

Human-Robot Collaborative Multi-Agent Path Planning using Monte Carlo Tree Search and Social Reward Sources

Marc Dalmasso, Anaís Garrell, José Enrique Domínguez, Pablo Jiménez and Alberto Sanfeliu

Abstract—The collaboration between humans and robots in an object search task requires the achievement of shared plans obtained from communicating and negotiating. In this work, we assume that the robot computes, as a first step, a multi-agent plan for both itself and the human. Then, both plans are submitted to human scrutiny, who either agrees or modifies it forcing the robot to adapt its own restrictions or preferences. This process is repeated along the search task as many times as required by the human. Our planner is based on a decentralized variant of Monte Carlo Tree Search (MCTS), with one robot and one human as agents. Moreover, our algorithm allows the robot and the human to optimize their own actions by maintaining a probability distribution over the plans in a joint-action space. The method allows an objective function definition over action sequences, it assumes intermittent communication, it is anytime and suitable for on-line replanning. To test it, we have developed a human-robot communication mobile phone interface. Validation is provided by real-life search experiments of a Parcheesi token in an urban space, including also an acceptability study.

I. INTRODUCTION

Collaborative human-robot planning is an important functionality for many application tasks as, for example, search & rescue [17] or object searching. It integrates the teammates' contribution to the task, and they have to make their own plans, communicate and negotiate the best plan. Moreover, these processes have to be done on-line, and often in real-time. In mixed human-robot teams the complexity increases, as robots and humans do not share the same knowledge, but they have to generate a common shared plan. The robot may not know the human plan, and one way to solve this is by predicting the human's intention and sharing such prediction with the human in order to receive some feedback. Alternatively, the human anticipates his/her plan, allowing the robot to look for the best shared plan. In this work we focus on shared path planning, and assume that the robot first creates its own plan using a prediction of the human path plan, and shares the resulting plan with the human. Then the human revises both plans and may either agree with them, or propose a modified plan, in which case the robot will have to modify its own plan.

We propose to use a multi-agent planner in human-robot collaborative settings. Building a shared plan for all the

Work supported under the Spanish State Research Agency through the María de Maeztu Seal of Excellence to IRI (MDM-2016-0656), ROCOTRANS project (PID2019-106702RB-C21 / AEI / 10.13039/501100011033), TERRINet (H2020-INFRAIA-2017-1-two-stage-730994) and AI4EU (H2020-ICT-2018-2-825619)

The authors are with the Institut de Robòtica i Informàtica Industrial (CSIC-UPC). Llorens Artigas 4-6, 08028 Barcelona, Spain. {mdalmasso, agarrell, jdominguez, jimenez, sanfeliu}@iri.upc.edu

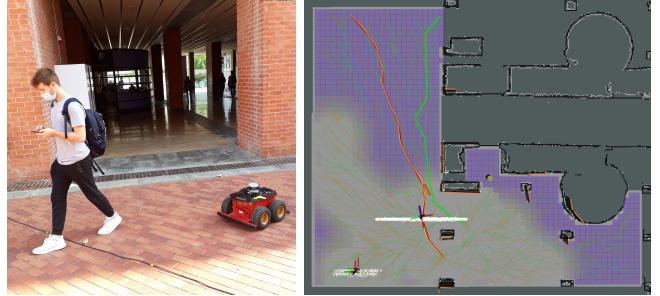


Fig. 1. Multiagent Human-Robot Collaborative Search. *Left*: A volunteer and a robot collaborating in a search task. *Right*: Map of an ongoing search task. Here, the object location probability (violet grid) and the current shared plan (green and red paths) are displayed.

team members, even if not explicitly communicated, provides knowledge on each agent's probable contribution. Using the insight from multi-robot and multi-agent literature, we aim to obtain a team-aware motion shared plan that combines the human exploration ideas with the robot optimal exploration policy. Moreover, the robot will make the shared plan take into account either the human's intention or the human's feedback. This process will be repeated as many times it is required until the goal is reached.

In this work, we have developed a multi-agent shared plan that is based on a decentralized variant of Monte Carlo Tree Search (MCTS) [4] and also uses the Social Reward Sources (SRS) model [8]. The SRS model allows to use sources' target sets, define shared goals and build agents' mental models. Moreover, the team members themselves may be defined as sources of reward to model inter-agent dependencies.

The present work is organized as follows. Section II explains the related work. Section III explains the collaborative search plan and the general method. Section V describes the HRI interface for the human, the experiments and the achieved results. Finally, section VI summarizes the conclusions.

II. RELATED WORK

In this section, the state of the art is presented. The overall goal of this research is to develop robots that work cooperatively in shared plans with people. Roughly speaking, we review human and robot task and goal representations, current approaches to human-robot collaborative navigation settings and several multi-agent and multi-robot works where parallelisms to the human-robot collaboration case can be found.

A. Task Representation in Human-Robot Teams

When facing human-robot joint action, studies on human-human joint action can be inspiring for human-robot team work [12], [7]. According to [24], human interpretation of robot actions in terms of action goals and means to achieve them is similar to how other human's actions are conceived. This, however, does not mean this representation is understandable by robots or, equivalently, that humans may understand the robots' internal knowledge representation.

Shared task representations require fluid human-robot communication, to share knowledge about the current status of the world. This may include gesture recognition [18] and natural language [23].

In [1], authors introduce applications of social eye gaze, [2] explores eye-hand behaviours in human-robot shared manipulation and [19] reviews intent detection, arbitration and communication aspects of shared control for physical human-robot collaboration.

B. Human Robot Collaborative Navigation

Human-robot collaboration studies those systems where humans and robots work together to achieve shared goals, a broad field sustained by many pillars: knowledge representation, planning, communication, plan sharing, decision making, agreement and adaptation. Within this framework, human-robot collaborative navigation (HRCN) focuses on shared navigation tasks, ranging from goal allocation over a number of agents to complex synchronous movements, like dances or acrobatics. Concretely, in this paper, we tackle human-robot collaborative search, we go one step further considering the problem of collaborative exploration making use of a robot and a person.

Some works see the human as a manager [20], [22] Hong et al. [13] refer to it as a “busy genius” and they achieve better performance in multi-robot search applying semi-autonomous teleoperation. The TRADR project [17] aims to enhance human-robot teams collaboration in USAR, highlighting report generation studies and work agreement handling and evaluation [21], though the agreement generation process is yet to be included. Johnson et al. [15] present a new design process for human-robot collaborative applications focusing on identifying joint activity interdependences: the coactive design. Other alternative approaches include co-driving, as the collaborative teleoperation of a robot through dialogue [10] or the collaborative control of wheelchair [6].

Finally, side-by-side navigation [9], [11], [14] is one of the first steps into collaborative models, but they cannot be extended to other applications as they are task-focused.

C. Multi-Agent Path Planning

In multi-agent planning, there is a family of sampling algorithms which are commonly used: the Monte Carlo Tree Search (MCTS). Browne et al. [5] reviews the usage of these methods and a large set of possible variations and enhancements present in the literature. One good recent example of a multi-robot MCTS implementation for path planning is the work of Best et al. [4]. They adapt the MCTS

to an interleaved distributed setting while presenting a new tree expansion policy (discounted UCT). In this work, each robot expands a global tree for all the agents and shares its computed probabilities for a set of dynamically selected available plans over a PRM connected through Dubins paths. Their implementation enables a robust distributed global plan expansion.

III. COLLABORATIVE SEARCH

One recurrent human activity is searching something, an object or a person in a specific environment. People use their environment knowledge, experience and strategy to do the searching task. When the search is done in a team of people, the strategies are different and usually overall efficiency is enhanced via a coordination among the team members.

In this work, we approach the challenge of designing robots capable of collaboratively participate in search tasks while being part of human-robot mixed teams. In general, searching may include a large spectrum of actions such as active perception and object manipulation. In our approach, we consider the search as a navigation task where the environment is assumed to be fixed (no possible physical interaction with the environment for occlusion removal), the robot sensors are in the robot platform and the target object is on ground level. On the other hand, both the robot and the human can move freely, they can communicate through a mobile phone app and both team members may follow any search policy and even lose contact.

A. System Overview

The human-robot collaborative system consist of two agents, the robot and the human. As a first approach, we have considered that the task begins with the robot making a first plan of both itself and the human. Then the robot displays these plans to the human in the HRI interface. The human can decide to follow the robot proposal or, conversely, ask for another one. Additionally, the human may convey to the robot its intentionality and/or the robot goal before asking for a new plan. In any case, they will ask for a replan and the robot will make another proposal plan. The agreement and decision process will finish when human accepts the shared plan. In the following sections we will explain the MCTS planner and detail the HRI interface. This process is repeated until the searched object is found.

B. Problem Statement

The human-robot collaborative search task may be defined as the process through which a given team of agents A , comprised of both humans and robots, explores a known space to locate an object O . During this process, agents can update their belief over the object location through the exploration of the environment, the observation of their colleagues' actions and the received information through active communication. The task is assumed as finished when the object is found.

In the experiments presented in this thesis, the team consists of one person and one robot. We focus on searching

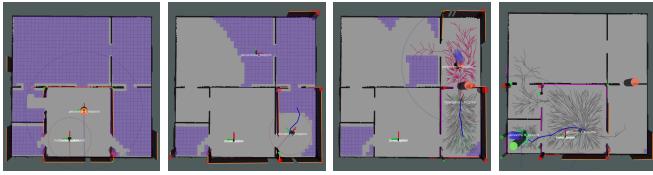


Fig. 2. **Collaborative Search Testbed.** From top to bottom, left to right: a) The robot infers the unexplored zone from its detection range (red circle) and the person's (blue circle). b) People detection is impossible when the person is out of sight, hence no inference is done. c) The person indicates the robot to avoid searching through that zone, as either it is already explored or the person will do it on their own. d) The person finds the object, thus indicates the robot to come.

both human and robot accessible spaces, so both are assumed capable of navigating through it autonomously. To perform this task, we build an observability graph upon a discretised representation of the search space, as shown in Fig. 2.

In the following sections, we formalise the core mechanics of this problem.

1) Agent Detection Model: During the search process, the object location belief is continuously updated based on the actors' actions. Hence, we should model the probability of an agent $a \in A$ detecting an object O at a certain location $\vec{p} = (r_a, \theta_a)$, where (r_a, θ_a) are the polar coordinates of the location p in the human a reference.

$$P(D_a(\vec{p}, \Delta t) | O(\vec{p})) \quad (1)$$

where $\Delta t = t_f - t_0$ is the search time and $D_a(\vec{p}, \Delta t)$ and $O(\vec{p})$ state the object being detected by agent a and actually being at the given location, respectively. We make a number of assumptions:

Assumption 1. Detection models are independent of their initial time t_{0_i} . In other words, human detection capability does not change over time. So for one agent:

$$P(\overline{D}_a(\vec{p}, \Delta t) | O(\vec{p})) = \prod_i P(\overline{D}_a(\vec{p}, \Delta t_i) | O(\vec{p})) \quad (2)$$

where the overline in \overline{D}_a expresses the complementary statement, i.e. being “undetected by agent a ”, and

$$t_{0_{i_0}} = t_0, \quad t_{f_{i_f}} = t_f, \quad t_{f_i} = t_{0_{i+1}} \quad i = i_0, \dots, i_{f-1}$$

Assumption 2. Detection models are independent of $\vec{p}_a \forall a \in A$. Human detection capability is independent on the perceiving human position, as long as (r_a, θ_a) is visible, and on all other participating agents' position. Whereas, change of focus or occlusions in its field of view due to other teammates proximity are not considered. So for each location:

$$P(\overline{D}(\vec{p}, \Delta t) | O(\vec{p})) = \prod_{a=1}^A P(\overline{D}_a(\vec{p}, \Delta t) | O(\vec{p})) \quad (3)$$

2) Object Location Probability: At a given time t , where t_0 is the task beginning and $t_f = t$, the updated object probability on each location given the current accumulated global search is:

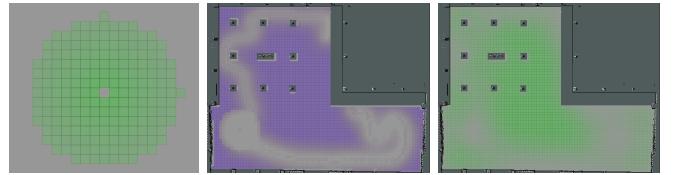


Fig. 3. **Search Social Reward Sources.** From left to right: a) Individual search social reward source proportional to the probability of detecting the object in its center area. b) Visualization of the object location probability in an intermediate timestamp of a two-agent collaborative search. c) Additive visualization of the rewards generated by all the search reward sources in the previous timestamp.

$$P(O(\vec{p}) | \overline{D}_t) = \frac{P(\overline{D}_t | O(\vec{p})) \cdot P(O(\vec{p}))}{P(\overline{D}_t)} \quad (4)$$

To update the global object location belief, we make the following assumption:

Assumption 3. Humans make no false positives while searching or, on another perspective, they filter them automatically. Consequently:

$$P(O(\vec{p}) | \overline{D}_t) = \frac{P(\overline{D}_t(\vec{p}) | O(\vec{p})) \cdot P(O(\vec{p}))}{1 - \sum_{\vec{p}} P(D_t(\vec{p}) | O(\vec{p})) \cdot P(O(\vec{p}))} \quad (5)$$

The previous formula can be further simplified for the uniform prior case. If we are working on a uniform space discretisation, we may use:

$$P(O(\vec{p}) | \overline{D}_t) = \frac{P(\overline{D}_t(\vec{p}) | O(\vec{p}))}{N - \sum_{\vec{p}} P(\overline{D}_t(\vec{p}) | O(\vec{p}))} \quad (6)$$

where $P(\overline{D}_t(\vec{p}) | O(\vec{p}))$ is iteratively updated from observations using the detection model. One may obtain an efficient belief using a dynamic programming approach.

C. Task Modelling through SRS

The SRS modelling [8] of the collaborative search task Ψ_{CS} is quite straightforward. We represented the collaborative search as a number of SRSs equivalent to that of the search space discretisation. In other words, each possible object location generates a function of reward. Each source ψ_L generates a reward proportional to the probability of detecting the object on the source location $P(O(\vec{p}_{\psi_L}) | \overline{D}_t)$ along the search space (Fig. 3.a). All sources can be combined by the planner, for instance in an additive form, while being subject to independent consumable dynamics. This allows for a coherent reward evolution inference on the planner simulations whilst allowing the planner to be disentangled from any task knowledge, such as visualization restriction or detection models.

$$\Psi_{CS} = (\psi_L \mid P(O(\vec{p}_{\psi_L})) > 0) \quad (7)$$

Likewise, belief is updated along the search progress and sources are generated at each replanning phase. Fig. 3.b shows a timestamp on a two-person search process. On the other hand, Fig. 3.c is the visualization of merging all the resulting social reward sources, which is proportional to the probability of seeing an object from the given lookout (discretisation block).

IV. MULTI-AGENT MCTS PLANNER

This multi-agent planner is a decentralized variant of the Monte Carlo Tree Search (MCTS) similar to Best et al. [4], but incorporating robots and humans in the process instead of only robots. The inclusion of humans in the loop has huge implications. The global team objective function has to include both humans and robots, their differences and mutual interaction, into account. Moreover, since the communication between the human and the robot is asynchronous and slow, the model cannot include information exchange iterations in planning time. It often has to update the shared plans using the previous human information or predictions and the interaction information is added in the objective function through the usage of the SRS model.

A. Problem statement

We consider a team of A agents $1, 2 \dots A$, humans or robots, where each agent a plans its own sequence of future actions \mathbf{x}^a in a decentralised manner. Each action x_j^a has an associated cost c_j^a and a required completion time t_j^a . The feasible set of actions and associated costs at each step j are a function of the previous actions $(x_1^a, x_2^a, \dots, x_{j-1}^a)$. Thus, there is a predefined set \mathcal{X}^a of feasible action sequences \mathbf{x}^a for each agent a . We use \mathbf{x} to denote the set of action sequences for all humans and robots $\mathbf{x} := \{\mathbf{x}^1, \mathbf{x}^2, \dots, \mathbf{x}^A\}$ and \mathcal{X} to denote the set of all feasible \mathbf{x} . Likewise, $\mathbf{x}^{(a)}$ and $\mathcal{X}^{(a)}$ define the set of action sequences and feasible action sequences of all agents but a .

The aim is to maximize a global team objective function $g(\mathbf{x})$ that is a function of the action sequences of all agents. We assume each agent a knows the global objective function g , but does not know the action sequences $\mathbf{x}^{(a)}$ selected by the other agents. For most of our proposed approach, we assume g is deterministic given a known set of action sequences \mathbf{x} , though each agent a may only have a partial or inaccurate perception of g , a local objective function $f^a(\mathbf{x})$.

In the presented approach, the objective function g is indirectly build through the Social Reward Sources model [8]. Thus, the planner must be able to deal with an arbitrary dynamic objective function that may change due to the past action sequences $\mathbf{x}_{1..j-1}$.

As we assume some agents in the setting may be humans, the problem must be solved in a decentralized and online setting. Agents may be able to communicate and make shared plans expressed in the objective functions $f^{\{1..A\}}$. The agents being robots and humans, however, are unable to communicate during planning time to improve coordination. Therefore, each robot will plan based on the information it has available locally. Due to this, communication needs may be encoded in $g(\mathbf{x})$, enabling active communication plans.

B. Algorithm Overview

The algorithm runs simultaneously and asynchronously on all robots in the team; we present the algorithm from the perspective of one robot, being the agent 0 of the team. The algorithm cycles between the three phases illustrated in Figure 4: (1) incrementally grow a set of feasible action

sequences for each agent of the team \mathcal{X}^\dashv , (2) individually compute the probability distribution over each agent possible action sequences and (3) incrementally grow a search tree using MCTS while taking into account information about the other agents' objectives and plans.

C. Building the Search Space

In this MCTS implementation, we use the RRT (Rapidly exploring Random Trees) to generate a feasible restricted set of each agent's possible paths $\hat{\mathcal{X}}^a$. These paths can be seen as heterogeneous action sequences, each one leading to a different goal.

In Best et al. [4], they build a multi-agent MCTS planner over a common PRM. Their agents, however, are assumed to have the same mobile and planning capabilities, a premise that does not hold in human-robot settings. Using individual RRTs, each agent tree may be expanded with different restrictions and precision. Moreover, since we do not know the human's plan, we assume that the robot will predict it using also a RRT planner. This path plan together with the robot plan will be communicate through the HRI interface to the human for human's approval or rejection, and in the last case the human will ask for a replanning.

1) Agent Action Set: Each agent action is assumed to be a movement action and represented by a RRT node i . Movement actions are defined by their origin, the node's parent location, their goal, their own location, and the completion time. Additionally, each action node can store a distribution probability over its children electability on a satisfactory shared plan.

Every RRT node i with position \vec{p}_i^a and time t_i^a can have an unbounded number of children $ch(i)$, but the number of actions eligible after each agent action is bounded to N_a . Moreover, the node stores a probability distribution q_i^a over the eligible actions (further developed in section IV-D.2).

D. Building a Human-Robot Team Plan

We aim to build a collaborative navigation plan to tackle shared tasks over heterogeneous action sets with variable time horizons. To do so, the MCTS planner should ensure temporal coherence in the tree expansion, deal with coexistence of agent action sequences with different temporal length and provide a feasible reward propagation mechanism to deal with a dynamic environment.

1) MCTS State: MCTS states are defined over the agents' action RRT trees. Each state s is formed by a list of ongoing agent actions on time t_s , $\mathbf{x}_s = \{x_s^0, x_s^1, \dots, x_s^A\}$ and each one's remaining time to finish $\mathbf{rt}_s = \{rt_s^1, rt_s^2, \dots, rt_s^A\}$, as well as the current collaborative plan time of the state.

Each MCTS state s_k can have a determinate number of children states $ch(s_k)$. The number of children states is bounded by $|ch(s_k)| \leq \prod_a^{(a_i | rt_s^{a_i} = 0)} ch(x_s^a)$. In other words, each MCTS state can only have as many successors as the existent possible combinations of finished tasks' eligible children. Moreover, a new state should be generated

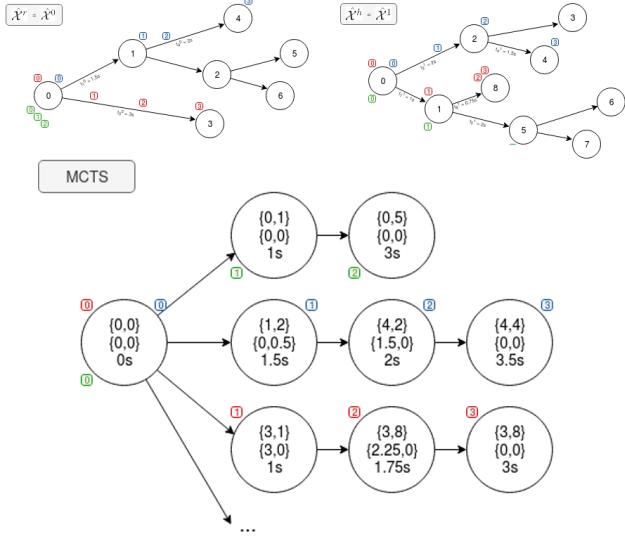


Fig. 4. **MCTS Tree expansion.** Example of a collaborative plan expansion for a two-member team. These corresponds to our tested case, being a robot and a human correspondingly. The green, blue and red markers illustrate three possible different plans and each agent RRT mapping of their MCTS nodes.

after each agent task is finished. More insight about this restrictions can be found in section IV-D.2.

2) MCTS expansion: First, an individual agent reward upper bound is calculated over each agent RRT. Rewards for each agent a action sequences $\mathbf{x}^a \in \mathcal{X}^a$ are updated using the method explained in [8]. Then, all these rewards are back-propagated and each tree node i stores the maximum attainable individual reward from it. For each, only the N_a children nodes with the highest upper bound are considered, further pruning the search tree, and a selection probability q_i^a weighted by those bounds is assigned to them.

The multi-agent plan is expanded from the root node s_0 , which is constituted by the root actions of each agent (0-time actions). In every iteration m , one of the expandable states s_m in the collaborative plan tree is randomly selected. From it, a chain of future states is continuously simulated until a final state is reached. Each new simulated state is sampled from a probability distribution q defined by each individual agent action node probability distribution q^a .

An example of an ongoing collaborative plan expansion of a two member human-robot team may be found in Fig. 4. Such team distribution is equivalent to the one used in the presented experiments.

Additionally, we added some predefined preliminary expansion to the MCTS tree. To ensure the viability of early plans we collect potentially rewarding goals from the individual agents' action sequences and combine them to define end state candidates.

E. Objective Function

We may define the collaborative plan objective function $g(\mathbf{x})$ as the additive combination of all the rewards influencing the team. This includes the rewards related to each agent

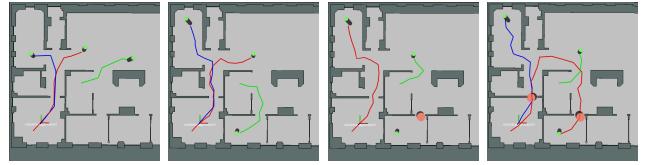


Fig. 5. **Human-robot multi-agent MCTS.** Examples of the presented MCTS method. From left to right: a) Global plan fulfilling three shared goals. b) Maximizing global rewards, even at the expense of increasing individual agent's effort. c) One of the shared goals is inaccessible due to some hazard. d) Two hazards in the scene, but they only affect one of the agents (green). The planner adapts to keep fulfilling the three shared goals.

actions cost and perceived influence of the environment, the shared task and team interaction.

$$g(\mathbf{x}) = \sum_{a=1}^{N_a} (R_A^a + R_E^a) + R_T + R_I \quad (8)$$

being

$$R_A^a(\mathbf{x}^a) = - \sum_{\substack{j \\ x_j^a \in \mathbf{x}^a}} c_j^a \quad (9)$$

$$R_E^a(\mathbf{x}^a) = \sum_{s_k}^{S(\mathbf{x}^a)} r_e(\Psi_e^a, x_{s_k}^a, t_{s_k}, \Delta t_{s_k}) \quad (10)$$

$$R_T(\mathbf{x}) = \sum_{s_k}^{S(\mathbf{x})} r_t(\Psi_t^\tau, \mathbf{x}_{s_0 \rightarrow k}, t_{s_k}, \Delta t_{s_k}) \quad (11)$$

$$R_I(\mathbf{x}) = \sum_{s_k}^{S(\mathbf{x})} \sum_{a=1}^{N_a} r_i(\Psi_{i_a}^\tau, \mathbf{x}_{s_0 \rightarrow k}^{\tau \cup a}, t_{s_k}, \Delta t_{s_k}) \quad (12)$$

where $S(\mathbf{x})$ is the set of MCTS states defined by the action sequences \mathbf{x} . $r_e(\Psi_e, x^a, t, \Delta t)$ is the reward generated by the environmental sources set Ψ_e to target agent a while performing action x^a during a period of Δt initiated at time t . Likewise, r_t and r_i are the rewards generated by sources Ψ_t^τ and $\Psi_{i_a}^\tau$, given the action sequences in $\mathbf{x}_{s_0 \rightarrow k}$ (action sequences that generate the tree branch connecting the initial state s_0 and the state s_k). Finally,

$$rt_s^a = \begin{cases} rt_s^a - \Delta t_s & \text{if } x_s^a = x_{p(s)}^a \\ t_s^a - \Delta t_s & \text{otherwise} \end{cases} \quad (13)$$

$$\Delta t_s = t_s - t_{p(s)} \quad (14)$$

where $p(s)$ is the parent state of s .

The images in Fig. 5 are presented to provide some qualitative example of the model. In them, a human-robot team of three members is given three shared goals and we are shown the plan built by one of the agents using the multi-agent MCTS model over the SRS representation of four different environments.

V. EXPERIMENTS

We validate our model using the BRL map from the *Barcelona Robot Lab Dataset*¹. The explorable area is discretised and all obstacles in the scene are assumed to block

¹<http://www.iri.upc.edu/research/webprojects/pau/datasets/BRL/>



Fig. 6. **Experimentation.** Human-robot pair collaboratively searching for a green Parcheesi token.

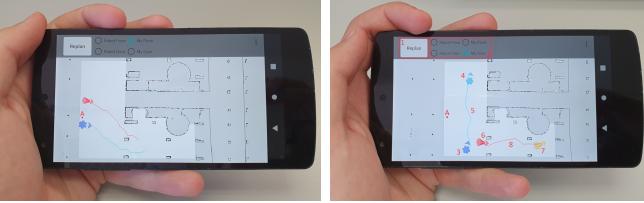


Fig. 7. **Mobile App's main screen.** *Left:* Execution of the system using only human's current position, the robot makes both plans autonomously considering *A* as the starting point of the experiment. *Right:* Execution using all system functionalities, also *A* as the starting point for the experiment. 1) Replan button. 2) Input data selection menu. 3) Robot's current position. 4) Human's desired goal for the robot. 5) Path calculated for the robot. 6) Human's current position. 7) Human's intended goal. 8) Path calculated for the human.

both the view of the robot and the human. We tested the model in a two-agent environment, where human participants collaborate with the robot in the a search task of a green Parcheesi token (Fig. 6).

A. HRI Interface and Experiments Setup

In order to allow both agents to communicate relevant information to each other as mentioned in section IV-A, we designed a mobile application based on the previous work from Kohler et al. [16]. The resulting app looks like Fig. 7.

Normal operation is as follows: the user indicates their position and, optionally, other relevant data such as their intention or the one they wish for the robot, and presses the Replan button. The robot calculates both plans and, if the user dislikes the result, this can be rejected asking for a new calculation by pressing the Replan button again. Otherwise, the robot will interpret that the user has accepted the plans made by the robot and will start moving. Using these replan policy the team can reach an agreement.

This interaction allows the appearance of different types of relationships: master-slave with the robot acting as master if the human only indicates their position and follows the path calculated by the robot, master-slave with the human as master if the human indicates the desired goal for the robot or a peer-to-peer relation if the human indicates their intention and leaves the robot to plan its path on its own.

Regarding the experiments setup, we selected an outdoor setting with two areas: one open space where it is easy to maintain direct line of sight between agents and another covered with multiple occlusions due to walls and columns making it necessary to maintain communication through the previous application. The robot is based on a Pioneer 2 mobile platform, the application is executed on a Nexus 5

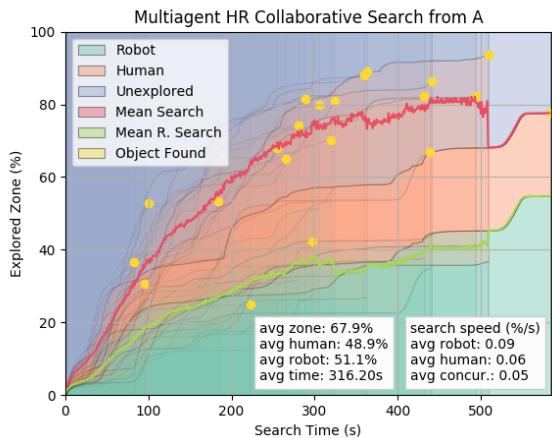


Fig. 8. **Multiagent Human-Robot Collaborative Search.** Experimental data.

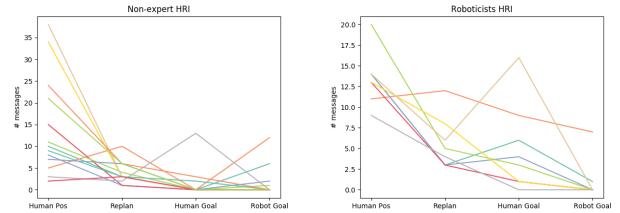


Fig. 9. **Interface Communication.** Human to robot message count per episode. From left to right: a) Results from people without contrastable knowledge of robotics. b) Results from people with studies or jobs related with robotics.

mobile with Android 10 and communications are carried out through the deployment of a local Wi-Fi network.

B. Results

A total of 19 volunteers, 4 women and 15 men, participated in the experiment. They were between 18 and 40 years old (mean: 21.95 std: 5.39) and, on a scale of 1 (None) to 5 (Expert), their average self-evaluated knowledge in robotics was 2.37 (std: 1.21). No one could practice using the setting or the mobile app, neither had any previous experience using it. Each of them participated in one or two episodes and, additionally, participants were surveyed after the interaction.

In Fig. 8, we can observe a graph depicting 21 search episodes. Each episode is characterized by two stripes, a green one representing the search area covered by the robot and an orange depicting the perceived contribution of the human. The light green and red lines represent the average value of each of them over all the episodes. Each episode ends when the object is found by any of the team members, such event being depicted with a yellow circle.

The task was finished on all episodes, with a mean time to completion of 316.20s and covering an average explored zone of 67.9% of the map. In the robot's perception, both the human and the robot explored more or less equivalent areas, a value that might be underestimated in the human case due to the asynchronous update of their location in occlusion situations.

Robot		mean	std	Mobile App		mean	std
Likeability	Dislike (1) - Like (5)	4.2	0.9	Likeability	Dislike (1) - Like (5)	3.9	0.9
	Unfriendly (1) - Friendly (5)	4.1	1.2		Unfriendly (1) - Friendly (5)	3.8	0.9
	Unpleasant (1) - Pleasant (5)	3.8	1.1		Unpleasant (1) - Pleasant (5)	3.7	1.0
	Awful (1) - Nice (5)	4.1	0.7		Awful (1) - Nice (5)	3.7	1.2
Robot Plans	Incompetent (1) - Competent (5)	4.0	0.9	Usefulness [from useless (1) to essential (5)]	Knowing the robot plan	4.2	0.9
	Ignorant (1) - Knowledgeable (5)	3.8	0.9		Asking the robot for another plan	4.3	0.7
	Irresponsible (1) - Responsible (5)	4.1	0.8		Telling the robot your intention	3.6	1.3
	Unintelligent (1) - Intelligent (5)	3.9	1.0		Telling the robot where to go	4.2	0.9
	Foolish (1) - Sensible (5)	3.7	0.9		Telling the robot your location	3.4	1.2
Perceived Safety	Anxious (1) - Relaxed (5)	4.1	1.3	General	Is our app helping you achieve your goals?	3.7	1.0
	Agitated (1) - Calm (5)	4.1	1.3				
	Quiescent (1) - Surprised (5)	3.3	1.4				

Fig. 10. **HRI Acceptability Study.** Participants' feedback concerning the robot and the mobile interface. Some of the questions are inspired from the *Godspeed* questionnaire [3].

In Fig. 9, it can be observed the number of human to robot interactions through the mobile on each episode. Here, we may differentiate between non-expert users (13) and people with technical or academic knowledge of robotics (6). Non-experts users indicate their position to the robot with great variability from almost never to assiduously while experts do make regular use of this functionality due to their intuition of the basic needs of the robot. Regarding the intention of the person transmitted to the robot through Human Goal messages, non-experts users tend not to make use of this functionality and either accept the plan calculated by the robot and establish a master-slave relationship with the robot as master or pulse again the Replan button looking for a route closer to their original intention. Expert users, though, do use this functionality showing a tendency towards peer-to-peer or master-slave relationships being these users the master.

Finally, in Fig. 10 are depicted the obtained results from participants' feedback. The results seem to indicate that participants appreciate that the robot fulfills its purpose in a competent way, but that there are doubts about how it does it. About the mobile app, the most valued functionality is to request another plan, probably because it is necessary for any type of relationship established with the robot. In any case, there is consensus that the application helps to solve the task.

VI. CONCLUSIONS

In this paper, we approach the usage of multi-agent motion planners in human-robot team settings. We developed a new human-robot MCTS multi-agent planner for collaborative navigation plans, including human robot interaction through its objective function. Human's intentionality and decision agreement are included in the system through the objective function and the replanning policy.

We tested the model using the collaborative search testbed in real world and evaluated human-robot interaction by both expert and non-expert participants. The experiments were conducted in the Barcelona Robotic Lab and human-robot communication was enabled through the development of a mobile phone app. All experiments conducted reached the

final objective of the search, finding a Parcheesi token, and we collected subjective perception of the process from the participants. The results concerning the acceptability of both the planner and the interface are promising and suggest further research.

REFERENCES

- [1] Admoni, H., Scassellati, B.: Social eye gaze in human-robot interaction: a review. *Journal of Human-Robot Interaction* 6(1), 25–63 (2017)
- [2] Aronson, R.M., Santini, T., Kübler, T.C., Kasneci, E., Srinivasa, S., Admoni, H.: Eye-hand behavior in human-robot shared manipulation. In: Proceedings of the ACM/IEEE International Conference on Human-Robot Interaction. pp. 4–13 (2018)
- [3] Bartneck, C., Kulic, D., Croft, E., Zoghbi, S.: Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots. *International journal of social robotics* 1(1), 71–81 (2009)
- [4] Best, G., Cliff, O.M., Patten, T., Mettu, R.R., Fitch, R.: Dec-mcts: Decentralized planning for multi-robot active perception. *The International Journal of Robotics Research* 38(2-3), 316–337 (2019)
- [5] Browne, C.B., Powley, E., Whitehouse, D., Lucas, S.M., Cowling, P.I., Rohlfshagen, P., Tavener, S., Perez, D., Samothrakis, S., Colton, S.: A survey of monte carlo tree search methods. *IEEE Transactions on Computational Intelligence and AI in games* 4(1), 1–43 (2012)
- [6] Carlson, T., Demiris, Y.: Collaborative control for a robotic wheelchair: evaluation of performance, attention, and workload. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)* 42(3), 876–888 (2012)
- [7] Clodic, A., Pacherie, E., Alami, R., Chatila, R.: Key elements for human-robot joint action. In: *Sociality and Normativity for Robots*, pp. 159–177. Springer (2017)
- [8] Dalmasso, M., Garrell, A., Jimenez, P., Sanfeliu, A.: Human-robot collaborative navigation search using social reward sources. arXiv preprint arXiv:1909.04768 (2019)
- [9] Ferrer, G., Garrell, A., Herrero, F., Sanfeliu, A.: Robot social-aware navigation framework to accompany people walking side-by-side. *Autonomous robots* 41(4), 775–793 (2017)
- [10] Fong, T., Thorpe, C., Baur, C.: Collaboration, dialogue, human-robot interaction. In: *Robotics Research*, pp. 255–266. Springer (2003)
- [11] Garrell, A., Garza-Elizondo, L., Villamizar, M., Herrero, F., Sanfeliu, A.: Aerial social force model: A new framework to accompany people using autonomous flying robots. In: *IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS, Vancouver, BC, Canada, September 24–28, 2017*. pp. 7011–7017 (2017)
- [12] Hoffman, G., Breazeal, C.: Collaboration in human-robot teams. In: *AIAA 1st Intelligent Systems Technical Conference*. p. 6434 (2004)
- [13] Hong, A., Igharoro, O., Liu, Y., Niroui, F., Nejat, G., Benhabib, B.: Investigating human-robot teams for learning-based semi-autonomous control in urban search and rescue environments. *Journal of Intelligent & Robotic Systems* 94(3-4), 669–686 (2019)

- [14] Jayawardena, C., Ardekani, I., et al.: A navigation model for side-by-side robotic wheelchairs for optimizing social comfort in crossing situations. *Robotics and Autonomous Systems* 100, 27–40 (2018)
- [15] Johnson, M., Bradshaw, J.M., Feltovich, P.J., Jonker, C.M., Van Riemsdijk, M.B., Sierhuis, M.: Coactive design: Designing support for interdependence in joint activity. *Journal of Human-Robot Interaction* 3(1), 43–69 (2014)
- [16] Kohler, D., Conley, K.: Rosjava - an implementation of ros in pure java with android support (2011), https://github.com/rosjava/rosjava_core
- [17] Kruijff-Korbatová, I., Colas, F., Gianni, M., Pirri, F., de Greeff, J., Hindriks, K., Neerincx, M., Ögren, P., Svoboda, T., Worst, R.: Tradr project: Long-term human-robot teaming for robot assisted disaster response. *KI-Künstliche Intelligenz* 29(2), 193–201 (2015)
- [18] Liu, H., Wang, L.: Gesture recognition for human-robot collaboration: A review. *International Journal of Industrial Ergonomics* 68, 355–367 (2018)
- [19] Losey, D.P., McDonald, C.G., Battaglia, E., O’Malley, M.K.: A review of intent detection, arbitration, and communication aspects of shared control for physical human-robot interaction. *Applied Mechanics Reviews* 70(1), 010804 (2018)
- [20] Marconi, L., Melchiorri, C., Beetz, M., Pangercic, D., Siegwart, R., Leutenegger, S., Carloni, R., Stramigioli, S., Bruyninckx, H., Doherty, P., et al.: The sherpa project: Smart collaboration between humans and ground-aerial robots for improving rescuing activities in alpine environments. In: *International Symposium on Safety, Security, and Rescue Robotics (SSRR)*. pp. 1–4. IEEE (2012)
- [21] Mioch, T., Peeters, M.M., Neerincx, M.A.: Improving adaptive human-robot cooperation through work agreements. In: *27th International Symposium on Robot and Human Interactive Communication (RO-MAN)*. pp. 1105–1110. IEEE (2018)
- [22] Silvagni, M., Tonoli, A., Zenerino, E., Chiaberge, M.: Multipurpose uav for search and rescue operations in mountain avalanche events. *Geomatics, Natural Hazards and Risk* 8(1), 18–33 (2017)
- [23] Thomason, J., Zhang, S., Mooney, R.J., Stone, P.: Learning to interpret natural language commands through human-robot dialog. In: *Twenty-Fourth International Joint Conference on Artificial Intelligence* (2015)
- [24] Wykowska, A., Chellali, R., Al-Amin, M.M., Müller, H.J.: Implications of robot actions for human perception. how do we represent actions of the observed robots? *International Journal of Social Robotics* 6(3), 357–366 (2014)