the *human-aware task planning problem*, \mathcal{P}_{rh} and *implicitly coordinated joint solution* defined (Definitions 5 & 6, respectively) in [15].

\mathbf{R} and \mathbf{H} have their own *action models*, *beliefs*, *task networks* (agendas), *plans*, and more, collectively comprising $\mathcal{P}_{rh} = \langle \mathcal{M}_R, \mathcal{M}_H \rangle$. More specifically, \mathbf{R} has its estimated beliefs, s_0^r . We consider it as the real *knowledge* of “ground truth” in the planner’s reference, versus what the robot estimates to be believed by the human, s_0^h , by perspective taking. s_0^h may include a literal that is not true (*false belief* – e.g., *prop1* in Fig. 1) from \mathbf{R} ’s perspective and can be *corrected*.

We extended HATP/EHDA in [16], which adeptly anticipates human false beliefs for better collaboration based on (non-) shared execution experience.

To achieve that, *situation assessment* processes based on co-presence are integrated into the planning framework of HATP/EHDA. This enhances the planner to be pertinent to capturing what humans can observe and infer in their surroundings. It assesses the detrimental effects of humans’ incorrect beliefs on the task at hand. As a result, \mathbf{R} plans to communicate minimally and proactively.

We demonstrated in our previous work how to handle false beliefs (of first order) and situate the research broadly within the literature. In this paper, we extend and model knowledge up to level two, enabling us to handle \mathbf{HR} collaboration more realistically and allowing us to incorporate communication in a more practical way. We detail all these aspects as we proceed.

4 The EHATP Planning Framework

We consider that the human maintains an estimated model for the robot \mathcal{M}_H^R , which can be *incorrect* compared to \mathcal{M}_R .

The epistemic HATP (EHATP) framework considers three models: \mathcal{M}_R , \mathcal{M}_H and \mathcal{M}_H^R . While \mathcal{M}_R guides the planning of \mathbf{R} 's actions and \mathcal{M}_H helps estimate and emulate \mathbf{H} 's decisions and actions. But, using \mathcal{M}_H^R , \mathbf{H} “*expects*” and “*predicts*” certain robot behavior (from their own perspective) both, respectively, when they are co-present and when they are not.

Majority of the models’ components remain static, but for each model, its task network (tn_ϕ) and belief ($Bel(\phi)$) components are dynamic, where ϕ denotes an agent (or agent perspective). Except for belief, we assume that components like robot’s action models and task network are accurately estimated by \mathbf{H} . This allows us to focus on the key aspects relevant to this paper. For other incorrectly estimated components of \mathcal{M}_H^R , we suspect a possible generalization utilizing concepts developed in [32] and intend to explore this in the future.

4.1 Planning Workflow

We focus on only the dynamic parts. The initial epistemic state s_0 (with the only world to begin with and that is also the designated world w_d) is provided as an input. In general, each world w_j in an epistemic state s_i represents $\langle (Bel(R), tn_r), (Bel(H), tn_h), (Bel(R_H), tn_{r_h}) \rangle$. It also includes the only designated world w_d always known to \mathbf{R} . Note that these worlds are indistinguishable for \mathbf{H} , but human knows that the robot can always distinguish them and that the robot can identify w_d . Also, the human knows that, if w_j is the designated world, then $Bel^{ij}(R_H)$, is the reality as they do not have access to the facts appearing in $Bel^{ij}(R)$. Here, we consider that $Bel(H)$ is equal to $Bel(R_H)$, but they can be different from $Bel(R)$ and can contain false (human) beliefs.

The robot, an epistemic state s_i and possible worlds w_j in it are considered. We compute the set of all possible primitive actions, computed by all feasible decompositions, based on $(Bel(R), tn_r)_{ij}$, and whether it is different than the set of primitive actions based on the allowed decompositions w.r.t. $(Bel(R_H), tn_{r_h})_{ij}$. The idea is to align these decompositions, w.r.t. each w_j , in a way that the human can correctly estimate the progress the robot may achieve, thus utilizing the human’s capacity for anticipating. If there is a difference, we identify the *relevant* facts in $Bel^{ij}(R)$ that need to be corrected in $Bel^{ij}(R_H)$, to align the decompositions. To achieve that, we adapt our earlier approach presented in [16]. That is, one can plan minimal communication, possible to schedule ahead of time during offline planning when communication is allowed. Eventually, communication will also fix $Bel^{ij}(H)$, accordingly. However, $Bel^{ij}(H)$ and $Bel^{ij}(R_H)$ can still have *non-relevant* false beliefs compared to the ground truth ($Bel^{ij}(R)$).

Next, the planner computes the \mathbf{R} 's next real action based on its task network tn_R^{id} in the designated world w_d of s_i , we call it the *designated* event. It also computes other non-designated events based on respective decompositions in each world w_j of s_i . (An event and a possible real action including *noops* are used interchangeably.) In other words, the planner computes a set of all possible decompositions based on what \mathbf{H} can anticipate, that means by taking into account each $(Bel(R_H), tn_{r_h})_{ij}$. These are all the anticipated events that can happen due to the robot acting, but the designated event may or may not be

assessed depending on *co-presence*. All the decompositions (*i.e.*, the set of the first primitive action in each refinement) together form an epistemic action a_i .

Executing an Epistemic Action in a State: Based on the cross-product (\otimes) operation, it computes $s_{i+1} = s_i \otimes a_i$. We model within the planning algorithm (Algorithm 1, Line 8) as, if human is *co-present* – an idea adapted from the literature and described, then they can distinguish between the designated event (real action on ground) performed by **R** with the other estimated actions, otherwise human considers each event as the **R**'s possible next move (*ref. Fig. 1*).

When *co-present*, **H** assesses the execution of **R**'s real action, thus narrowing down the possibilities over w'_j 's in s_{i+1} – captured in \otimes operation (*ref Fig. 3*).

Within each world of the new epistemic state, belief components, *i.e.*, $Bel(R)$, $Bel(H)$, and $Bel(R_H)$ are updated corresponding to the possible robot action (either *real* or *anticipated*) that is a part of epistemic action a_i . Also, the task networks concerning \mathcal{M}^R and \mathcal{M}_H^R are updated in each world, accordingly.

When The Human Acts: Humans act only if they believe that their next real action, corresponding to a possible decomposition, is applicable in all possible worlds. That means, for each w_j in s_{i+1} , applicability of the action is examined w.r.t. every $(Bel(H), tn_h)_{i+1,j}$. At this stage, there arise *two* key issues. First, humans can act based on a false belief (if consistent throughout all the worlds), or a true belief w.r.t. the ground truth in every w_j . We handle false belief scenarios the way it is addressed in the literature, that is, by finding out relevant belief divergence and handling it via communication [16].

Second, we also know that a boolean state variable, p , that **H** is uncertain about at this stage, which holds only in some worlds, is due to disrupted shared execution experiences. If p is a precondition of the task refinement process, then **H** can initiate communication, or **R** can inform **H** about p . And, if *co-present*, **R** can also act to implicitly share p 's value such that there is some correlation between that action and p . Here, we focus on explicit communication, while sharing p 's value by changing the environment is left for the future.

Handling H&R Communication. We introduce two types of actions and they become a part of the deliberation process. First, *ask-p* – human inquires about p from **R**, and, second *inform-p* – **R** informs them of the status of p .

At this stage, we create two specialized versions of state s_{i+1} : one prioritizing human inquiries, *ask-p*, and the other prioritizing robot updates, *inform-p*. Communication tasks are adjusted into respective networks appropriately.

Situation Assessment. Assessing the status of a state property depends on a broader context, which determines whether it can be observed or only inferred by attending the action execution affecting it. Knowledge rules were used to address this aspect [31]. For example, **H** can view the current status of the variable $inside(c_r, box_1)$ as *true* if they meet the requirements of the rule's antecedent formula, *e.g.*, being at the main table, box_1 is transparent, and c_r is inside box_1 .

Algorithm 1 *AND/OR Planner using Breadth-First Search. Two key subroutines are **Situation Assessment** and **Expand**.*

```

1: Input: A HAETP task
2: Output: A joint solution or failure
3:  $root\_epi\_state \leftarrow \langle \mathcal{M}, w_d \rangle$   $\triangleright$  (focusing just on the essential parts) each world in
    $w \in W$  contains  $(\langle s_0^r, tn_{r,0} \rangle, \langle s_0^h, tn_{h,0} \rangle, \langle s_0^{r_h}, tn_{r_h,0} \rangle)$ 
4:  $queue.enqueue(root\_epi\_state)$ 
5: while  $queue$  is not empty do
6:    $curr\_node' \leftarrow queue.dequeue()$ 
7:    $curr\_node \leftarrow \text{Situation Assessment}(curr\_node')$ 
8:    $successors \leftarrow \text{Expand}(curr\_node)$ 
9:   if  $successors \neq \emptyset$  then
10:    for successor in  $successors$  do
11:       $queue.enqueue(successor)$ 
12:    end for
13:   else
14:      $eval(curr\_node)$   $\triangleright$  assign it DONE or DEAD
15:      $propagate\_revised\_status(curr\_node)$ 
16:   end if
17:   if  $root\_solved(root\_epi\_state)$  then
18:     return  $extract\_joint\_solution()$ 
19:   end if
20: end while
21: return failure

```

Definition 1. The situational assessment (*SA*) process considers the observation process (described in [31]) and a state s_i , producing an updated epistemic state s'_i . This process iterates over each world w_j in s_i , removing it if it can be distinguished from w_d by the human.

We roughly show how the *SA* process works in Figure 3. Let $w_1 = \langle (\dots), (\{inside(c_r, box_1)\}, \dots), (\dots) \rangle$ and $w_2 = \langle (\dots), (\{inside(c_r, box_2)\}, \dots), (\dots) \rangle$, where w_1 and w_2 represent distinct worlds within an epistemic state s_i , with w_1 as the designated world. When boxes are transparent and the human is co-located with the main table, the updated epistemic state s'_i contains only w_1 .

5 AND/OR Search based HAETP Planner

Algorithm 1 takes the EHATP problem as input, producing an output as either a *failure* or an optimal-worst case *joint solution*. We assume agents' real and auxiliary actions, e.g., *NOOP*, to be instantaneous and of equal cost, as in classical planning. Algorithm 1 is an implementation of the classic AND/OR search using rooted graphs. Following the search, when the *root* node is *DONE*, the joint solution policy is extracted, $extract_joint_solution()$, in Lines 17 & 18.

We consider the root node ($root_epi_state$) and the subsequent actor, either **R** or **H**, to begin the plan exploration (Line 3). Within the loop, in Line 6, we select a node/state from $queue$, and next call the *Situation Assessment*()

subroutine. What we present in Line 7 is a *lazy* approach for doing *SA*. At this stage, the planner already knows whether agents were co-present and whether the designated action could be assessed by the human. This helps the planner ignore those worlds that can be distinguished from the designated world (*Definition 1*). The scenario where a human transitions to the **R**'s location and subsequently becomes co-present is particularly interesting. Another significant subroutine, *Expand()*, previously discussed in the EHATP framework's planning workflow, is called in Line 8. The children created after **R** expands the popped node are *AND* nodes, *OR* nodes, and vice versa for the case where **H** expands this node.

If there are no successors for the current node, it indicates either a goal node or a dead end. In Line 14, we evaluate the current node. If both tn_r and tn_h are fully decomposed in the designated world of s_i , we execute an auxiliary action with a precondition that the task network is fully decomposed. If both agents can execute it individually, it signifies that agents believe that the shared task has been achieved. In Line 15, we propagate the status of this node to its immediate parent, who then propagates its status to its parent depending on whether it is an *AND* node or an *OR* node.

The Post-processing Step: Post-processing of the joint solution is done based on whether **H&R** are co-present. When *co-present*, we follow a turn-taking approach, but when *not* co-present, their actions are parallelized. This involves executing the *AND/OR* policy, and identifying where **H&R** separate and reunite. We then group the agents' actions in between to form pairs. This step assumes that actions performed in parallel do not interact when actors are apart.

6 Related Work

Human Robot Collaboration (HRC): Generating the robot's behavior while considering the existence of humans, known as human-aware planning and decision-making [7, 1, 35, 20, 22, 10, 9]. Also, it can do reasoning for task allocation [27, 26]. Communication is an essential key to successful HRC, which is used to align an agent's belief, clarify its decision or action, fix errors, etc. [33, 22]. We extend this research line but have not found studies addressing human anticipation and divergent beliefs in disrupted execution experiences.

Models, Planning Approaches and Solution Plans: Several models are applied in the context of HRC planning, including HTNs [21, 27, 6], POMDPs (Partially Observable MDPs) [34, 27, 35], *AND/OR* graphs [11], etc. HTNs use both abstract and non-abstract tasks to form a hierarchical network, while *AND/OR* graphs cover the causal links among subtasks and depth-first search is used for planning [17].

Epistemic Planning: The epistemic planning framework, in [4], holds promise for capturing key elements of ToM in autonomous robots. For HRC, the framework lays the groundwork for implicit coordination through perspective shifts [12].

<i>inst</i>	<i>K</i>	<i>comm</i>	<i>#states</i>	<i> W </i>	<i>#leaves</i>	<i>time (ms) × 10⁵</i>
P1 (2,2,T)	2	N	218	4	3	0.089
P2 (2,2,O)	2	Y	236	4	3	0.141
P3 (3,2,T)	2	N	1643	7	6	5.906
P4 (3,2,O)	2	Y	2003	7	6	9.816
P5 (3,2,T)	4	N	4107	14	5	99.81
P6 (3,2,O)	4	Y	5607	14	5	125.3
Cooking 1	2	Y	603	3	5	0.382
Cooking 2	3	Y	1054	4	5	1.474
Cooking 3	4	Y	1800	5	5	5.301

Table 1. The planner’s performance is evaluated using different metrics. *inst* denotes instance description, while *comm* indicates whether communication is employed. The metrics include the total number of explored states (*#states*), the worst-case number of worlds evaluated in a state (*|W|*), the number of traces in the final AND/OR solution tree (*#leaves*), and the execution time (measured in 10^5 ms). The table is divided into two parts, containing instances from two domains.

By adopting this planning framework and focusing on the robot’s perspective, it could serve as a basis for addressing the core problem we aim to solve with the shared mental model [25], albeit without considering false beliefs.

Explainable AI Planning (XAIP): In general, XAIP focuses on human-aware systems providing explanations of their behavior [18]. For example, a system might explain the correctness of its plan and the reasoning behind its decision based on its own model. The model reconciliation method, introduced in [32], assumes that the human possesses a disparate model of the robot’s behavior (\mathcal{M}_H^R instead of \mathcal{M}_R). This approach avoids unnecessary explanations by identifying the specific differences between the two models and only generates explanations where needed. Essentially, it suggests changes to the human’s model to optimize the robot’s plan based on that revised model. The proposed planning approach calculates the optimal explanations by identifying relevant discrepancies and communicating only the necessary information to align the models. We adopt a similar approach, computing relevant divergences to communicate only what is necessary to align the decompositions.

7 Empirical Evaluation

We implemented our planning system using Algorithm 1 in Python. It is based on the latest version of HATP/EHDA code [5].

As far as we know, there are not any standard planners available for comparison. However, we occasionally gauge the performance of our planner against the one presented in [16], which is more effective in scenarios with disrupted shared execution experiences. Nonetheless, it is worth noting that directly comparing their runtime would not be entirely fair, as our planner operates with a richer representation.

Domain Description: We test the planner in our use case domain and the cooking domain adapted from [16], on a variety of problems.

In the adapted scenario, both **H** and **R** are tasked with preparing dinner. The main activities involve *cutting* (R), *washing* (R) vegetables, *putting* (R) them on the stove with a pan and *seasoning* (R) them. Depending on the vegetables, seasoning can occur before or after they are placed in the pan, but always after washing. **H** is responsible for *bringing* (H) spices and other ingredients from the pantry and *mixing* (H) them in the pan, but only after the vegetables have been boiled (i.e., the effect of the *putting* action). *Serving* (H) dinner can only happen after the spices and seasoning have been mixed. Actors appear in (). Effects of washing and seasoning are non-observable.

In our context, the action of bringing ingredients is modeled such that it temporarily separates **H** from **R**. Despite the adaptation, **H** can still choose when to leave the kitchen for the pantry.

Analyzing the Impact of Parameter (K) and Non-Determinism: Algorithm 1 highlights a rapid growth in the size of the epistemic state in terms of the number of worlds which directly correlates with the *number of actions* (K) the robot can perform when the experience is not shared. The sequencing of actions significantly influences the range of potential worlds **H** expects to see.

K is considered to assess its impact on the planner’s performance. We assume that whenever the shared execution experience is disrupted, **R** can execute a *maximum* of K actions, including the option of doing nothing. For example, when the human is away to fetch the cube and has a *fixed* length and sequence of actions to perform. The exact number of real ontic actions **R** performs ranging from 0 to K , including which of those allowed ones and their potential sequences, will depend on the scenario at hand, environment dynamics (e.g., the observability factor), and the optimization criteria. The option for the robot to limit its real actions whenever required is integrated into the task description, aligning with the turn-taking nature of the underlying planner. Consequently, the planner is engineered to optimize the robot’s policy tree branching on uncontrollable human choices, including a communication action, to meet our objective.

Qualitative Analysis: In our use case domain, we explore different plan traces the planner can come up with depending on scenarios that arise. We start with two cubes, c_r and c_w , placed initially on tables mt and ot , respectively. Initially, there is only one designated world, w_d , in the initial epistemic state, s_0 . The environment otherwise remains unchanged. **H** can decide to go and retrieve the white cube, while the robot begins to work on other parts of the shared task.

Two plan traces from an AND/OR tree are shown in Figure 4. **H** starts to execute. **H**&**R** are co-present and the boxes are opaque. (*SA* is a systematic subroutine, but is shown only at relevant places.)

Let us focus on (a): after the human shifts focus to ot , both agents are not co-present until they reunite later in the trace, during which they act simultaneously. (*In this situation, agents must be at the same table and simultaneously focus on it to be considered co-present.*) In the first broad rectangular box, the

human moves to *ot*. They anticipate that the robot may have picked c_r or done nothing, but in reality, the robot picks c_r , resulting in two possibilities that will be maintained within the robot. Similarly, in the following box, the human picks c_w at *ot* and anticipates that if the robot had picked c_r , it could have placed it in one of the boxes or held onto it, or c_r is still on the table. Together, these create four possibilities, with the reality being that c_r is inside box_1 . At this point, the robot currently has no feasible action to execute, and the shared task has been not achieved yet, too. Upon the human’s return, as per their initial agreement on K , the robot has prepared itself with four possible worlds (with a designated world that only the robot knows). Perspective-taking and situation assessment help the robot eliminate two worlds where c_r is not on *mt* or in its hand.

We present two approaches to proceed with the task. In trace (a), the robot waits for human inquiry, while in trace (b), the human does nothing. Consequently, the robot decides to inform that box_2 is empty, resulting in only the designated world remaining. Here, $empty(box_2)$ is a precondition for the human to place c_w in it, which is true in one world and not another. Our proposed method considers a situation where the human waits for information without taking any action, such as nodding or making eye contact with the robot, as a distinct condition (*trace (b)*). Meanwhile, \mathbf{R} also has the option to modify the world to signal the variable’s status, which we intend to address in the future.

In our three-cube scenario, if the red cube is already in box_1 and the robot is holding the c_y , it can choose to place the c_y in box_1 in the presence of the human. This action results in the creation of a state with only the designated world as the next action ordered in the task network (tn_{r_n}) of that world does not allow the robot to execute $place(c_y, box_1)$. The robot can only be clever if it can fully explore its options. Depending on the situation, it might not always be preferable to place the yellow cube while the human is away and rely on communication or other means later on.

In contrast, in [16], \mathbf{R} communicates after agents become co-present again. This assumes that \mathbf{H} can choose to place c_w in box_1 due to their outdated belief w.r.t. changes they missed. In some practical cases, not communicating may lead to detrimental effects.

Quantitative Results and Analysis: Refer to Table 6. For the first domain, the first column indicates the instance number, along with the count of cubes and boxes, and whether the boxes are transparent (T) or opaque (O), respectively. In each instance, at least one cube is positioned on *ot*, which \mathbf{H} must retrieve. We show how the factor K influences the overall runtime.

We observe that $|W|$ and K contribute to longer runtime in both domains. Instances requiring communication tend to take slightly longer compared to those where communication is not required.

8 Conclusion

Our framework allows the robot to implement a ToM not only at execution time but also at planning time and hence explores what would be the beliefs of the

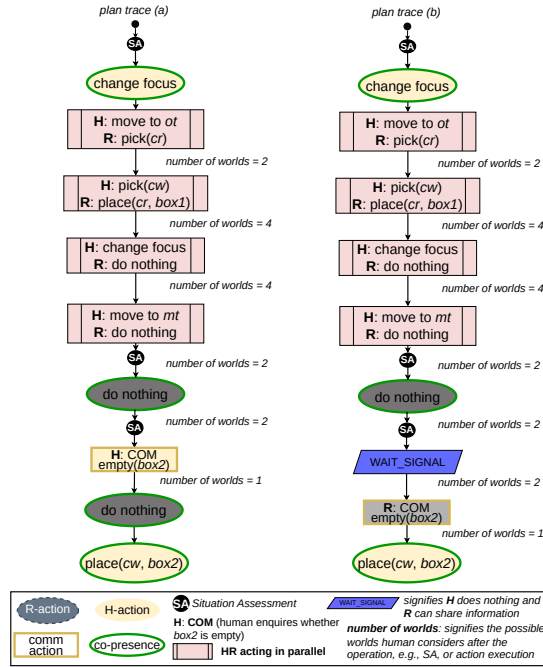


Fig. 4. Two branches from an AND/OR joint solution are shown: (a) R informs H proactively, thus leaving only the designated world for them to continue with $place(c_w, box_2)$. (b) R waits to inform H about the condition $empty(box_2)$.

human and the robot depending on which course of action. This is done thanks to the use of epistemic reasoning, the notion of shared experience, and observable and non-observable facts, which allow anticipation of H 's situation assessment along the various non-deterministic shared plan traces of H and R .

It allows R to adapt its choices to H 's diverging beliefs over time, for example by choosing to communicate to inform the human or elicit an action or to choose a particular context to act.

We acknowledge that scaling such abilities can pose complexity challenges for planners, which can be evident in [4]. Hence, we take care to precisely identify the context in which our approach can be effectively used which is dealing in a refined manner with short-term interactions and intricate H & R face-to-face situations. Also, we intend to test the current system in different domains with realistic H & R co-activities. We aim to enhance planner's practical efficiency and explore incremental task planning.

User Study: A user study has been conducted, validating the HATP framework, where planning anticipates human decisions and actions [13, 14]. Users acknowledged R 's pertinent decisions and compliance with their preferences.

Next, we are extending the plan execution framework to assess the pertinence of HATP equipped with epistemic reasoning as shown in relevant task contexts.

References

1. Alili, S., Warnier, M., Ali, M., Alami, R.: Planning and plan-execution for human-robot cooperative task achievement. In: Proc. of ICAPS (2009)
2. Baratta, A., Cimino, A., Gnoni, M.G., Longo, F.: Human robot collaboration in industry 4.0: a literature review. *Procedia Computer Science* **217**, 1887–1895 (2023)
3. Bolander, T., Andersen, M.B.: Epistemic planning for single and multi-agent systems. *J. Appl. Non Class. Logics* **21**(1), 9–34 (2011)
4. Bolander, T., Dissing, L., Herrmann, N.: DEL-based Epistemic Planning for Human-Robot Collaboration: Theory and Implementation. In: Proc. of KR (2021)
5. Buisan, G., Favier, A., Mayima, A., Alami, R.: HATP/EHDA: A Robot Task Planner Anticipating and Eliciting Human Decisions and Actions. In: ICRA (2022)
6. Cheng, Y., Sun, L., Tomizuka, M.: Human-aware robot task planning based on a hierarchical task model. *IEEE Robot Autom. Lett.* **6**(2), 1136–1143 (2021)
7. Cirillo, M., Karlsson, L., Saffiotti, A.: A human-aware robot task planner. In: Proc. of ICAPS 2009 (2009)
8. Coupeté, E., Moutarde, F., Manitsaris, S.: Gesture recognition using a depth camera for human robot collaboration on assembly line. *Procedia Manufacturing* **3**, 518–525 (2015)
9. Cramer, M., Kellens, K., Demeester, E.: Probabilistic decision model for adaptive task planning in human-robot collaborative assembly based on designer and operator intents. *IEEE Robotics AL* **6**(4), 7325–7332 (2021)
10. Darvish, K., Simetti, E., Mastrogiovanni, F., Casalino, G.: A hierarchical architecture for human–robot cooperation processes. *IEEE Transactions on Robotics* **37**(2), 567–586 (2020)
11. Darvish, K., Simetti, E., Mastrogiovanni, F., Casalino, G.: A hierarchical architecture for HR cooperation processes. *IEEE Trans. Robot.* **37**(2), 567–586 (2021)
12. Engesser, T., Bolander, T., Mattmüller, R., Nebel, B.: Cooperative epistemic multi-agent planning for implicit coordination. In: Proc. of WS on EPTCS (2017)
13. Favier, A.: Human-aware robot task planning: theory of mind and anticipation of human decisions and actions. Ph.D. thesis, Université de Toulouse (2024)
14. Favier, A., Alami, R.: A Model of Concurrent and Compliant Human-Robot Joint Action to Plan and Supervise Collaborative Robot Actions. In: ACS (2024)
15. Favier, A., Shekhar, S., Alami, R.: Anticipating false beliefs and planning pertinent reactions in human-aware task planning with models of theory of mind. *PlanRob 2023, The ICAPS Workshop* (2023)
16. Favier, A., Shekhar, S., Alami, R.: Models and algorithms for human-aware task planning with integrated theory of mind. In: RO-MAN. IEEE (2023)
17. Gombolay, M.C., Jensen, R., Stigile, J., Son, S., Shah, J.A.: Apprenticeship scheduling: Learning to schedule from human experts. In: Proc. of IJCAI (2016)
18. Kambhampati, S., Sreedharan, S., Verma, M., Zha, Y., Guan, L.: Symbols as a lingua franca for bridging human-ai chasm for explainable and advisable AI systems. In: AAAI. pp. 12262–12267. AAAI Press (2022)
19. Kourtis, D., Knoblich, G., Woźniak, M., Sebanz, N.: Attention allocation & task representation during joint action planning. *J. Cogn. Neurosci.* **26**, 2275–86 (2014)
20. Lallement, R., De Silva, L., Alami, R.: HATP: An HTN planner for robotics. arXiv preprint arXiv:1405.5345 (2014)
21. Lallement, R., de Silva, L., Alami, R.: HATP: hierarchical agent-based task planner. In: Proc. of AAMAS (2018)

22. Lemaignan, S., Warnier, M., Sisbot, E.A., Clodic, A., Alami, R.: Artificial Cognition for Social Human-Robot Interaction: An Implementation. *Artificial Intelligence* **247**, 45–69 (2017)
23. Liang, C.J., Wang, X., Kamat, V.R., Menassa, C.C.: Human-robot collaboration in construction: Classification and research trends. *Journal of Construction Engineering and Management* **147**(10), 03121006 (2021)
24. Nieto Agraz, C., Pflingsthor, M., Gliesche, P., Eichelberg, M., Hein, A.: A survey of robotic systems for nursing care. *Frontiers in Robotics and AI* **9**, 832248 (2022)
25. Nikolaidis, S., Shah, J.: Human-robot teaming using shared mental models. *ACM/IEEE HRI* (2012)
26. Ramachandruni, K., Kent, C., Chernova, S.: UHTP: A user-aware hierarchical task planning framework for communication-free, mutually-adaptive human-robot collaboration. *ACM Transactions on Human-Robot Interaction* (2023)
27. Roncone, A., Mangin, O., Scassellati, B.: Transparent role assignment and task allocation in human robot collaboration. In: *Proc. of ICRA* (2017)
28. Schmitz, L., Vesper, C., Sebanz, N., Knoblich, G.: Co-representation of others' task constraints in joint action. *Journal of Experimental Psychology: Human Perception and Performance* **43**(8), 1480 (2017)
29. Sebanz, N., Knoblich, G.: Prediction in joint action: What, when, and where. *Topics in cognitive science* **1**(2), 353–367 (2009)
30. Semeraro, F., Griffiths, A., Cangelosi, A.: Human-robot collaboration and machine learning: A systematic review of recent research. *Robotics and Computer-Integrated Manufacturing* **79**, 102432 (2023)
31. Shekhar, S., Favier, A., Alami, R., Croitoru, M.: A knowledge rich task planning framework for human-robot collaboration. In: *Proc. of 43rd SGAI* (2023)
32. Sreedharan, S., Chakraborti, T., Kambhampati, S.: Foundations of explanations as model reconciliation. *Artif. Intell.* **301**, 103558 (2021)
33. Tellex, S., Knepper, R.A., Li, A., Rus, D., Roy, N.: Asking for help using inverse semantics. In: *Proc. of RSS* (2014)
34. Unhelkar, V.V., Li, S., Shah, J.A.: Semi-supervised learning of decision-making models for human-robot collaboration. In: *Proc. of CoRL* (2019)
35. Unhelkar, V.V., Li, S., Shah, J.A.: Decision-making for bidirectional communication in sequential human-robot collaborative tasks. In: *Proc. of HRI* (2020)