



**HAL**  
open science

# Planning Concurrent Actions and Decisions in Human-Robot Joint Action Context

Anthony Favier, Rachid Alami

► **To cite this version:**

Anthony Favier, Rachid Alami. Planning Concurrent Actions and Decisions in Human-Robot Joint Action Context. HRI'24 Workshop Symbiotic Society with Avatars (SSA), Mar 2024, Boulder, United States. hal-04511507

**HAL Id: hal-04511507**

**<https://hal.science/hal-04511507>**

Submitted on 19 Mar 2024

**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.

# Planning Concurrent Actions and Decisions in Human-Robot Joint Action Context

Anthony Favier  
LAAS-CNRS Univ. de Toulouse ANITI  
Toulouse, France  
anthony.favier@laas.fr

Rachid Alami  
LAAS-CNRS Univ. de Toulouse ANITI  
Toulouse, France  
rachid.alami@laas.fr

## ABSTRACT

We address the task planning problem for robots collaborating with humans to achieve a shared task. In this context, humans often want to be free to choose *what* they want to do, as well as *how* according to their internal preferences. In this work, we propose a new human-aware task planner that explicitly takes into account a *model of execution* inspired by joint action schemes while capturing the uncontrollability factor of the human operators and social signals. Its output is the robot's behavioral policy, which describes the robot's action to perform in every state such that the action is congruent and compliant with the human's decision and action, including being passive, to be executed in parallel. The generation of the robot policy using our proposed framework is described, highlighting the planning process and the joint action model it is based on. We provide empirical evidence and brief user study results through simulated BlocksWorld scenarios, validating its ability to produce compliant concurrent robot behaviors.

## CCS CONCEPTS

- **Human-centered computing** → **Collaborative interaction**;
- **Computing methodologies** → *Robotic planning*; *collaboration and coordination*.

## KEYWORDS

Human-Robot Collaboration, Planning for Joint Action, Robot Planning, Dynamic Execution

### ACM Reference Format:

Anthony Favier and Rachid Alami. 2018. Planning Concurrent Actions and Decisions in Human-Robot Joint Action Context. In *Proceedings of 3rd International Workshop of Symbiotic Society With Avatars (SSA'24)*. ACM, New York, NY, USA, 7 pages. <https://doi.org/XXXXXXX.XXXXXXX>

## 1 INTRODUCTION

Human-robot collaboration (HRC) is a current research focus due to the rise in the number of robot-assisted applications where collaborative robots obviously contribute some value [27]. Robots collaborating with humans has a utility in real-world domains like household [30], or a robot assisting surgeons [15], collaborating at

workshops [28], and helping astronauts in the space [11]. Hence, robot's autonomy is useful to have an efficient collaboration with humans on a joint task [23]. Autonomy can be enhanced by planning the robot's behavior beforehand while also considering human operators, which is broadly known as human-aware planning and decision-making [4, 7, 30]. This is the subject of this paper.

In the context of HRC for a shared task [27], we believe, based on the literature on joint action [6, 13, 25, 26], that the key towards a seamless interaction is, to consider the human as an uncontrollable agent and to be fully and concurrently compliant with them. The human should not be dictated which action they must perform, as in [3, 23], and the robot must comply with possible human decisions and actions during execution.

To collaborate with such humans with their (hidden) preferences, one can devise an online planning scheme coupled with a plan executor. However, in order to maintain real-time performance, online planning generally keeps a restricted horizon. Therefore, decisions taken online may lead to a dead end or may not lead to an optimal solution. Offline planning overcomes these issues.

We propose a new *offline* task planner which extends an existing human-aware planning system addressed in [3]. The new planner is designed to take into account a *Model of Execution*, which is in the form of an automaton and mainly inspired by the joint actions schemes. The model captures humans' latitude in their decisions. The planner's output is the robot's behavioral policy, which describes the robot's action in a state such that the action is congruent and compliant with the human's decision in this state and their estimated preferences, and that it is also legal to be executed in parallel. Our framework also allows humans at any time during execution to share their preferences to adapt online and accordingly the robot's policy. In addition, our approach considers social signals to enhance execution by minimizing uncertainties. Both humans and robots issue signals to clarify situations such as performing an action, waiting for the other agent's necessary actions, or indicating a desire to remain passive.

In this paper, we discuss relevant related work before describing the joint action model of execution that is central to our approach. We then describe the task planning problem and then introduce our novel framework. The following two sections explain how the robot policy is generated by a three-step process: *exploration*, *characterization*, and *generation*. We empirically evaluated our approach in simulation. With a BlocksWorld scenario, we show how our approach can effectively produce a concurrent robot behavior that

---

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

SSA'24, March 15, 2024, Boulder, Colorado, USA

© 2018 Copyright held by the owner/author(s). Publication rights licensed to ACM.

ACM ISBN 978-1-4503-XXXX-X/18/06

<https://doi.org/XXXXXXX.XXXXXXX>

---

This work has been partially funded by the Agence Nationale de la Recherche through the ANITI ANR-19-PI3A-0004 grant.

is compliant with human online decisions and preferences. Eventually, before we conclude, we briefly present the user study we conducted validating our approach.

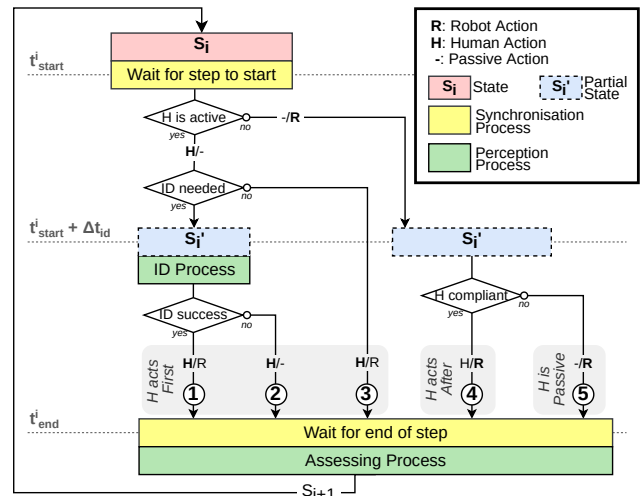
## 2 RELATED WORK

Using task models is a typical method of representing computer interaction and human behavior at a high abstraction level. Annett and Duncan were the first to use the hierarchical structure of human action to their advantage [2]. They claimed that up until a particular criterion is satisfied, activities can be described at various levels of abstraction. Thus, each task can be broken down into smaller ones that describe the steps a person would take to complete a higher-level task. Various task models have been realized in the HR collaborative planning context, e.g., hierarchical task networks (HTNs) [18, 23], POMDPs [23, 29, 30], AND/OR graphs [9], etc. A hierarchical network is created using HTNs (abstract and non-abstract tasks) and AND/OR graphs to represent the inner coupling links of the subtasks [12], and the plan search occurs in a depth-first manner. In [14], the authors show how uncertainty can be dealt with in the evolution of the environment and agent behavior. The challenge lies with the hidden and implied state of the human agent [30], especially in POMDPs for HRC.

Past works motivated us to use HTNs for the task/action specification for agents, as in HATP/EHDA [3] too, because they have similarities to real-world human decision-making.

Unlike ours, HATP frameworks extending HTNs consider humans controllable [1, 18], while in [23], the framework considers planning at multiple abstraction levels (with a single HTN) with humans. The latter is capable of basic reasoning for role assignment and task allocation. Like in our setting, robots plan under state uncertainty with partially observable human intentions modeled and tackled (mainly) at the primitive task level. This framework currently does not allow for *concurrent* planning/execution. On the other hand, we can see our execution automaton as an extension of the POMDP model (used in [23] for deciding role allocation at the primitive task level) for handling concurrency. Generally, existing frameworks (unlike ours) produce explicitly coordinated, shared HR plans that are assumed to be acceptable by humans, hence these techniques rely more on prior negotiation with humans and the replanning aspect for successful execution in reality.

There have been a few attempts to cater to concurrent execution, but they deal with explicit time to manage concurrency [5, 16]. However, due to the context uncertainties, it is often challenging to estimate accurately tasks and actions duration. In [4], the robot does not plan actions for humans but forecasts their actions/plans from their activities and bases its own decision on the distribution of possible human plans. Here robots can perform actions concurrently, estimating/managing the completion time of the agents' actions carefully. We can see the human activity recognition part as a form of our *identification (ID)* process of the automaton used. The robot needs such a plan/goal recognition technique to be compliant with the human's decisions. But, unlike ours, they do not consider an explicit shared goal among the agents, and hence humans are not concerned with stuff robots might be interested in during collaboration. We believe that a shared goal creates a different context



**Figure 1: The Model of Execution, in the form of an automaton and here simplified, captures the latitude of uncontrollable humans in their actions and guides our task planning approach. In this paradigm, the two agents can act concurrently but one is always compliant with the other's decision to act. Here, the human is always free to decide whether to start acting first, or after the robot, or not to act at all. To be compliant, the robot attempts to identify human decisions using perception and situation assessment as well as possible collaborative human signaling acts (e.g., gestures or speech).**

in HRC than the robot just being compliant with an estimated human's goals/plans. In another work, both *recognition* and *adaptation* take place simultaneously and comprehensively [19]. It deals with action scheduling of an already generated contingent plan comprising human's and robot's actions. It outputs schedules for the robot actions that can execute concurrently but to do that explicit temporal constraints are considered.

Another approach [22] proposes a communication free human-robot collaborative approach for an adaptive execution. In this work, the robot observes and supports human decisions, actively selecting the actions to optimize collaborative task efficiency. Unlike our approach, they introduce an extended collaborative HTN representation with role assignment for planning and state tracking during execution, which is more in line with [23]. In contrast, we employ two distinct HTNs for robot and human capabilities and use an AND/OR tree for exploration and execution tracking. While their online planning may enhance scalability, optimality is not guaranteed. Also, our scheme accommodates both verbal and non-verbal communication, allowing the human to express preferences that update the robot policy online.

## 3 THE MODEL OF EXECUTION

Our task planning approach uses a model of execution to improve the fluency and amenability of HRC. This model is in the form of an execution automaton as shown in Figure 1, and is based on several key notions and mechanisms borrowed from studies on

joint actions [17, 21], and adapted to Human-Robot Joint Action [6, 8]. The key idea is that co-acting agents co-represent the shared task context and integrate task components of their co-actors into their own task representation [24, 31]. Also, coordination and role distribution rely strongly on reciprocal information flow, e.g., social signals [8], prediction of other's next action [20]. By abstracting the model of execution and implementing it in our planning approach, our scheme provides human full latitude in their action choices.

Consider an example to clarify the execution automaton. Assume a human and a robot have to pick up two blocks,  $A$  and  $B$ , that both can reach. Both blocks can be picked up concurrently unless the agents try to pick the same block, the actions are conflicting. As a result, despite being executable in parallel, the actions are interdependent, and in order to avoid conflicts, one agent must be compliant with the other. However, if we consider a third block  $C$  that only the robot can reach, it can always pick up this block without any risk of conflicts with the human's choice.

In a state, a human decision can result in one of three outcomes. First, the human can choose to act first (*left subtree*). If the robot's best action is not in conflict with the human action (e.g., *pick C*), the robot can safely perform this action concurrently with the human operator (*branch 3*). However, if the robot's best action is either *pick A* or *pick B*, the human action must be identified first with a subroutine in order to be compliant with it. If this subroutine is successful the robot can perform any action which is congruent with the identified human action (*branch 1*). This includes the robot's choice to be *passive* and let the human act alone. However, if the robot is unable to identify the human action, it must remain passive in order to avoid potential conflicts (*branch 2*). Then, the human can either decide to be *passive* or to act after the robot (*right subtree*). In both cases, the human is *passive* at the beginning, making the robot to start performing alone a feasible action. While the robot is acting, the human is free to remain *passive* until the next step (*branch 5*), or to choose a congruent action to act concurrently (*branch 4*). As a result, the human can always choose to 1) act first, 2) act after the robot, or 3) not act at all. The robot will always be compliant with these online human decisions.

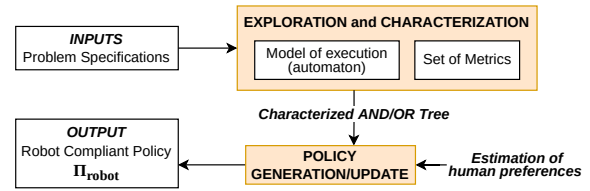
When both agents finish their actions, the step is considered as "over". Then, another subroutine assesses the new world state ( $s_{i+1}$ ), which is the result of the concurrent actions being executed in the state  $s_i$ , before repeating the whole process until the task is solved.

Note that if both agents are passive, e.g. the human decides to be passive when the robot cannot act, then the step is repeated.

## 4 PROPOSED FRAMEWORK DESCRIPTION

Before we delve into the details of the problem description, let us list in one place all the important assumptions we make with respect to human operators.

- Humans and robots are not equal. The robot is here to assist the human while aiming for efficiency.
- Humans are naturally uncontrollable agents, or at least they do not like otherwise. So, their behavior can only be estimated and emulated.
- Humans are cooperative, rational, and congruent but their commitment and preferences regarding the shared task are unknown and only estimated.



**Figure 2: An overview of the planning process. Taking as input the problem specification, the *Exploration and Characterization* process produces an AND/OR tree s.t. each branch is a successful plan and is characterized by a set of metrics. Then, an estimation of the human preferences enables ranking these branches and generating the robot's policy,  $\Pi_{robot}$ .**

- Humans are expected to signal the robot their decisions so that the robot can comply and act in parallel.
- Models describing human capabilities, world dynamics, and their understanding of common ground is given.

### 4.1 The Task Planning Problem

We address the human-aware task planning problem in a joint action context as described in [3]. The ability of the robot to simulate human behavior, beliefs, and decision-making processes enhances the traditional non-deterministic planning problem that this issue represents. Here, we consider distinct action models, beliefs, and agendas for the human and the robot.

### 4.2 The Planning Algorithm

Our planning framework is depicted in Figure 2. It extends the planning scheme appearing in [3], allowing the agents to act concurrently and produce solutions that always respect humans' online decisions and preferences.

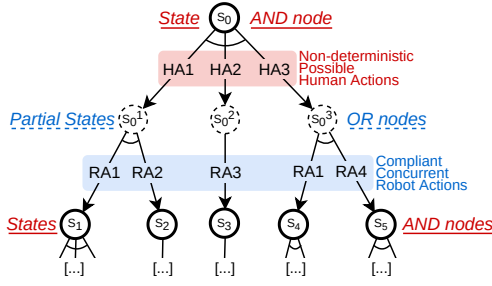
As Input, the framework takes the problem specifications consisting in initial agendas and beliefs, and action models. We assume that during planning, each feasible pair of concurrent actions *instantaneously* produces  $s_{i+1}$  from  $s_i$  as in classical planning. Hence, we make a clear distinction between the usages of the same automaton in planning time and online execution. Also, we do not perform reasoning with explicit temporal constraints or use durative actions in planning (unlike [4]).

The output of our approach is a compliant robot policy  $\Pi_{robot}$  which indicates for every possible human action in every state and for every the action the robot should perform concurrently to the human one. Each robot action has been selected to optimally satisfy the estimated human preferences regarding the task.

The search space must be computed offline. However, the policy can be generated online, and thus, it can be updated online with newly estimated human preferences.

## 5 EXPLORATION AND CHARACTERIZATION

The model of execution (Figure 1) is abstracted and used to model the assumptions w.r.t. the human operator, and explore relevant courses of actions to eventually produce the robot policy  $\Pi_{robot}$ .



**Figure 3: The AND/OR tree represents the search space. Meaningful states are connected to each other with one human action and one robot action executed concurrently. To visualize it better, while expanding a state  $s_i$  in search, each human action  $HA_k$  is followed by a partial state  $s_i^k$  from which the robot can only select actions  $\{RA_1, RA_2, \dots, RA_p\}$  non-conflicting with  $HA_k$  and can be executed concurrently in  $s_i$ .**

### 5.1 Computing Concurrent Actions

Concurrency is critical in our context enabling agents to act in parallel, but the non-determinism involved due to uncontrollable human decisions challenges the planning process differently. Beginning from the initial state, we depict in Figure 3 how our deliberation process expands and builds the search space in the form of an AND/OR tree. From each real state (AND) node  $s_i$ , based on the human agenda and action model ( $HTN_{human}$ ), all the actions that the human is likely to perform are retrieved. Each possible action expands the tree with partial state nodes  $s_i^j$  and updates the agenda and beliefs accordingly. Next, for each AND node, the robot actions returned by its action model ( $HTN_{robot}$ ) which are non-mutually exclusive to the particular human action expand the tree with new real state (AND) nodes.

Note that if an action does not cause any explicit change in the environment it is called *passive*. Such actions are useful in our context to allow agents to not always act when they can, and thus, explore further courses of actions. To anticipate the human choice to *not-act* (branch 5) a *passive* action is always inserted as a possible human action. It also helps to anticipate situations when the human wants to let the robot start and act concurrently. We also consider systematic *passive* robot action for two main reasons. First, if a mandatory identification fails, the robot must remain passive to avoid any conflicts with human online choices (branch 2). Second, the robot must be proactive to satisfy the human preferences. Hence, it is sometimes better to not perform an action (and be passive) to let the human perform it now or later.

### 5.2 Characterization Metrics

Each branch of the AND/OR tree is a feasible plan. To generate the robot policy we first characterize every plan by computing the following set of metrics:

- **Time of End of Human Duty:** Time step after which the human can remain passive.
- **Human Effort:** Sum of the cost of all human actions.
- **Time of Task Completion:** The time step at which the task is fully achieved.

- **Global Effort:** The sum of the costs of all actions.

Here, we consider that each non-passive action has a cost of 1.0, and that every passive action has a cost of 0.0.

We proposed a reasonable set of metrics that are relevant to characterize each plan but it could (and should) be complemented by other social metrics yet to be defined. The computed values are stored with the fully explored search space which is then referred to as “*Characterized AND/OR Tree*”. It constitutes the output of the offline exploration.

However, these metrics are not sufficient to determine the robot’s policy since it depends on how the human wants to prioritize these metrics while collaborating. This prioritization is referred to as the *estimated human preferences*.

## 6 THE ROBOT’S POLICY ( $\Pi_{robot}$ )

### 6.1 Estimation of human preferences

We are interested in finding an appropriate robot’s action in each (partial) state ( $s_i$  or  $s_i^j$ ) in the characterized AND/OR tree as depicted in Figure 3. To do so, we first have to be able to compare plans with each other to aim for the best one. We do so by using the estimated human preferences as a mean of comparison. Suppose no explicit preferences were told, so to start with, the robot estimates the human preferences. Such estimation is done by minimizing or maximizing each characterization metric in a defined priority order. For instance, one of the example human preferences we defined is *HUMAN-MIN-WORK* corresponding to the following ordering: Minimal Human Effort > Earliest End of Human Duty > Best Overall Cost > Earliest End of Task. When two plans have the same *Minimal Human Effort*, then the second metric in the ordered list is used to compare the plans, and so on.

### 6.2 Generation

Estimating the human preferences allow us to compare and rank each plan from best to worst. To improve legibility, we normalize the branches’ ranks to obtain a score such that a plan with the lowest rank has a score of 0.0 while the highest rank corresponds to a score of 1.0. This makes it easier to visualize the “quality” of a plan, regardless of the instance’s size. Then, by propagating the best reachable rank in the upper nodes we can identify the best robot action ( $RA^*$ ) to pick w.r.t. every human possible action.

The first draft of the robot’s policy maps each partial state  $s_i^j$  of the tree to the best robot action to perform ( $RA^*$ ), i.e.,  $\Pi_{robot} : \{s_i^j\} \rightarrow RA^*$ . The partial states are identified at execution time if either the *ID* process is successful or human remains passive. Such policy is complemented to handle identification failures, noted as the state  $\lambda$ , s.t.,  $\pi(\lambda) = PASS$ . Additionally, for a given  $s_i$ , if the identification is not needed then we can directly map the state itself  $s_i$  to the best robot action. In such cases, we override the policy defined above as follows  $\pi(s_i) = RA^*$ .

### 6.3 Execution

The execution of the policy stems from the execution automaton and is depicted in Algorithm 1.

---

**Algorithm 1:** The robot’s online behavioral policy.

---

```

(starting)  $s \leftarrow \text{initial state};$ 
while  $\neg \text{IsTaskAchieved}(s)$  do
   $\text{wait\_start\_step}();$ 
  if  $s \in \text{Domain}(\pi)$  then
     $\text{execute}(\pi(s));$ 
  else
    if  $\text{IsHumanPassive}()$  then
       $\text{execute}(\pi(s_i^j));$ 
    else
       $\text{id\_state} \leftarrow \text{IDProcess}(); \quad \# \in \{\lambda\} \cup \{s_i^j\}$ 
       $\text{execute}(\pi(\text{id\_state}));$ 
    end
  end
   $\text{wait\_end\_step}();$ 
   $s \leftarrow \text{AssessmentProcess}();$ 
end

```

---

## 7 EMPIRICAL EVALUATION

We provide results obtained after simulating symbolically the execution of robot policies produced with our approach.

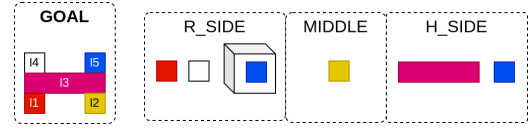
### 7.1 Simulated Experiment

The execution is symbolically simulated using an implementation of Alg. 1 and by mocking the human behavior. A human policy is produced using our approach with perfectly estimated human preferences. Then, the robot policy is produced by using an estimation, more or less accurate, of these preferences.

Here we have access to both the real human preferences and their estimation by the robot, thus we use the following notations. A plan’s score regarding the estimation of the robot will be referred to as an R-score. The plan’s score regarding the real human preferences will be referred to as the H-score. In practice, the R-score is an estimation of the H-score by the robot, and by maximizing the R-score the robot tries to maximize the H-score. Thus, it is important to note that it is acceptable to reach a low R-score as long as a high H-score is obtained.

Incorrect robot estimations of the human preferences can have various effects such as being *adversarial* to the real preferences. When the robot has a *correct* estimation, maximizing the R-score will necessary maximize the H-score as-well. However, in case of *incorrect* estimation, maximizing the R-score can lead to a wide range of H-score, including low scores. Eventually, an *adversarial* estimation implies that the R-score and H-score cannot be high at the same time. In such case, most often, a high R-score will necessarily lead to a low H-score, making the robot behavior unintentionally adversarial to the human preferences.

We evaluated our approach in the BlocksWorld domain. Figure 4 shows one problem instance. The human and the robot are on opposite sides of a table and their shared task is to stack colored cubes to match the given goal pattern. Initially, all colored cubes are disposed on the table into three zones. Each agent has a dedicated zone (*RZ* & *HZ*) and a common zone (*CZ*) is in the middle. Each



**Figure 4:** An instance of the BlocksWorld domain. The ideal plan is strongly influenced by the human preferences. For the earliest end of the task, the human prevents using the box. A lazy human will only place the required pink bar from their side. A human in a hurry will place concurrently the yellow cube to place the pink bar early and be able to leave.

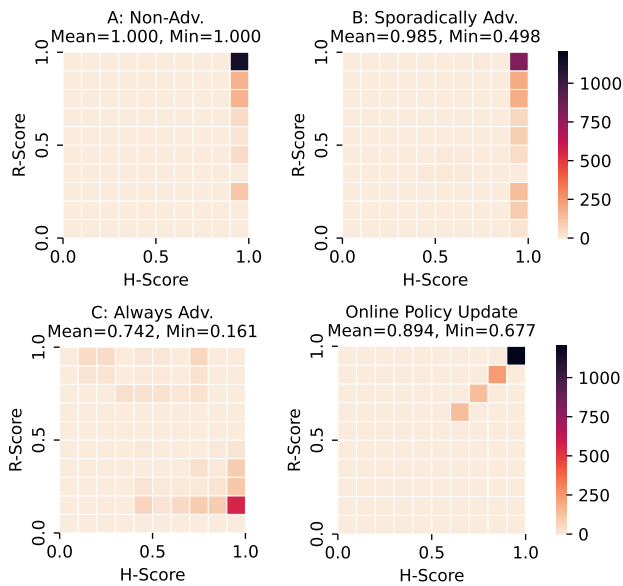
agent can only pick cubes from either their own zone and from the common zone *CZ*. There is a box in *RZ* in which cubes can be inserted. To pick such cubes the robot must first perform a dedicated action to open the box before being able to use the cubes inside as regular ones.

For the simulations, we first generated three problems with different initial states and shared tasks, and we produced for each their corresponding characterized AND/OR tree. After, we created three sets of pairs of human preferences and their estimation. In Set A, the estimations are never adversarial and mostly correct. Set B includes incorrect estimations which are sometimes adversarial. And Set C only contains adversarial estimations. Then, for each problem, we generated the robot and human policies for each preference-estimation pair from each set, and their execution was simulated. The R-score and H-score of every simulated executed plan are shown as heatmaps for each set of pairs in Figure 5.

### 7.2 Simulation Results

In Set A, the estimation of the robot is close to the real human preferences and is never adversarial. Some plans had an R-score lower than 1.0 showing that the estimation was not perfect. Yet, the compliance to the human actions and the non-adversarial choices of the robot allows to always reach a maximal H-score of 1.0. With Set B, the wrong estimation induced some adversarial choices of the robot preventing the human from always reaching a H-score of 1.0. But the average H-score obtained is 0.985, thus the human’s preferences were pretty much satisfied. Set C captures the worst possible estimations making the robot to always make choices detrimental to the H-score. This can be seen by the lower average H-score obtained of 0.742 and the very low minimal H-score obtained (0.161). Yet, we can notice that the average H-score is still high, showing that making the robot compliant with human actions is very effective and compensates significantly (of course not totally) for a very poor estimation of human preferences.

We can reasonably complement the human policy, which is so far only based on preferences, with a *rule* stating that whenever the robot performs an action that degrades significantly the best reachable H-Score then the human reacts by updating online the robot preferences, and align them with their own. The *Bottom-Right* part of Figure 5 shows the new distribution of scores obtained with Set C but using the rule. We notice that aligning preferences online avoids very low human scores (minimum of 0.677), and increases significantly the average human score as compared when the robot’s policy is not updated online (from 0.742 to 0.894). Hence, making



**Figure 5: R-scores and H-scores of the obtained executed plans after simulating the execution of the robot and the human policy generated by considering three problems and three sets of pairs of preferences/estimations. The estimations in each set are (A) Never, (B) Sporadically, and (C) Always adversarial. On the right, it is shown the scores obtained using an enhanced human policy that can correct online the robot’s estimation while using the set (C).**

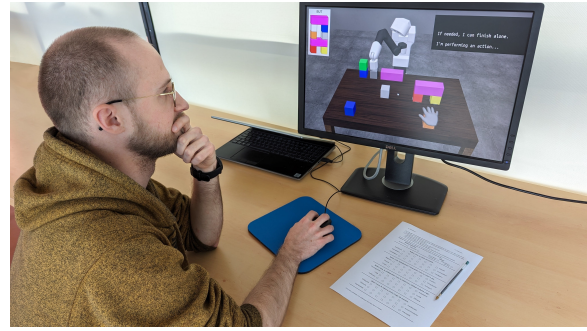
the robot compliant with online preferences is very effective in improving the quality of the joint plan executed.

Overall, we can see that the compliant robot behavior regarding both online human actions and preferences benefits the collaboration thanks to the high human scores obtained.

We can easily imagine and discuss about a counter-case where the robot never gives the initiative to humans and always execute its best action. In contrast with our approach, humans would be forced to comply with the robots which is less acceptable and more restricting. As it is evident from our simulation results in adversarial setups, this time, the robot would have a stronger impact on the solution than the human. Thus, wrong robot choices would significantly degrade the H-score without the human being about to react. As a result, being compliant and adjusting to online decisions and preferences can be seen as some social factor that robots should maximize, and our framework helps achieving that.

## 8 USER STUDY VALIDATION

We developed an interactive simulation in which a human can control an avatar and collaborate with a simulated robot following a policy generated with our approach. We used this simulator to conduct a user study with 25 participants (fig. 6) to validate our overall approach and the use of the model of execution, referred to as the Human-First (HF) regime. We used a baseline consisting of a robot executing the produced policy while always taking the



**Figure 6: A participant doing the experiment.**

initiative and forcing the human to be compliant, referred to as the Robot-First (RF) regime.

As preferences, the participants were given in each scenario the objective to either finish the collaborative stacking task fast or to free themselves as early as possible, letting the robot to finish the task alone. Then, we produced 3 pairs of preferences-estimations where only 1 pair has a correct estimation and the 2 others have adversarial estimations, producing 3 robot policies. Then, for each policy, the participants faced both a robot following the HF regime and then the baseline RF regime.

By asking the participants to answer a reduced version of the PeRDITA questionnaire [10] after each scenario, and by recording the execution and computing objective metrics, we were able to show the following. Among the 6 scenarios, the human preferences were significantly less satisfied in the two scenario using the baseline with erroneous estimations. This means that the preferences were satisfied similarly despite the erroneous estimations when using the HF regime or when the RF regime has a correct estimation.

Additionally, in two scenarios with lower preferences satisfaction, the participants perceived the interaction as significantly less *Positive*, the collaboration as significantly less *Adaptive* and *Efficient*, and their found the robot decisions significantly less *Appropriate* and *Accommodating*.

## 9 CONCLUSION

We addressed the complex challenge of concurrent task planning for a shared goal in the context of human-robot collaboration, acknowledging the inherent need for autonomy in humans’ choices of ‘what’ and ‘how’ aspects during task execution. Based on studies about joint action, we formulate an execution model, and we present a new human-aware task planner designed to accommodate the uncontrollability factor inherent in human agents while employing this execution model leveraging social signals to facilitate the exploration of human-robot joint actions and a smooth execution. As a result, the planner produces the behavioral policy for a robot that complies with online human decisions and their (online) provided estimated preferences, ensuring a sound execution of concurrent joint action in reality. We provide a detailed account of the novel planning process and joint action model, demonstrating its effectiveness through simulated BlocksWorld scenarios. We also briefly presents results obtained by a user study we conducted which validates our approach.

## REFERENCES

- [1] Rachid Alami, Aurélie Clodic, Vincent Montreuil, Emrah Akin Sisbot, and Raja Chatila. 2006. Toward Human-Aware Robot Task Planning. In *AAAI spring symposium: to boldly go where no human-robot team has gone before*.
- [2] John Annett and Keith D Duncan. 1967. Task analysis and training design. *Journal of Occupational Psychology* (1967).
- [3] Guilhem Buisan, Anthony Favier, Amandine Mayima, and Rachid Alami. 2022. HATP/EHDA: A Robot Task Planner Anticipating and Eliciting Human Decisions and Actions. In *Proc. of ICRA*.
- [4] Marcello Cirillo, Lars Karlsson, and Alessandro Saffiotti. 2009. A Human-Aware Robot Task Planner. In *Proc. of ICAPS 2009*.
- [5] Marcello Cirillo, Lars Karlsson, and Alessandro Saffiotti. 2009. Human-aware task planning for mobile robots. In *ICAR 2009*.
- [6] Aurélie Clodic, Elisabeth Pacherie, Rachid Alami, and Raja Chatila. 2017. Key Elements for Human Robot Joint Action. (2017), 159–177.
- [7] Martijn Cramer, Karel Kellens, and Eric Demeester. 2021. Probabilistic Decision Model for Adaptive Task Planning in Human-Robot Collaborative Assembly Based on Designer and Operator Intent. *IEEE Robotics AL* 6, 4 (2021), 7325–7332.
- [8] Arianna Curioni, Cordula Vesper, Günther Knoblich, and Natalie Sebanz. 2019. Reciprocal information flow and role distribution support joint action coordination. *Cognition* 187 (2019), 21–31.
- [9] Kourosh Darvish, Enrico Simetti, Fulvio Mastrogiovanni, and Giuseppe Casalino. 2021. A Hierarchical Architecture for Human-Robot Cooperation Processes. *IEEE Trans. Robotics* 37, 2 (2021), 567–586.
- [10] Sandra Devin, Camille Vrignaud, Kathleen Belhassein, Aurelie Clodic, Ophelie Carreras, and Rachid Alami. 2018. Evaluating the Pertinence of Robot Decisions in a Human-Robot Joint Action Context: The PeRDITA Questionnaire. In *2018 27th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. IEEE, Nanjing, 144–151. <https://doi.org/10.1109/ROMAN.2018.8525785>
- [11] Myron A Diftler, JS Mehling, Muhammad E Abdallah, Nicolaus A Radford, Lynn B Bridgwater, Adam M Sanders, Roger Scott Askew, D Marty Linn, John D Yamokoski, FA Permenter, et al. 2011. Robonaut 2-the first humanoid robot in space. In *Proc. of ICRA*.
- [12] Matthew C. Gombolay, Reed Jensen, Jessica Stigile, Sung-Hyun Son, and Julie A. Shah. 2016. Apprenticeship Scheduling: Learning to Schedule from Human Experts. In *Proc. of IJCAI*.
- [13] Jeremy Gordon, Guenther Knoblich, and Giovanni Pezzulo. 2023. Strategic Task Decomposition in Joint Action. *Cognitive Science* 47, 7 (2023), e13316.
- [14] Marcus Hörger, Hanna Kurniawati, and Alberto Elfes. 2019. Multilevel Monte-Carlo for Solving POMDPs Online. In *Proc. of ISRR*.
- [15] Mithun George Jacob, Yu-Ting Li, George A Akingba, and Juan P Wachs. 2013. Collaboration with a robotic scrub nurse. *Commun. ACM* 56, 5 (2013), 68–75.
- [16] Uwe Köckemann, Federico Pecora, and Lars Karlsson. 2014. Grandpa hates robots- interaction constraints for planning in inhabited environments. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 28.
- [17] Dimitrios Kourtis, Günther Knoblich, Mateusz Woźniak, and Natalie Sebanz. 2014. Attention allocation and task representation during joint action planning. *Journal of Cognitive Neuroscience* 26, 10 (2014), 2275–2286.
- [18] Raphaël Lallement, Lavindra de Silva, and Rachid Alami. 2018. HATP: hierarchical agent-based task planner. In *Proc. of AAMAS*.
- [19] Steven James Levine and Brian Charles Williams. 2014. Concurrent plan recognition & execution for human-robot teams. In *ICAPS*.
- [20] Luke McEllin, Natalie Sebanz, and Günther Knoblich. 2018. Identifying others' informative intentions from movement kinematics. *Cognition* 180 (08 2018), 246–258.
- [21] John Michael, Natalie Sebanz, and Günther Knoblich. 2016. Observing Joint Action: Coordination Creates Commitment. *Cognition* 157 (09 2016), 106–113.
- [22] Kartik Ramachandruni, Cassandra Kent, and Sonia Chernova. 2023. UHTP: A User-Aware Hierarchical Task Planning Framework for Communication-Free, Mutually-Adaptive Human-Robot Collaboration. *ACM Transactions on Human-Robot Interaction* (2023).
- [23] Alessandro Roncone, Olivier Mangin, and Brian Scassellati. 2017. Transparent role assignment and task allocation in human robot collaboration. In *Proc. of ICRA, 2017*.
- [24] Laura Schmitz, Cordula Vesper, Natalie Sebanz, and Günther Knoblich. 2017. Co-Representation of Others' Task Constraints in Joint Action. *Journal of experimental psychology. Human perception and performance* 43 (04 2017).
- [25] Natalie Sebanz, Harold Bekkering, and Günther Knoblich. 2006. Joint action: bodies and minds moving together. *Trends in cognitive sciences* 10, 2 (2006), 70–76.
- [26] Natalie Sebanz and Günther Knoblich. 2009. Prediction in Joint Action: What, When, and Where. *Top. Cogn. Sci.* 1, 2 (2009), 353–367.
- [27] Mario Selvaggio, Marco Cognetti, Stefanos Nikolaidis, Serena Ivaldi, and Bruno Siciliano. 2021. Autonomy in physical human-robot interaction: A brief survey. *IEEE Robotics Autom. Lett.* (2021).
- [28] Vaibhav V Unhelkar, Przemyslaw A Lasota, Quirin Tyroller, Rares-Darius Buhai, Laurie Marceau, Barbara Deml, and Julie A Shah. 2018. Human-aware robotic assistant for collaborative assembly: Integrating human motion prediction with planning in time. *IEEE Robotics and Automation Letters* 3, 3 (2018), 2394–2401.
- [29] Vaibhav V. Unhelkar, Shen Li, and Julie A. Shah. 2019. Semi-Supervised Learning of Decision-Making Models for Human-Robot Collaboration. In *Proc. of CoRL*.
- [30] Vaibhav V. Unhelkar, Shen Li, and Julie A. Shah. 2020. Decision-Making for Bidirectional Communication in Sequential Human-Robot Collaborative Tasks. In *Proc. of HRI*.
- [31] Motonori Yamaguchi, Helen J. Wall, and Bernhard Hommel. 2019. The roles of action selection and actor selection in joint task settings. *Cognition* 182 (2019), 184–192.