



HAL
open science

A Model of Concurrent and Compliant Human-Robot Joint Action to Plan and Supervise Collaborative Robot Actions

Anthony Favier, Rachid Alami

► **To cite this version:**

Anthony Favier, Rachid Alami. A Model of Concurrent and Compliant Human-Robot Joint Action to Plan and Supervise Collaborative Robot Actions. *Advances in Cognitive Systems (ACS)*, Nov 2024, Arlington (Virginia), United States. hal-04609442

HAL Id: hal-04609442

<https://hal.science/hal-04609442>

Submitted on 12 Jun 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.

A Model of Concurrent and Compliant Human-Robot Joint Action to Plan and Supervise Collaborative Robot Actions

Anthony Favier

ANTHONY.FAVIER@LAAS.FR

Rachid Alami

RACHID.ALAMI@LAAS.FR

LAAS-CNRS, ANITI, Universite de Toulouse, CNRS, 31031 Toulouse, France

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 and *how*, according to their internal preferences. We propose a concurrent and compliant joint action model based on social and collaborative signals. This model captures subtle possible agents' coordination and the human's inherent uncontrollability. We use an abstracted version of this model to explore relevant courses of action and guide our planning approach. The latter produces a behavioral policy that describes the robot's action to perform concurrently in every state, such that the action is congruent and compliant with any online human's decision and action, including being passive. The policy also aims to satisfy, at best, an estimation of the human internal preferences, which can be updated online. In this paper, we present our proposed joint action model and describe a novel human-aware task planning approach based on this model. Then, using BlocksWorld scenarios, we provide empirical evidence depicting the compliant concurrent robot behaviors produced by our approach. Additionally, we implemented our joint action model as an execution scheme into an interactive simulator used to conduct a user study. Compared with a baseline where robot decisions are always imposed on humans, our approach was significantly more appreciated and better satisfied the participants' preferences.

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 contribute some value (Selvaggio et al., 2021). Robots collaborating with humans has a utility in real-world domains like household (Unhelkar et al., 2020), assisting surgeons (Jacob et al., 2013), collaborating at workshops (Unhelkar et al., 2018), and helping astronauts in space (Diftler et al., 2011). Hence, robots' autonomy is helpful for efficiently collaborating with humans on a joint task (Roncone et al., 2017). Autonomy can be enhanced by planning the robot's behavior beforehand while also considering human operators, broadly known as human-aware planning and decision-making (Cirillo et al., 2009a; Cramer et al., 2021; Unhelkar et al., 2020). This is the subject of this paper.

In the context of HRC for a shared task (Selvaggio et al., 2021), we believe, based on the literature on joint action (Sebanz et al., 2006; Sebanz & Knoblich, 2009; Clodic et al., 2017; Gordon et al., 2023), that the key towards a seamless interaction is to consider the human as an uncontrol-

lable agent and to be concurrently fully compliant with them. The human should not be dictated which action they must perform, as in Roncone et al. (2017); Buisan et al. (2022), 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, 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 that extends an existing human-aware planning system addressed in Buisan et al. (2022). Our planning approach is guided by a concurrent and compliant joint action model based on social and collaborative signals exchanged between the agents. The model captures the agent’s coordination at execution while preserving humans’ latitude in their decisions. The planner’s output is a behavioral policy indicating in every state the robot’s action to execute concurrently with human activity. Each such action is congruent and compliant with the human’s decision and their estimated preferences in the given state. Our framework also allows humans to share their preferences at any time during execution and adapt the robot’s policy accordingly online. 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.

This paper first discusses relevant related work before describing an abstracted version of our joint action model that guides our planning approach. We then describe our task planning problem and introduce our novel framework. The following two sections explain how a three-step process generates the robot policy: *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 compliant with human online decisions and preferences. Additionally, we integrated the complete joint action model as an execution scheme into an interactive simulator. Using this simulator, we conducted a user study with 25 human participants. Before concluding, we discuss the results obtained by the study, which validates 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 & Duncan (1967) were the first to use the hierarchical structure of human action to their advantage. They claimed that 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) (Lallement et al., 2018; Roncone et al., 2017), POMDPs (Unhelkar et al., 2019; Roncone et al., 2017; Unhelkar et al., 2020), AND/OR graphs (Darvish et al., 2021), etc. In Hörger et al. (2019), 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 (Unhelkar et al., 2020), especially in POMDPs for HRC.

Past works motivated us to use HTNs for task and action specifications for agents because they resemble real-world human decision-making. The HATP framework (Alami et al., 2006; Lallement et al., 2018) extends HTNs for collaborative task planning producing a joint plan optimizing given performance and social criteria. In Roncone et al. (2017), the framework considers planning with humans at multiple abstraction levels (with a single HTN). The latter is capable of basic reasoning for role assignment and task allocation using a POMDP model. Like numerous existing frameworks, such approaches produce explicitly coordinated shared Human-Robot plans that humans are assumed to accept and follow. Hence, these techniques rely on prior negotiation, assuming in a way that humans are controllable, and on the replanning aspect for successful execution in reality. To tackle this issue, Buisan et al. (2022) proposed the HATP/EHDA framework, extending HATP, which explicitly models distinct human reasoning processes (\mathcal{M}^H) in addition to the robot ones (\mathcal{M}^R). This model is used to anticipate all actions humans are likely to perform and plan for each the best robot actions to execute. This accounts for humans’ inherent uncontrollability factor but is based on turn-taking without concurrent execution. We decided to extend the HATP/EHDA approach to plan collaborative concurrent robot actions while preserving human latitude of choice.

A few attempts have been made to plan concurrent execution using an explicit time representation (Cirillo et al., 2009b; Köckemann et al., 2014). However, estimating tasks and action duration is often challenging due to context uncertainties. In Cirillo et al. (2009a), 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, carefully estimating/managing the completion time of the agents’ actions. The robot needs a plan/goal recognition technique to comply with human 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 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 (Levine & Williams, 2014). It deals with the action scheduling of an already generated contingent plan comprising human and robot actions. It outputs schedules for the robot actions that can execute concurrently, but to do that, explicit temporal constraints are considered.

Another approach (Ramachandrani et al., 2023) proposes a communication-free human-robot collaborative approach for adaptive execution. In this work, the robot observes and supports human decisions, actively selecting 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 (Roncone et al., 2017). 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. MODEL of CONCURRENT and COMPLIANT JOINT ACTION for PLANNING

Our task planning approach is guided by a model of concurrent and compliant joint action execution, which improves HRC’s fluency and amenability. 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 (Michael et al., 2016; Kourtis et al., 2014), and adapted to Human-Robot Joint Action (Clodic et al., 2017; Curioni et al., 2019). 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 (Schmitz et al., 2017; Yamaguchi et al., 2019). Also, coordination and role distribution rely strongly on reciprocal information flow, e.g., social signals (Curioni et al., 2019), prediction of other’s next action (McEllin et al., 2018). By abstracting and implementing the joint action model in our planning approach, our scheme allows humans full latitude in their action choices, resulting in seamless collaboration.

Consider the following example to clarify the use of this joint action model. Assume a human and a robot must 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, actions *pick A* and *pick B* are interdependent despite being executable in parallel. One agent must be compliant with the other to avoid conflicts. Now, consider a third block *C* that only the robot can reach. It can always be picked up without the risk of conflicting with human choice.

Our step-based model begins by synchronizing agents, waiting for a human decision. The robot waits for a social signal from the human to start the step. This comprises starting to act or making an explicit hand gesture to indicate the desire to be passive. Note that, at this stage, the robot only detects the start of a human action, but not which particular action it is. After a few seconds without a signal, the robot considers the human passive and pursues. The human decision can result in one of three outcomes. First, the human can choose to act first (Fig.1: *left subtree*). If the robot’s best action is not in conflict with any human action (e.g., *pick C*), the robot can safely perform it concurrently with the human (Fig.1: *branch 3*). However, when the robot’s best action depends on the human decision (e.g., *pick A* or *pick B*), the human decision must be identified first with a subroutine (*ID process*) in order to be compliant with it. If this subroutine is successful, the robot can perform any action that is congruent with the identified human action (Fig.1: *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 (Fig.1: *branch 2*). Then, the human can decide to be *passive* or to act after the robot (Fig.1: *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 (Fig.1: *branch 5*), or to choose a congruent action to act concurrently (Fig.1: *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 (*SA Process*) assesses the current situation and the new world state s_{i+1} , which results from the concurrent actions being executed in the state s_i , before repeating the whole process until a goal state is reached where 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.

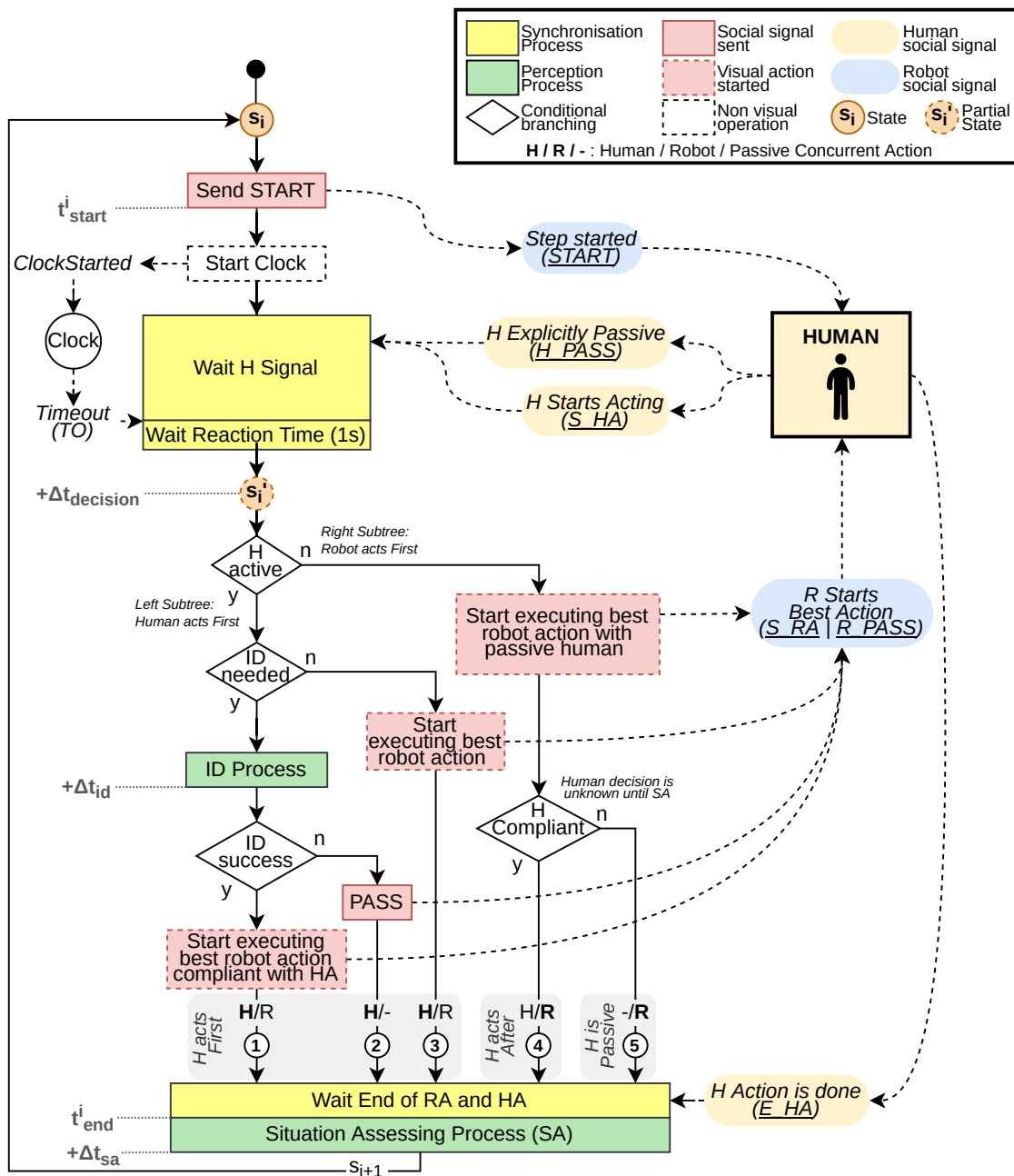


Figure 1. Abstracted Concurrent and Compliant Joint Action Model. In the form of an automaton, it captures the latitude of uncontrollable humans in their actions and guides our task planning approach. In this paradigm, the two agents can act concurrently, but one is always compliant with the other’s decision to act. Here, humans are always free to decide whether to start acting first (1,2,3), after the robot (4), or not to act at all (5). 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).

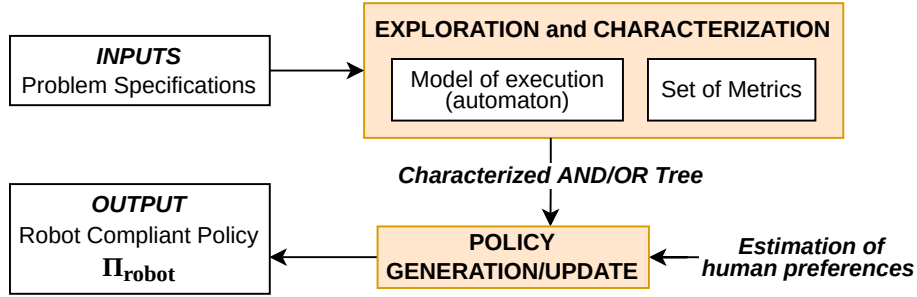


Figure 2. An overview of the planning process. Taking as *input* the problem specification, the *Exploration and Characterization* process produces an AND/OR tree. Each branch is a successful concurrent implicitly coordinated plan characterized by a set of metrics. Then, an estimation of the human preferences enables ranking these branches and generating the robot’s policy: Π_{robot} .

4. PROPOSED FRAMEWORK OVERVIEW

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.
- 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.

An overview of our planning framework is depicted in Figure 2. It extends the planning scheme appearing in Buisan et al. (2022), allowing the agents to act concurrently and produce solutions that always respect humans’ online decisions and preferences.

We address the human-aware task planning problem in a joint action context as described in Buisan et al. (2022). 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. As input, the framework takes the problem specifications consisting of distinct human and robot initial agendas, 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,

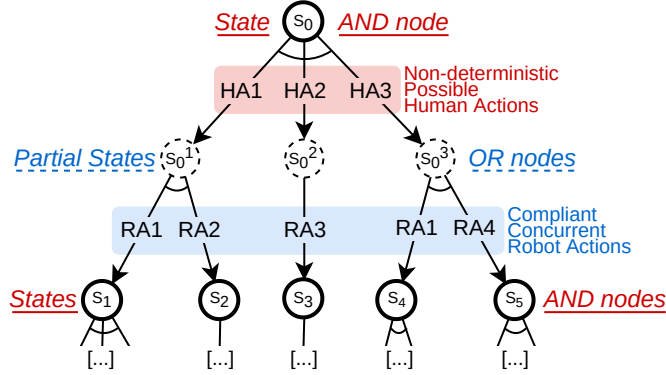


Figure 3. The AND/OR tree represents the search space. States are connected to each other with one human action and one concurrent compliant robot action. To visualize it better, while expanding a state s_i in search, each human action HA_k is followed by a partial state s_i^k from which the robot can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

we clearly distinguish 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 in Cirillo et al. (2009a)).

The output of our approach is a compliant robot policy Π_{robot} , which indicates for every possible human action in every state the action the robot should perform concurrently with the human one. Each robot action has been selected to optimally satisfy the estimated human preferences regarding the task, eventually solving the latter.

Our joint action model captures the assumptions regarding humans and guides our planning approach to exploring relevant courses of concurrent action. The next section describes how this offline exploration is conducted. The process of extracting the robot policy from the produced search space is described after. Since this process can be done online, the robot’s policy can be updated in real-time with newly estimated human preferences.

5. EXPLORATION and CHARACTERIZATION

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 as an AND/OR tree.

From each real state (AND) node s_i , based on the human agenda and action model (\mathcal{M}^H), 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 . Next, for each partial state (OR) node s_i^j , we compute the robot actions that are compliant and executable in parallel with the corresponding human action HA_k . This is done by using the robot agenda and action model (\mathcal{M}^R) to compute possible robot actions, and by

using the following assumption to check concurrency compatibility. We consider that if two actions A_1 and A_2 do not use the same shared resource (e.g., tool) and if they can be executed sequentially in both orders, then they are executable concurrently. The obtained valid robot actions expand the tree with new AND state nodes, including updated agendas and beliefs.

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 action. 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 human preferences. Hence, it is sometimes better not to perform an action (and be passive) and 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.

We consider non-passive actions to have a cost of 1.0 and passive ones to have 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 in the AND/OR tree leaves, 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 (Π_{robot})

6.1 Estimated Human Preferences

We are interested in finding an appropriate robot action for each possible human action, i.e., defining the best robot choice (OR node) for each possible human choice (AND node) in the characterized AND/OR tree. 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 an estimation of human preferences as a means of comparison. We assume this estimation is done externally and is given to our system. We consider that such preferences minimize or maximize each characterization metric in a defined priority order. For instance, one

example of human preferences 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

The estimated 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 (HA_k) or partial state (s_i^j).

The first draft of the robot’s policy maps each partial state s_i^j of the tree to the best robot action (RA^*) to perform, 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 λ , s.t., $\pi(\lambda) = PASS$. Additionally, for a given s_i , if the identification is not needed (*branch 3*), 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 $\pi(s_i) = RA^*$.

7. EMPIRICAL EVALUATION

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

7.1 Simulated Experiment

The execution of the policy stems from the execution automaton and is implemented in a simple controller. With it, we symbolically simulated the robot’s behavior and mocked the human’s 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 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 necessarily maximize the H-score as well. However, in the 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 cases, most often, a high R-score will necessarily lead to a low H-score, making the robot’s behavior unintentionally adversarial to human preferences.

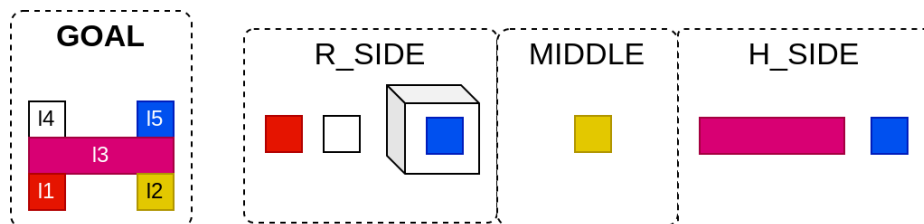


Figure 4. An instance of the BlocksWorld domain. The ideal plan is strongly influenced by human preferences. For the earliest end of the task, the human prevents using the box. A lazy human will only place the required pink bar from their side. A human in a hurry will place concurrently the yellow cube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

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 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 produced their corresponding characterized AND/OR trees. After, we created three sets, including 576 different 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 simulated their execution. The R-score and H-score of each of the 1728 simulated plans are shown as heatmaps for each set of pairs in Figure 5 and are discussed just below.

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 imperfect. Yet, the robot’s compliance with human actions and non-adversarial choices always allows it to reach an H-score of 1.0. With Set B, the wrong estimation induced some adversarial robot choices preventing the human from always reaching an 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 where the robot always makes choices detrimental to the H-score. This can be seen by the lower average H-score (0.742) and the very low minimal H-score (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

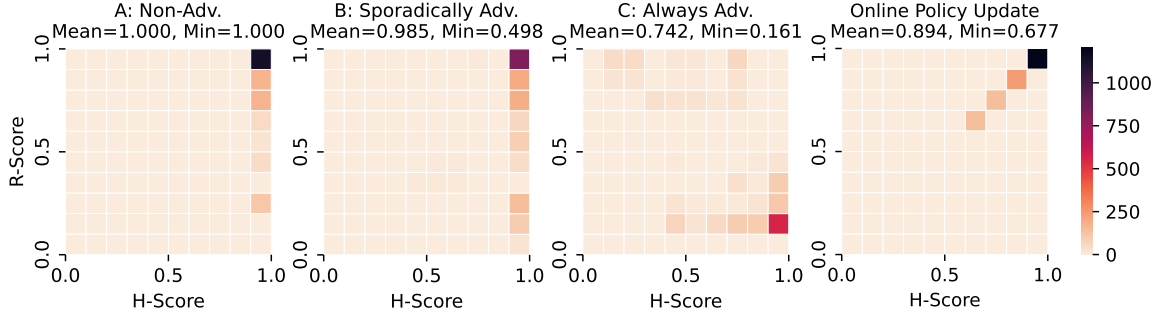


Figure 5. R-scores and H-scores of the obtained executed plans after simulating the execution of the robot and the human policy generated by considering three problems and three sets of pairs of preferences/estimations. The estimations in each set are (A) Never, (B) Sporadically, and (C) Always adversarial.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

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 significantly increases the average human score compared to when the robot’s policy is not updated online (from 0.742 to 0.894). Hence, making the robot compliant with online preferences effectively improves 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 a counter-case as a baseline in which the robot never gives the initiative to humans and always executes 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 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, compliance and adjusting to online decisions and preferences can be seen as a social factor that robots should maximize, and our framework helps achieve that.

8. USER STUDY VALIDATION

In addition to the empirical results described above, we conducted a user study to provide more results and validate our proposed joint action model and planning approach. For that, we refined the proposed joint action mode. In addition to the robot automaton, the complete model explicitly depicts another automaton describing the human’s possible decisions. The two automatons are synchronized with explicit social signals exchanged between the two agents. The abstracted model guides our planning approach to better anticipate execution. Meanwhile, the complete model has been implemented as an execution scheme, allowing us to supervise the execution of policies produced by our approach. This execution scheme has been integrated into an interactive simulator, in



Figure 6. A participant in a collaborative scenario with the simulated robot.

which a human can control an avatar in real-time and collaborate with a simulated robot following a policy generated by our approach. A BlocksWorld problem with 8 blocks is used.

We used this simulator to conduct a user study with 25 participants (fig. 6) to validate our overall approach, and the use of our joint action model referred to as the Human-First (HF) regime. We used a baseline corresponding to the counter-case discussed at the end of the previous section. It consists of a robot executing the produced policy while always taking the initiative and forcing the human to comply, referred to as the Robot-First (RF) regime. A detailed execution is shown here¹.

As preferences, the participants were given in each scenario the objective to either (1) finish the collaborative stacking task fast or (2) to free themselves as early as possible, letting the robot to finish the task alone. Then, we produced 3 pairs of preferences-estimations: Pair A corresponds to objective (1) with a correct estimation; Pair B to objective (1) with an erroneous estimation; Pair C to objective (2) with an erroneous estimation. Incorrect estimation of human preferences can lead to adversarial robot decisions w.r.t. the actual preferences. However, as shown by the previous results, we expect the robot’s compliance to allow humans to impose their decisions to satisfy anyway their preferences. A robot policy is generated for each pair. Then, for each policy, the participants faced a robot following the HF regime and the baseline RF regime, leading to 6 different scenarios (S1-S6). Each participant faced a randomized ordering of the 6 scenarios.

After each scenario, we saved and computed objective execution metrics. This permits to evaluate how much human preferences are satisfied by measuring the task completion time in pairs A and B, and the time for the human to be free in pair C. These metrics are depicted in figure 7. Pair A is the only one where the difference between the HF and RF regimes isn’t significant. RF even tends to perform better. However, in case of erroneous estimation (pairs B and C), HF satisfy significantly better the preferences. Hence, erroneous estimations seem to have a significant detrimental effect

1. <https://cloud.laas.fr/index.php/s/vX1YRjSnbngcBDQ>

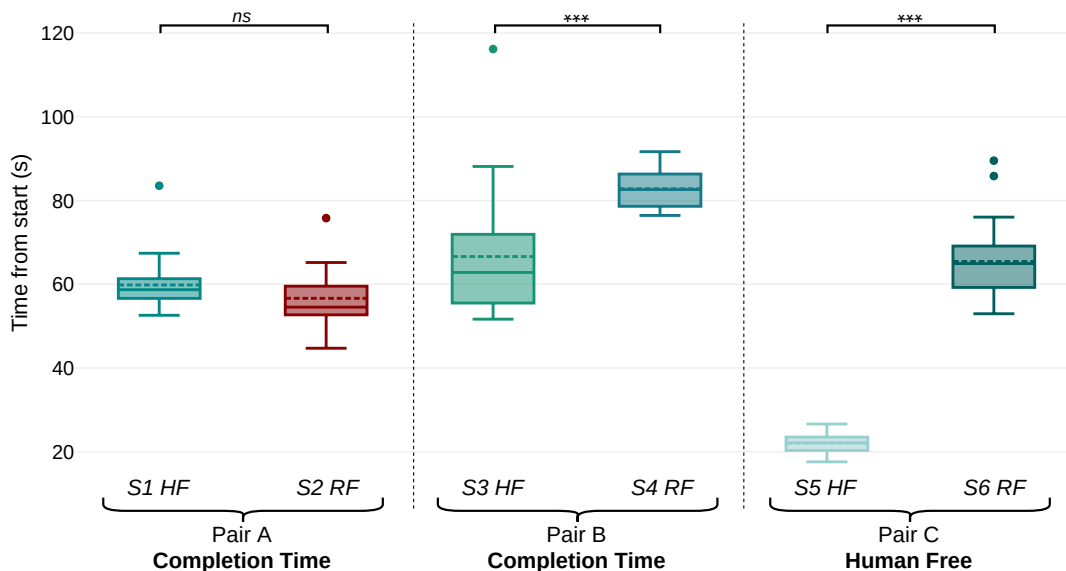


Figure 7. Execution metrics representing participants’ preferences satisfaction. S4 & S6 are significantly worse.

on preference satisfaction when using the RF regime. By being compliant with the human online decisions, HF satisfies better human preferences, even in case of erroneous estimation.

Additionally, after each scenario, participants were asked to answer a reduced version of the PeRDITA questionnaire proposed in Devin et al. (2018). It comprises 12 numbered items depicted in table 1. These items are organized into 4 categories, each evaluating respectively the overall robot perception, the interaction, the collaboration, and the robot actions.

Robot perception		Interaction		Collaboration		Acting	
1	Responsive	4	Positive	7	Adaptive	10	Appropriate
2	Competent	5	Simple	8	Useful	11	Accommodating
3	Intelligent	6	Clear	9	Efficient	12	Predictable

Table 1. Questionnaire 12 numbered items, grouped in four categories.

The questionnaire average answers and associated standard deviations are shown in figure 8. The answers confirm the objective results above. In pair A with correct estimation, the RF regime (S2) is perceived as good as the HF regime (S1). However, with erroneous estimation, the RF scenarios (S4 and S6) are perceived as significantly worse. Thanks to an ANalysis Of the VAriance (ANOVA) with repeated measures, we showed that in S4 and S6, the participants perceived: the interaction as significantly less *Positive*, the collaboration as significantly less *Adaptive* and *Efficient*, and they found the robot decisions significantly less *Appropriate* and *Accommodating*. These findings suggest that the RF regime can be pertinent, but the lack of latitude given to humans can be detrimental. The HF regime makes the participants feel more in control, which is appreciated and allows them to satisfy their preferences much better.

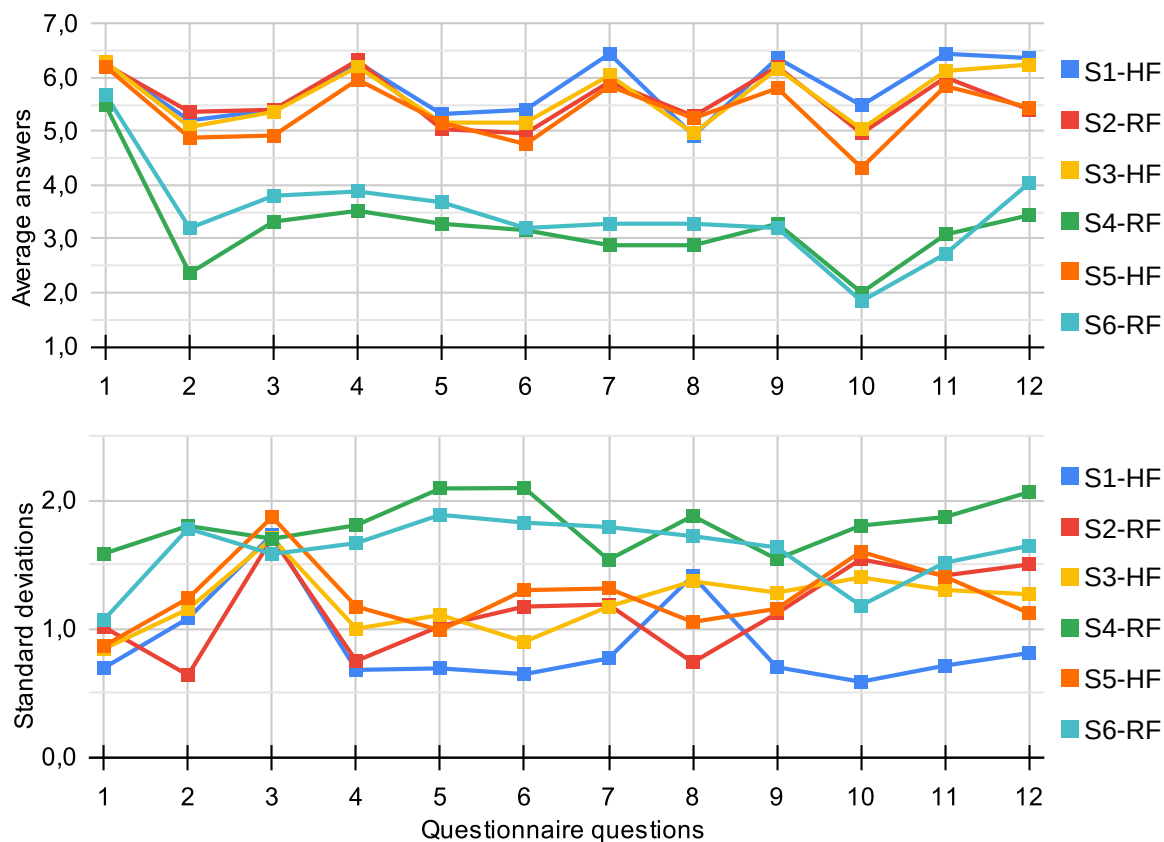


Figure 8. Average questionnaire answers (top) and associated standard deviations (bottom) for each scenario. The two RF scenarios with erroneous estimation (S4 and S6) have significantly lower answers and higher standard deviations than the HF regime.

9. CONCLUSION

We addressed the complex challenge of concurrent task planning for a shared goal in the context of human-robot collaboration, acknowledging humans’ inherent need for autonomy in their choices of ‘what’ and ‘how’ during task execution. Based on studies about joint action, we formulate an execution model and present a new human-aware task planner designed to accommodate the uncontrollability factor inherent in human agents. This execution model leverages social signals to facilitate the exploration of human-robot joint actions and smooth execution. As a result, the planner produces the behavioral policy for a robot that complies with online human decisions and their estimated dynamic 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 and demonstrate its effectiveness through simulated BlocksWorld scenarios. We also present results obtained after conducting a user study with a baseline in which the robot’s decisions are imposed on humans. We showed that our approach leads to better satisfaction of human preferences and a more positive, adaptive, and appropriate collaboration.

Acknowledgements

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

References

- Alami, R., Clodic, A., Montreuil, V., Sisbot, E. A., & Chatila, R. (2006). Toward human-aware robot task planning. *AAAI spring symposium: to boldly go where no human-robot team has gone before*.
- Annett, J., & Duncan, K. D. (1967). Task analysis and training design. *Journal of Occupational Psychology*.
- Buisan, G., Favier, A., Mayima, A., & Alami, R. (2022). HATP/EHDA: A Robot Task Planner Anticipating and Eliciting Human Decisions and Actions. *ICRA*.
- Cirillo, M., Karlsson, L., & Saffiotti, A. (2009a). A human-aware robot task planner. *ICAPS*.
- Cirillo, M., Karlsson, L., & Saffiotti, A. (2009b). Human-aware task planning for mobile robots. *ICAR*.
- Clodic, A., Pacherie, E., Alami, R., & Chatila, R. (2017). Key elements for human-robot joint action. *Sociality and normativity for robots: philosophical inquiries into human-robot interactions*.
- Cramer, M., Kellens, K., & Demeester, E. (2021). Probabilistic decision model for adaptive task planning in human-robot collaborative assembly based on designer and operator intents. *IEEE Robotics AL*.
- Curioni, A., Vesper, C., Knoblich, G., & Sebanz, N. (2019). Reciprocal information flow and role distribution support joint action coordination. *Cognition*.
- Darvish, K., Simetti, E., Mastrogiovanni, F., & Casalino, G. (2021). A hierarchical architecture for human-robot cooperation processes. *IEEE Trans. Robotics*.
- Devin, S., Vrignaud, C., Belhassein, K., Clodic, A., Carreras, O., & Alami, R. (2018). Evaluating the Pertinence of Robot Decisions in a Human-Robot Joint Action Context: The PeRDITA Questionnaire. *IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. Nanjing: IEEE.
- Diftler, M. A., et al. (2011). Robonaut 2-the first humanoid robot in space. *ICRA*.
- Gordon, J., Knoblich, G., & Pezzulo, G. (2023). Strategic task decomposition in joint action. *Cognitive Science*.
- Hörger, M., Kurniawati, H., & Elfes, A. (2019). Multilevel monte-carlo for solving pomdps online. *ISRR*.
- Jacob, M. G., Li, Y.-T., Akingba, G. A., & Wachs, J. P. (2013). Collaboration with a robotic scrub nurse. *Communications of the ACM*.
- Köckemann, U., Pecora, F., & Karlsson, L. (2014). Grandpa hates robots-interaction constraints for planning in inhabited environments. *AAAI Conference on Artificial Intelligence*.

- Kourtis, D., Knoblich, G., Woźniak, M., & Sebanz, N. (2014). Attention allocation and task representation during joint action planning. *Journal of Cognitive Neuroscience*.
- Lallement, R., de Silva, L., & Alami, R. (2018). Hatp: hierarchical agent-based task planner. *AAMAS*.
- Levine, S. J., & Williams, B. C. (2014). Concurrent plan recognition & execution for human-robot teams. *ICAPS*.
- McEllin, L., Sebanz, N., & Knoblich, G. (2018). Identifying others' informative intentions from movement kinematics. *Cognition*.
- Michael, J., Sebanz, N., & Knoblich, G. (2016). Observing joint action: Coordination creates commitment. *Cognition*.
- Ramachandruni, K., Kent, C., & Chernova, S. (2023). Uhtp: A user-aware hierarchical task planning framework for communication-free, mutually-adaptive human-robot collaboration. *ACM Transactions on Human-Robot Interaction*.
- Roncone, A., Mangin, O., & Scassellati, B. (2017). Transparent role assignment and task allocation in human robot collaboration. *ICRA*.
- Schmitz, L., Vesper, C., Sebanz, N., & Knoblich, G. (2017). Co-representation of others' task constraints in joint action. *Journal of experimental psychology. Human perception and performance*.
- Sebanz, N., Bekkering, H., & Knoblich, G. (2006). Joint action: bodies and minds moving together. *Trends in cognitive sciences*.
- Sebanz, N., & Knoblich, G. (2009). Prediction in joint action: What, when, and where. *Topics in cognitive science*.
- Selvaggio, M., Cognetti, M., Nikolaidis, S., Ivaldi, S., & Siciliano, B. (2021). Autonomy in physical human-robot interaction: A brief survey. *IEEE Robotics Autom. Lett.*
- Unhelkar, V. V., Lasota, P. A., Tyroller, Q., Buhai, R.-D., Marceau, L., Deml, B., & Shah, J. A. (2018). Human-aware robotic assistant for collaborative assembly: Integrating human motion prediction with planning in time. *IEEE Robotics and Automation Letters*.
- Unhelkar, V. V., Li, S., & Shah, J. A. (2019). Semi-supervised learning of decision-making models for human-robot collaboration. *CoRL*.
- Unhelkar, V. V., Li, S., & Shah, J. A. (2020). Decision-making for bidirectional communication in sequential human-robot collaborative tasks. *HRI*.
- Yamaguchi, M., Wall, H. J., & Hommel, B. (2019). The roles of action selection and actor selection in joint task settings. *Cognition*.