

# Planning for Serendipity

Tathagata Chakraborti<sup>1</sup> Gordon Briggs<sup>2</sup> Kartik Talamadupula<sup>3</sup> Yu Zhang<sup>1</sup>  
Matthias Scheutz<sup>2</sup> David Smith<sup>4</sup> Subbarao Kambhampati<sup>1</sup>

**Abstract**—Recently there has been a lot of focus on human robot co-habitation issues that are often orthogonal to many aspects of human-robot teaming; e.g. on producing socially acceptable behaviors of robots and de-conflicting plans of robots and humans in shared environments. However, an interesting offshoot of these settings that has largely been overlooked is the problem of planning for serendipity - i.e. planning for stigmergic collaboration without explicit commitments on agents in co-habitation. In this paper we formalize this notion of planning for serendipity for the first time, and provide an Integer Programming based solution for this problem. Further, we illustrate the different modes of this planning technique on a typical Urban Search and Rescue scenario and show a real-life implementation of the ideas on the Nao Robot interacting with a human colleague.

## I. INTRODUCTION

Automated planners are increasingly being used to endow robots with autonomous planning capabilities in joint human-robot task scenarios [1]. As the efficiency and ubiquity of planners used in these scenarios increases, so does the complexity of the various tasks that the planner can handle on behalf of the robot. Specifically, in cooperative scenarios (including human-robot teaming), the planner’s role is no longer limited to only generating new plans for the robot to execute. Instead, contingent on the availability of the right information, the planner can anticipate, recognize, and further predict the future plans of other agents. Recent work [2] has seen the successful deployment of this idea in scenarios where a robotic agent is trying to coordinate its plan with that of a human, and where the agents are competing for the same resource(s) and must have their plans de-conflicted in some principled manner. Indeed there has been a lot of work under the umbrella of “human-aware” planning, both in the context of path planning [3], [4] and in task planning [5], [6], that aim to provide social skills to robots so as to make them produce plans conforming to desired behaviors when humans and robots operate in

shared settings. However, as we will show in this paper, the robots can be more proactive in their choices to help, and there can be different modes of collaboration in such settings that are not predefined behavioral traits of the robots. Indeed very little attention has been paid to an important phenomenon that often occurs in the course of cooperative behavior amongst agents – *serendipity*.

In this context, serendipity can be seen as the occurrence or resolution of facts in the world such that the future plan of an agent is rendered easier in some measurable sense. Note that there is no explicit team being formed here, and as such the agents do not have any commitments to help each other - this type of assistance can thus be seen as an instance of stigmergic collaboration between robots and humans in co-habitation and a way for robots (to the extent that they only exist in the setting as assistive agents to the humans) to exhibit goodwill to their human “colleagues”. If the planner knows enough about the model, intentions, and state of the other agent in the scenario, it can try to *manufacture* these serendipitous circumstances. To the other agent, conditions that appear serendipitously will look remarkably similar to (positive) exogenous events [7], and that agent may replan to take these serendipitous facts into account, thus hopefully reducing the cost of its own plan.

In this paper, we define for the first time, the notion of planning for serendipity, and outline a general framework for modeling the different modes of such behavior. We will be using a typical USAR (Urban Search and Rescue) setting as the motivating scenario throughout the discussion to illustrate most of these ideas. The rest of the paper is organized as follows. After a brief discussion of the related work, we provide an overview of the current use case, and formalization of the agent models. Then we discuss the issues involved in planning for serendipity, and provide an IP-based planner that models these constraints. Finally we look at the how the planner responds to different situations and provide a demonstration of the ideas on the Nao robot.

## Related Work

Human-aware planning in fact holds a unique position within the multi-agent planning paradigm, as illustrated in Figure 1. When we move from classical planning over to multi-agent planning, we have to deal with challenges in coordination, capability and commitment modeling, handling concurrency, etc. But even within the confines of multi-agent planning, the presence of a human-in-the-loop or aspects of teaming introduces its own typical challenges. For example, introducing a human in the context of teaming behavior

\*This research is supported in part by the ARO grant W911NF-13-1-0023, and the ONR grants N00014-13-1-0176, N00014-13-1-0519 and N00014-15-1-2027.

<sup>1</sup>Tathagata Chakraborti, Yu Zhang and Subbarao Kambhampati are with the Department of Computer Science, Arizona State University, Tempe, AZ 85281, USA {tchakra2, yzhan442, rao}@asu.edu

<sup>2</sup>Gordon Briggs and Matthias Scheutz are with the HRI Laboratory, Tufts University, Medford, MA 02155, USA {gbriggs, mscheutz}@cs.tufts.edu

<sup>3</sup>Kartik Talamadupula is with the Cognitive Learning Department at IBM Thomas J. Watson Research Center, Yorktown Heights, NY, USA krtalamad@us.ibm.com

<sup>4</sup>David Smith is with the Autonomous Systems & Robotics Intelligent Systems Division, NASA Ames Research Center, M/S 269-2, Moffett Field, CA 94035, NY, USA david.smith@nasa.gov

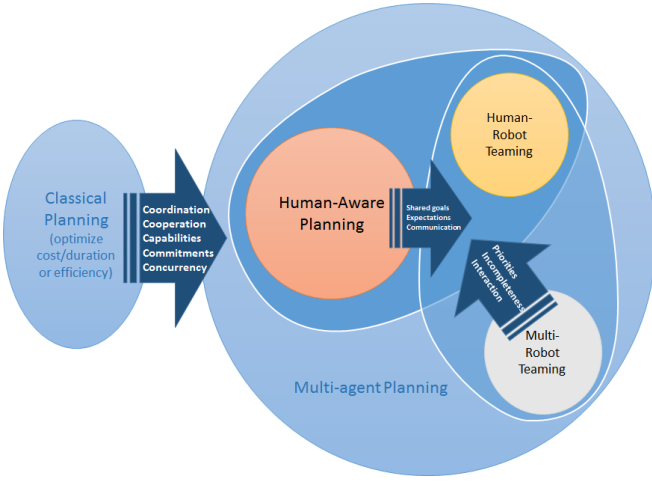


Fig. 1. Challenges involved in the different flavors of multi-agent planning.

with robots often implies having to address issues with model incompleteness, priorities and interaction. Further, the presence or absence of a team itself determines if we can or cannot make such assumptions as shared goals and expectations and communication protocols. Thus, we can think of human-aware planning as a sub-category of multi-agent planning which includes the flavors associated with human-robot interactions, but mostly excludes the assumptions often made in explicit teaming scenarios. However, much of the previous work in such settings, as discussed previously, have been primarily aligned with the notion of indirect coordination - e.g., avoidance of conflicts or producing plans conforming to human expectations or intentions.

## II. OVERVIEW AND PRELIMINARIES

Figure 2 shows a typical USAR setting, unfolding inside a building with interconnected rooms and hallways, with a human commander  $\text{CommX}$  and a robot. The commander has capabilities to move and conduct triage at specified locations, and he can also meet with other agents, as well as pickup, drop off or handover medkits to accomplish their task. The robot can similarly move about, search rooms, or handover or change the position of the medkits. It can thus have its own goals (maybe from being directly assigned by the commander himself or due to long term task specifications), but can also help the commander in accomplishing his goals by fetching the medkits for him. All of these agents are autonomous agents working together or independently in the same environment. The specific problem we look at in this work is an interesting spin-off of such a setting - can the robot choose to help without being told to do so explicitly? What forms of assistance can this involve? Before we discuss ways to model these behaviors, we will first define a few terms related to planning in this setting.

### A. Agent Models

Each agent  $\alpha$  (in the current scenario  $\alpha$  is either  $H$  or  $R$  referring to the human or the robot respectively) is described by a domain model  $D_\alpha = \langle T_\alpha, V_\alpha, S_\alpha, A_\alpha \rangle$ , where  $T_\alpha$  is a set

of object types;  $V_\alpha$  is a set of variables that describe objects that belong to  $T_\alpha$ ;  $S_\alpha$  is a set of named first-order logical predicates over the variables  $V_\alpha$  that describe the world state  $\mathbb{W}$ ; and  $A_\alpha$  is a set of operators available to the agent. The action models  $a \in A_\alpha$  are represented as  $a = \langle \mathbb{N}_a, \mathbb{C}_a, \mathbb{P}_a, \mathbb{E}_a \rangle$  where  $\mathbb{N}_a$  denotes the name of that action;  $\mathbb{C}_a$  is the cost of that action;  $\mathbb{P}_a$  is the list of pre-conditions that must hold for the action  $a$  to be applicable; and  $\mathbb{E}_a = \{eff^+(a), eff^-(a)\}$  is a list of predicates in  $S_\alpha$  that indicates the effects of applying the action. The transition function  $\delta(\cdot)$  determines the next state after the application of action  $a$  in state  $s$  as  $\delta(a, s) = (s \setminus eff^-(a)) \cup eff^+(a), s \subseteq S_\alpha$ . When a sequence of actions is applied to the world state, the transition function determines the resultant state by applying the actions one at a time as follows  $\delta(\langle a_1, a_2, \dots, a_n \rangle, s) = \delta(\langle a_2, \dots, a_n \rangle, \delta(a_1, s))$ .

Further, the robot can also maintain belief models  $Bel_\alpha$  about the other agents in its environment as described in more detail in [2]. For the purposes of this paper, we assume that agents have complete information about the environment, and the robot has the complete domain as well as belief model of the humans. We will also assume that agents do not have communication or observation actions between themselves and the only way they can update their beliefs is through direct interaction (e.g. handing over a medkit) or exceptions in the world state during plan execution. We will relax this assumption later.

### B. Semantics of Individual vs Composite Planning

The agents can of course, given their current state and goal, produce plans based on their own action models. However, given that the robot and the commanders are co-existing (even though independently) in the same environment with potentially mutually helpful capabilities, they could form teams or coalitions or even come up with impromptu interactions consisting of one or more agents in order to achieve a common goal. We will now talk about a formalism to reason with plans in such a setting.

*Definition 1.0* : An individual plan  $\pi_\alpha$  of an agent  $\alpha$  with the domain model  $D_\alpha$  is a mapping  $\mathbb{I}_\alpha \times \mathbb{G}_\alpha \times D_\alpha \mapsto \pi_\alpha$  from the initial state  $\mathbb{I}_\alpha \subseteq S_\alpha$  and the goal state  $\mathbb{G}_\alpha \subseteq S_\alpha$  to an ordered sequence of actions  $\pi_\alpha = \langle a_1, a_2, \dots, a_n \rangle, a_i \in A_\alpha$  such that  $\delta(\mathbb{I}_\alpha, \pi_\alpha) \models \mathbb{G}_\alpha$ . The plan is *optimal* if whenever  $\delta(\mathbb{I}_\alpha, \pi'_\alpha) \models \mathbb{G}_\alpha, C(\pi_\alpha) \leq C(\pi'_\alpha)$  (where  $C(\pi_\alpha) = \sum_{a \in \pi_\alpha} \mathbb{C}_a$  is the cost of the plan).

*Definition 1.1* : A composite plan  $\pi_\mathbb{A}$  of a set of agents  $\mathbb{A} = \{R, H\}$ , referred to as the *super-agent* with a composite domain  $D_\mathbb{A} = \bigcup_{\alpha \in \mathbb{A}} D_\alpha$ , is defined as a mapping  $\mathbb{I}_\mathbb{A} \times \mathbb{G}_\mathbb{A} \times D_\mathbb{A} \mapsto \pi_\mathbb{A}$  from the initial state  $\mathbb{I}_\mathbb{A} = \bigcup_{\alpha \in \mathbb{A}} \mathbb{I}_\alpha$  and the goal state  $\mathbb{G}_\mathbb{A} = \bigcup_{\alpha \in \mathbb{A}} \mathbb{G}_\alpha$  of the super-agent to an ordered sequence of action sets  $\pi_\mathbb{A} = \langle \mu_1, \mu_2, \dots, \mu_n \rangle$ , where  $\mu = \{a_1, \dots, a_{|\mathbb{A}|}\}, \mu(\alpha) = a \in A_\alpha \forall \mu \in \pi_\mathbb{A}$  such that  $\delta'(\mathbb{I}_\mathbb{A}, \pi_\mathbb{A}) \models \mathbb{G}_\mathbb{A}$ , where the modified transition function  $\delta'(\mu, s) = (s \setminus \bigcup_{a \in \mu} eff^-(a)) \cup \bigcup_{a \in \mu} eff^+(a)$ . Similarly, the composite plan  $\pi_\mathbb{A}$  is said to be optimal if whenever  $\delta'(\mathbb{I}_\mathbb{A}, \pi'_\mathbb{A}) \models \mathbb{G}_\mathbb{A}, C(\pi_\mathbb{A}) \leq C(\pi'_\mathbb{A})$  (where  $C(\pi_\mathbb{A}) = \sum_{\mu \in \pi_\mathbb{A}} \sum_{a \in \mu} \mathbb{C}_a$  is the cost of the plan).

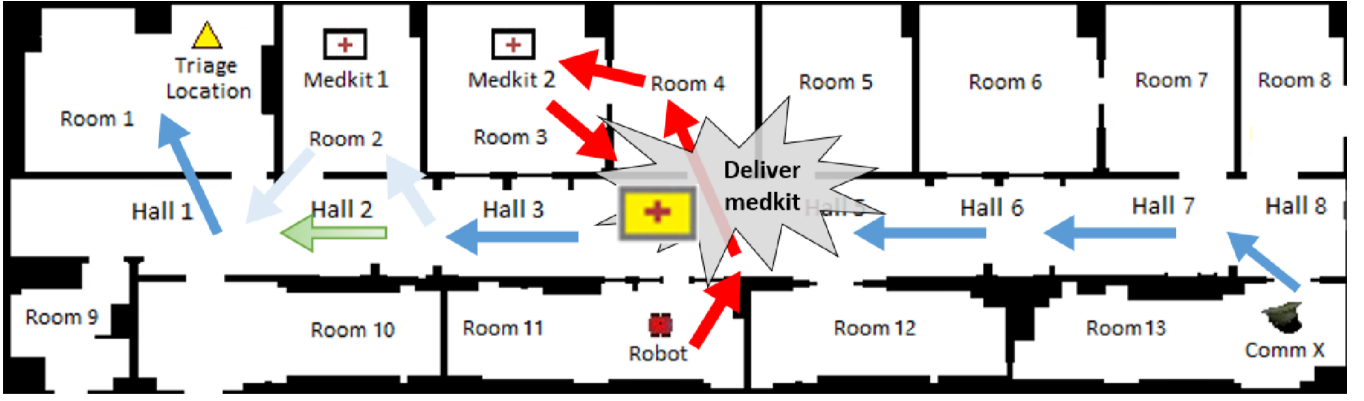


Fig. 2. USAR setting involving a human commander and a robot. The commander has a goal to conduct triage in `room1`, the robot can help him by intercepting him with a medkit in the hallway so he need not fetch one himself - planning for serendipity!

A composite plan can thus be viewed as a composition of individual plans such that they together achieve a particular goal. The characteristics of the composite plan in terms of how these individual plans are composed is determined by what kind of behavior (planning for teaming or collaboration vs serendipity) we desire from the agents involved in the composite plan. Note that the contribution of the human to the composite plan is not the same as the plan he is currently executing - while generating the composite plan, the robot only ensures that this is the plan that ends up being executed (subject to different constraints discussed in detail in the Sections III-A and III-B), given the human’s individual plan currently “planned” for execution.

*Lemma 1.1a* : A composite plan  $\pi_{\mathbb{A}} = \langle \mu_1, \mu_2, \dots, \mu_T \rangle = \bigcup_{\alpha \in \mathbb{A}} \pi_{\alpha}$  can thus be represented as a union of plans  $\pi_{\alpha}$  contributed by each agent  $\alpha \in \mathbb{A}$  so that we can represent  $\alpha$ ’s component as  $\pi_{\mathbb{A}}(\alpha) = \langle a_1, a_2, \dots, a_n \rangle$ ,  $a_i = \mu_i(\alpha) \forall \mu_i \in \pi_{\mathbb{A}}$ .

*Lemma 1.1b* : A composite plan  $\pi_{\mathbb{A}} = \langle \mu_1, \mu_2, \dots, \mu_T \rangle$  with  $\delta'(\bigcup_{\alpha \in \mathbb{A}} \mathbb{I}_{\alpha}, \pi_{\mathbb{A}}) \models \mathbb{G}_{\mathbb{A}}$  guarantees that the world state  $\mathbb{W} \models \mathbb{G}$  at  $t = T$ , if every agent  $\alpha \in \mathbb{A}$  starting from the initial state  $\mathbb{I}_{\alpha}$  at  $t = 0$ , executes  $a_t = \mu_t(\alpha)$ ,  $\forall \mu_t \in \pi_{\mathbb{A}}$  at each time step  $t \in [1, T]$ . It follows that at  $\forall t < T$ ,  $[\pi_{\alpha}]_{\text{execution}}$  is not necessarily same as  $\langle a_1, a_2, \dots, a_T \rangle$ ,  $a_i = \mu_i(\alpha)$ ,  $\mu_i \in \pi_{\mathbb{A}}$ .

The challenge then, for coming up with any combined effort (*planning for serendipity* in the rest of the discussion), is to find the right composition of the individual plans into the composite plan, under constraints defined by the context. We will continue using this notion of a composite plan for the super-agent in the discussion on planning for serendipity in the next section. Here, the robot uses this formalism to come up with composite plans that can benefit its human colleague without their a prior knowledge.

### III. PLANNING FOR SERENDIPITY

In the current section we will look to formalize exactly what it means to be planning for serendipity. If the robotic agents are the ones who bring about the serendipitous moments for the human, then these moments would essentially appear as outcomes of positive exogenous events during the

execution of the human’s plan. Remember that, even though the humans and the robotic agents are cohabitants of the same environment, this is not a team setting, and there is no explicit commitment to help from the robots - and so the human cannot expect or plan to exploit these exogenous events. This means that, given that there are no guarantees or even expectations from the other agents, the human agent can at best only be optimal by himself. This also means that the robots, if they want to make positive interventions, must produce composite plans that are valid given the current human plan under execution. Thus it becomes incumbent on the robot to analyze the original human plan in order to determine which specific parts of the plan can be changed and which parts need to be preserved. Indeed, we will see that these notions of plan interruptibility and plan preservation are crucial to the aspect of planning for serendipity. In the following discussion we will define the semantics of planning for serendipity in terms of plan interruptibility and plan preservation. Before we do that, however, it is worth noting at this point, that the notion of plans being enabled by positive external events is closely associated with the use of triangle table kernels [8] during plan execution. However, triangle table kernels only enable positive effects internal to a plan, and cannot capture the variety of modalities in stigmergic collaboration, specifically ones that involve changes outside of the original plan under execution.

#### A. Plan Interruptibility

We start off by noting that it only makes sense to produce composite plans that have a lesser global cost than the single optimal plan of the human. However, just having a better cost does not guarantee a useful composite plan in the current context. Consider the following example. Suppose the initial positions of `medkit1` and `medkit2` are `room7` and `room3` respectively (refer to Figure 2), and `CommX` has a goal to conduct triage in `room1`. Also, suppose that the robot knows that the commander plans to pick up `medkit1` from `room7` on his way to the triage location (this being the optimal plan), while a cheaper composite plan is available if the robot chooses to pick up `medkit2` from `room3` and hands it over to `commX` in

hall4 which falls in his path. One possible way to make this happen would be to maybe lock the door to room7 so that commX cannot execute his original plan any more, and switches to a plan that happens to conform to the composite optimum. However, since there is no active collaboration between the agents, in this case CommX might very well go looking for the keys to enter room7, and the serendipity is lost. Indeed, this leads us to the notion of identifying parts of the human plan as interruptible, so as to lend itself to such serendipitous execution, as follows -

*Definition 2.0* : If plan  $\pi_H = \langle a_1, a_2, \dots, a_T \rangle$  of the human  $H$  with  $\delta(\mathbb{I}_H, \pi_H) \models \mathbb{G}_H$ , then any subplan  $\pi_H^{ij} = \langle a_i, \dots, a_j \rangle, 1 \leq i < j \leq |\pi_H|$  is *positively removable* iff  $\exists \pi_{\mathbb{A}}$  for the set of agents  $\mathbb{A} = \{R, H\}$  (R being the robot) such that  $\delta'(\bigcup_{\alpha \in \mathbb{A}} \mathbb{I}_{\alpha}, \pi_{\mathbb{A}}) \models \mathbb{G}_H$  where, for some  $i' > i$ ,  $\pi_{\mathbb{A}}(H) = (\subseteq \pi_H[1 : i - 1]) \cdot \pi_{\mathbb{A}}(H)[i : i'] \cdot (\subseteq \pi_H[j + 1 : |\pi_H|])$  and  $C(\pi_{\mathbb{A}}(H)) < C(\pi_H)$  (here  $\cdot$  means concatenation).

*Definition 2.1* : A plan is *interruptible* iff it has at least one positively removable subplan.

Thus time steps  $i \leq t \leq i'$  is when the (serendipitous) exceptions can occur. Note that we specify the rest of the plan to be subsequences of the original plan which ensures that the human does not need to go outside his original plan sans the part where the actual exceptions occurs.

The notion of serendipitous exceptions is closely tied to the issue of what is actually visible to the human and whether such exceptions are immediately recognizable to the human or not. While this is hard to generalize in such non-proximal settings, one measure of this might be the length of the exception. Going back to the previous example, if the exception is just finding the locked door, then this cannot be a positive interruption because when the human goes looking for the keys then this detour is not a subplan of his original plan anymore. However, the exception can always be made to be long enough to accommodate the entire detour, but such exceptions are penalized because it is likely to be harder for the human to come up with such newer plans, partly because the entire world might not be visible to him. Note that this might mean that the formulation would sometimes prefer that the robot does not do the entire job for the human even if it were possible - this is particularly relevant to situations when the human has implicit preferences or commitments in his plan and thus shorter detours are preferable. The exact trade-off between longer interruptions (and possible interference being perceived by the human) and cheaper plans is determined by the objective function of the planner. We intend to do HRI studies similar to [9] to see what kind of exceptions humans really respond to. If we assume that the human replans optimally (and independently) after the serendipitous exceptions, we can modify Definition 2.0 to accommodate such adaptive behavior as follows -

*Definition 2.0a* : If plan  $\pi_H = \langle a_1, a_2, \dots, a_T \rangle$  of the human  $H$  with  $\delta(\mathbb{I}_H, \pi_H) \models \mathbb{G}_H$ , then any subplan  $\pi_H^{ij} = \langle a_i, \dots, a_j \rangle, 1 \leq i < j \leq |\pi_H|$  is *positively removable*

iff  $\exists \pi_{\mathbb{A}}$  for the set of agents  $\mathbb{A} = \{R, H\}$  such that  $\pi_{\mathbb{A}}(H)[1 : i - 1] \subseteq \pi_H[1 : i - 1] = \langle a_1, a_2, \dots, a_{i-1} \rangle$  and  $\pi_{\mathbb{A}}(H)[i' + 1 : |\pi_{\mathbb{A}}(H)|]$  is the optimal plan such that  $\delta'(\bigcup_{\alpha \in \mathbb{A}} \mathbb{I}_{\alpha}, \pi_{\mathbb{A}}) \models \mathbb{G}_H$  and  $C(\pi_{\mathbb{A}}(H)) < C(\pi_H)$ .

We will now see what kinds of positive interruptibility accommodates serendipity for the human, and define constraints on top of Definitions 2.0 and 2.1 to determine opportunities to plan for such serendipitous moments.

### B. Preservation Constraints

Let us now go back to the setting in Figure 2. Suppose the initial position of medkit1 is now room5, and CommX still has a goal to conduct triage in room1. Clearly, one of the optimal human plans is to pick up medkit1 from room5 on his way to the triage location, while a cheaper composite plan is again available if the robot chooses to pick up medkit2 from room3 and hands it over to commX in hall4 which falls in his path. However, the CommX does not know that the robot plans to do this, and will continue with his original plan, which makes the robot's actions redundant, and the composite plan, though cheaper, is not a feasible plan in the current setting. Specifically, since there is no expectation of interventions, the robot must preserve the plan prefix of the original plan (that appears before and independently of the intervention) in the final composite plan. This then forms the first preservation constraint -

*Definition 3.0* : The composite plan  $\pi_{\mathbb{A}}$  that positively removes subplan  $\pi_H^{ij}$  from the original plan  $\pi_H$  of the human is a serendipitous plan iff  $\pi_{\mathbb{A}}(H)[1 : i - 1] = \pi_H[1 : i - 1]$ , where  $i = \text{argmin}_i [a = \pi_{\mathbb{A}}(H)[i] \wedge a \notin \pi_H], \forall a \in A_H$ .

Further, the composite plan must ensure that the effects of the actions of the robot  $R$  preserve the world state for the human's plan to continue executing beyond the serendipitous moment (because there is no commitment from the robot to help in future, and the human cannot plan to exploit future assistance), which provides our second preservation constraint below -

*Definition 3.1* : The composite plan  $\pi_{\mathbb{A}}$  that positively removes subplan  $\pi_H^{ij}$  from the original plan  $\pi_H$  of the human is a serendipitous plan iff  $\delta'(\bigcup_{\alpha \in \mathbb{A}} \mathbb{I}_{\alpha}, \pi_{\mathbb{A}}[1 : i']) \models \delta(\mathbb{I}_H, \pi_H[1 : j])$ .

We will now introduce a planner that can take into account all these constraints and produce serendipitous composite plans. Given the plan the human is executing, the robot decides on what serendipitous exceptions to introduce to make the cost of that plan lower. In doing this, the robot searches over a space of exceptions during execution time for the human's plan, as well as the length of the detours that those exceptions will cause (by "simulating" what it thinks the human will do in replanning).

### C. The Planner

The planning problem of the robot, defined in terms of the super-agent  $\mathbb{A} = \{R, H\}$  - given by  $\Pi = \langle D_{\mathbb{A}}, \theta_{\mathbb{A}}, \pi_H \rangle$

- consists of the domain model  $D_{\mathbb{A}}$ , the problem instance  $\theta_{\mathbb{A}} = \langle \mathbb{O}_{\mathbb{A}}, \mathbb{I}_{\mathbb{A}}, \mathbb{G}_{\mathbb{A}} \rangle$  (where  $\mathbb{O}$  are the objects or constants in the domain, and  $\mathbb{I}_{\mathbb{A}}$  and  $\mathbb{G}_{\mathbb{A}}$  are the initial and goal states of the super-agent respectively) and the original plan  $\pi_H$  of the human. Recall that we assumed completely known belief models, which means that in our current scenario, the robot starts with the full knowledge of the human's goal(s) and can predict the plan he is currently following (assuming optimality) - this forms  $\pi_H$ . Though we assume here that the intentions of the human are completely known, it is easy to have a plan recognition module that provides  $\pi_H$  as the most likely plan from a distribution over possible plans given observations up to that point. Another way to handle uncertainty in goals is to convert information from the set of possible plans to resource profiles as in [10]. However, our primary intention in this work, is to lay down the foundation of what it means to plan for serendipity given a known plan.

Planning for serendipity involves, as we discussed in the previous section, modeling complicated constraints between the human's plan and the composite plan being generated, which is not directly suited to be handled by conventional planners. We adopt the principles of planning for serendipity outlined thus far and propose the following IP-based planner (partly following the technique for IP encoding for state space planning outlined in [11]) to showcase these modes of behavior in our scenario.

Henceforth, when we refer to the domain  $D_{\alpha}$  of agent  $\alpha$ , we will mean the grounded (with objects  $\in \mathbb{O}$ ) version of its domain. Note that this might mean that some of the inter-agent actions (like handing over medkits) are now only available to the super-agent  $\mathbb{A}$ . i.e.  $\bigcup_{\alpha \in \mathbb{A}} A_{\alpha} \subseteq A_{\mathbb{A}}$ .

For the super agent, we define a binary action variable for action  $a \in A_{\mathbb{A}}$  at time step  $t$  as follows:

$$x_{a,t} = \begin{cases} 1, & \text{if action } a \text{ is executed by the super-agent } \mathbb{A} \\ & \text{at time step } t \\ 0, & \text{otherwise; } t \in \{1, 2, \dots, T\} \end{cases}$$

Also, for every proposition  $f$  at step  $t$  a binary state variable is introduced as follows:

$$y_{f,t} = \begin{cases} 1, & \text{if proposition is true in plan step } t \\ 0, & \text{otherwise; } \forall f \in S_{\mathbb{A}}, t \in \{0, 1, \dots, T\} \end{cases}$$

We also introduce two variables  $\xi_1, \xi_2 \in [1, T], 1 \leq \xi_1 < \xi_2 \leq T$  to represent the beginning and end of the subplan that gets positively removed by Definition 2.0. Finally, we add a new "no-operation" action  $A_{\alpha} \leftarrow A_{\alpha} \cup a_{\phi} \forall \alpha \in \mathbb{A}$  such that  $a_{\phi} = \langle \mathbb{N}, \mathbb{C}, \mathbb{P}, \mathbb{E} \rangle$  where  $\mathbb{N} = \text{NOOP}$ ,  $\mathbb{C} = 0$ ,  $\mathbb{P} = \{\}$  and  $\mathbb{E} = \{\}$ . The IP formulation modeling the interruptibility and preservation constraints is given by:

$$Obj : \min \sum_{a \in A_{\mathbb{A}}} \sum_{t \in \{1, 2, \dots, T\}} \mathbb{C}_a \times x_{a,t} + K \|\xi_2 - \xi_1\|$$

$$y_{f,0} = 1 \quad \forall f \in \bigcup_{\alpha \in \mathbb{A}} \mathbb{I}_{\alpha} \quad (1)$$

$$y_{f,0} = 0 \quad \forall f \notin \bigcup_{\alpha \in \mathbb{A}} \mathbb{I}_{\alpha} \quad (2)$$

$$y_{f,T} = 1 \quad \forall f \in \mathbb{G}_H \quad (3)$$

$$x_{a,t} \leq y_{f,t-1} \quad \forall a \in A_{\mathbb{A}}, s.t. f \in \mathbb{P}_a, t \in \{1, \dots, T\} \quad (4)$$

$$y_{f,t} \leq y_{f,t-1} + \sum_{a \in \text{add}(f)} x_{a,t} \\ s.t. \text{add}(f) = \{a | f \in \text{eff}^+(a)\}, a \in A_{\mathbb{A}}, \forall f \in S_{\mathbb{A}}, \\ t \in \{1, 2, \dots, T\} \quad (5)$$

$$y_{f,t} \leq 1 - \sum_{a \in \text{del}(f)} x_{a,t} \\ s.t. \text{del}(f) = \{a | f \in \text{eff}^-(a)\}, a \in A_{\mathbb{A}} \quad \forall f \in S_{\mathbb{A}}, \\ t \in \{1, 2, \dots, T\} \quad (6)$$

$$\xi_1 \leq \sum_t (t \times x_{a,t}) (1 - \sum_t \sum_{a \in \pi_{\alpha_H}} x_{a,t}) \\ + T(1 - \sum_t x_{a,t}) + T(\sum_t \sum_{a \in \pi_{\alpha_H}} x_{a,t}) \\ \forall a \in A_H, t \in \{1, 2, \dots, T\} \quad (7a)$$

$$x_{a,t} \geq \frac{1}{T}(\xi_1 - t) \quad \forall a \in \pi_H, t \in \{1, 2, \dots, T\} \quad (7b)$$

$$x_{a,t} \leq 1 + \frac{1}{T}(\xi_2 - t) \quad \forall a \in A_R, t \in \{1, \dots, T\} \quad (8)$$

$$x_{a,t} + x_{a_{\phi},t} \geq \frac{1}{T}(t - \xi_2) \quad \forall a \in \pi_H, t \in \{1, 2, \dots, T\} \quad (9)$$

$$\sum_{a \in A_{\alpha}} x_{a,t} + \sum_{a \in A_{\mathbb{A}} \setminus \bigcup_{\alpha \in \mathbb{A}} A_{\alpha}} x_{a,t} \leq 1 \\ \forall \alpha \in \mathbb{A}, t \in \{1, \dots, T\} \quad (10)$$

$$\sum_{a \in A_{\mathbb{A}}} \sum_{t \in \{1, 2, \dots, T\}} \mathbb{C}_a \times x_{a,t} \leq \text{cost}(\pi_H) \quad (11)$$

$$\xi_1, \xi_2 \in \{1, 2, \dots, T\}, \xi_2 \leq \xi_1 + 1 \quad (12)$$

$$y_{f,t} \in \{0, 1\} \quad \forall f \in S_{\mathbb{A}}, t \in \{0, 1, \dots, T\} \quad (13)$$

$$x_{a,t} \in \{0, 1\} \quad \forall a \in A_{\mathbb{A}}, t \in \{1, 2, \dots, T\} \quad (14)$$

where  $K$  is a large constant and  $T$  is the planning horizon.

Here, the objective function minimizes the sum of the cost of the composite plan and the length of the proposed positively removable subplan. Here we assume unit cost actions, i.e.  $\mathbb{C}_a = 1 \quad \forall a \in A_{\mathbb{A}}$  and then investigate the effect of varying the cost of the robot's actions with respect to the human's. Constraints (1) through (3) model the initial and goal conditions, while constraints (4) through (6) enforce the state equations that maintain the preconditions, and add and delete effects of the actions. Constraint (7a) specifies the value of  $\xi_1$  as per Definition 3.0. Specifically,  $\xi_1 = \text{argmin}_i [a = \pi_{\mathbb{A}}(H)[i] \wedge a \notin \pi_H], \forall a \in A_H$ . Thus  $\forall a \in A_H$  we write the following inequalities (the equations are written in a way so as to render all such constraints that do not belong to cases in Definition 3.0, trivially satisfied - since we already have  $1 \leq \xi_1 \leq T$  as the total range) -

$$\xi_1 \leq \begin{cases} \infty, & \text{if } a \notin \pi_{\mathbb{A}}(H) \\ \infty, & \text{if } a \in \pi_H \\ t, & \text{if } a = \pi_{\mathbb{A}}(H)[t] \wedge a \notin \pi_H \end{cases}$$

and constraint (7b) imposes Definition 3.0 as

$$x_{a,t} \begin{cases} > 0 \implies 1, & \text{if } a \in \pi_H \text{ and } t < \xi_1 \\ \in \{0, 1\}, & \text{otherwise} \end{cases}$$

Similarly, constraint (8) models Definition 3.1 by stopping

actions from the robot for  $t > \xi_2$  as follows -

$$x_{a,t} \begin{cases} < 1 \implies 0, & \text{if } a \in A_R \text{ and } t > \xi_2 \\ \in \{0, 1\}, & \text{otherwise} \end{cases}$$

Constraint (9) is optional and models Definition 2.0 (when ignored, Definition 2.0a is implied) as follows -

$$x_{a,t} \begin{cases} > 0 \implies 1, & \text{if } a \in \pi_H \cup a_\phi \text{ and } t > \xi_2 \\ \in \{0, 1\}, & \text{otherwise} \end{cases}$$

Constraint (10) imposes non concurrency on the actions of each agent (or inter-agent actions) during every epoch. Constraint (11) specifies that the generated composite plan should have lesser cost than the original human plan (again, this is optional). Finally constraints (12) to (14) provide the binary ranges of the variables. The constant  $K$  penalizes larger subplans from being removed (so as to minimize interference with the human's plan).

We will discuss the implementation of our planner and the statistics of its behavior in the next section. In the following discussion, we will illustrate how the implemented planner handles different configurations of the running example.

Going back to Figure 2, we note that the optimal plan for CommX in order to perform triage in `room1` involves picking up `medkit1` from `room2` -

```
1 - MOVE_COMMX_ROOM13_HALL8
2 - MOVE_REVERSE_COMMX_HALL8_HALL7
3 - MOVE_REVERSE_COMMX_HALL7_HALL6
4 - MOVE_REVERSE_COMMX_HALL6_HALL5
5 - MOVE_REVERSE_COMMX_HALL5_HALL4
6 - MOVE_REVERSE_COMMX_HALL4_HALL3
7 - MOVE_REVERSE_COMMX_HALL3_HALL2
8 - MOVE_REVERSE_COMMX_HALL2_ROOM2
9 - PICK_UP_MEDKIT_COMMX_MK1_ROOM2
10 - MOVE_COMMX_ROOM2_HALL2
11 - MOVE_REVERSE_COMMX_HALL2_HALL1
12 - MOVE_REVERSE_COMMX_HALL1_ROOM1
13 - CONDUCT_TRIAGE_COMMX_ROOM1
```

However, the robot can be proactive and decide to fetch `medkit2` and hand it over to him on his way towards `room1`. Indeed, this is the plan that the planner produces -

```
1 - MOVE_COMMX_ROOM13_HALL8
1 - MOVE_REVERSE_ROBOT_ROOM4_ROOM3
2 - MOVE_REVERSE_COMMX_HALL8_HALL7
2 - PICK_UP_MEDKIT_ROBOT_MK2_ROOM3
3 - MOVE_REVERSE_COMMX_HALL7_HALL6
4 - MOVE_REVERSE_COMMX_HALL6_HALL5
4 - MOVE_ROBOT_ROOM3_ROOM4
5 - MOVE_REVERSE_COMMX_HALL5_HALL4
5 - MOVE_ROBOT_ROOM4_HALL4
6 - HAND_OVER_ROBOT_COMMX_MK2_HALL4
6 - HAND_OVER_ROBOT_COMMX_MK2_HALL4
7 - MOVE_REVERSE_COMMX_HALL4_HALL3
8 - MOVE_REVERSE_COMMX_HALL3_HALL2
9 - MOVE_REVERSE_COMMX_HALL2_HALL1
10 - MOVE_REVERSE_COMMX_HALL1_ROOM1
11 - CONDUCT_TRIAGE_COMMX_ROOM1
```

#### D. Planning with Communication

The dynamics of the setting change somewhat when we allow certain forms of communication to exist between the

agents. Going back to the previous example, now it is no longer necessary for the robot to ensure that the prefix of the original human plan is respected (for example, the robot can inform `commX` that it is going to be in `hall4` to hand over `medkit2` to him), so that planning with communication changes the desiderata in terms of the preservation constrains in the plan generation process.

One immediate upshot of being able to communicate is that it is no longer necessary for the robot to preserve plan prefixes, and Definition 3.0 and correspondingly constraints (7b) and (7b) are no longer required. If, however, we wish to impose the interruptibility constraints from Definition 1.0 as  $\pi_A(H)[1:i-1] \subseteq \pi_H[1:i-1]$  (for a positively removable subplan  $\pi_H^i$ ) on the plan prefix, constraint 7b may now updated to the following -

$$x_{a,t} \leq 1 + \frac{1}{T}(t - \xi_1) \\ \forall a \in A_H \wedge a \notin \pi_H, t \in \{1, 2, \dots, T\} \quad (7b)$$

Finally, communication comes at a cost - too much communication might feel like interference from the point of view of the human. With this in mind, we can define the communication cost to be proportional to the number (or cost) of actions that the robot changes in the composite plan with respect to the original human plan.

*Definition 3.2* : The communication cost in the composite plan  $\pi_A$  is given by  $\mathcal{C} \propto \sum C_a \forall a \in \pi_A(H) \wedge a \notin \pi_H$ .

Thus we update the objective function of the IP with  $Obj \leftarrow Obj + \mathcal{C}$  (and remove constraints (7a) and (7b)).

Going back again to the world state in Figure 2, but now with `medkit1` in `room7`, we note that the optimal plan for CommX in order to perform triage in `room1` involves picking up `medkit1` from `room7`, as follows -

```
1 - MOVE_COMMX_ROOM13_HALL8
2 - MOVE_REVERSE_COMMX_HALL8_HALL7
3 - MOVE_REVERSE_COMMX_HALL7_ROOM7
4 - PICK_UP_MEDKIT_COMMX_MK1_ROOM7
5 - MOVE_COMMX_ROOM7_HALL7
6 - MOVE_REVERSE_COMMX_HALL7_HALL6
7 - MOVE_REVERSE_COMMX_HALL6_HALL5
8 - MOVE_REVERSE_COMMX_HALL5_HALL4
9 - MOVE_REVERSE_COMMX_HALL4_HALL3
10 - MOVE_REVERSE_COMMX_HALL3_HALL2
11 - MOVE_REVERSE_COMMX_HALL2_HALL1
12 - MOVE_REVERSE_COMMX_HALL1_ROOM1
13 - CONDUCT_TRIAGE_COMMX_ROOM1
```

The plan from the previous section is no longer a valid serendipitous plan because it violates Definition 3.0, as confirmed by the planner. However, the robot can choose to communicate its intention to handover `medkit2` and indeed, the planner once again produces the plan outlined in the previous section when communication is allowed.

## IV. EXPERIMENTAL RESULTS

In the following section we will go through simulations to illustrate some of the salient aspects of planning for serendipity, and provide a real world execution of the ideas discussed so far on the Nao Robot. The IP-planner has

TABLE I

COMPARISON OF COSTS OF TEAM PLANS B/W W/ COMM. VS W/O COMM.  
(AS COMPARED TO AVERAGE COST OF 9.825 FOR INDIVIDUAL PLANS)

Discount	w/o comm.	w/ comm. $\equiv$ comp. plan
0%	9.82 (1)	9.72 (13)
10%	9.81 (7)	9.65 (23)
30%	9.79 (7)	9.48 (34)
50%	9.76 (12)	9.25 (40)
70%	9.68 (29)	8.93 (62)
90%	9.55 (32)	8.51 (70)

TABLE II

RUNTIME PERFORMANCE OF THE PLANNER

	w/o comm.	w/ comm.	Comp. Optimal
Time (in sec)	28.68	33.93	44.96

been implemented on the IP-solver `gurobi`. The planner is available at <http://bit.ly/1zzvFB8>. The simulations were conducted on an Intel Xeon(R) CPU E5-1620 v2 @ 3.70GHz $\times$ 8 processor with a 62.9GiB memory. For the simulations, we build a suite of 200 test problems on the domain shown in Figure 2, by randomly generating positions for the two medkits and the positions of the two agents, and also randomly assigning a triage goal to the commander.

#### A. Different Flavors of Collaboration

In Table I we look at the full spectrum of costs incurred (to the *entire* team) by planning for individual plans to planning for serendipity (with and without communication) to optimal global plans and compare gains associated with each specific type of planning with respect to the individual optimal plans. Note that communication costs are set to zero in these evaluations so as to show the maximum gains potentially available by allowing communication. Also, for composite planning, the number of planning epochs was set to the length of the planning horizon of the original individual plan. Of course with higher planning horizons we will get more and more composite plans that make the robot do most of the work with discounted actions costs. Notice the gains in cost achieved through the different flavors of collaboration. The results indicate that, *for the given scenario*, planning for serendipity with communication essentially boiled down to bounded length composite optimal plans. The results also outline the expected trend of decreasing costs of the composite plan with respect to increasing discounts on the cost of the robot’s actions, as expected.

Table I also shows the effect of varying the discount factor on the percentage of problem instances that supported opportunities for the robot to plan for serendipity. That the numbers are low is not surprising given that we are planning for cases where the robot can help without being asked to, but notice how more and more instances become suitable for serendipitous collaboration as we reduce the costs incurred by the robot, indicating there is sufficient scope of exhibiting such behaviors for relatively lower costs of the robot’s actions as compared to the human’s.

Table II shows the runtime performance of the four types of planning approaches discussed above. The performance is evidently not affected much by the different modes of planning. Note that the time for generating the single plan is contained in these cases (for the composite plan also, the individual plan is produced to get the planning horizon).

#### B. Implementation on the Nao

We now illustrate the ideas discussed so far on the Nao Robot operating in a miniature implementation of the USAR scenario in Figure 2. We reproduce the scenarios mentioned in Sections III-C and III-D, and demonstrate how the Nao produces serendipitous moments during execution of the human’s plan. A video of the demonstration is available at [https://youtu.be/VCb0TaM\\_RJY](https://youtu.be/VCb0TaM_RJY).

## V. CONCLUSION

In this paper we propose a new planning paradigm - planning for serendipity - and provide a general formulation of the problem with assumptions of complete knowledge of the world state and agent models. We also illustrate how this can model proactive and helpful behaviors of autonomous agents towards humans operating in a shared setting like USAR scenarios. This of course raises interesting questions on how the approach can be adopted to a probabilistic framework for partially known models and goals of agents, and how the agents can use plan recognition techniques with observations on the world state to inform their planning process - questions we hope to address in future.

## REFERENCES

- [1] K. Talamadupula, J. Benton, S. Kambhampati, P. Schermerhorn, and M. Scheutz, “Planning for human-robot teaming in open worlds,” *ACM Trans. Intell. Syst. Technol.*, vol. 1, no. 2, pp. 14:1–14:24, Dec. 2010.
- [2] K. Talamadupula, G. Briggs, T. Chakraborti, M. Scheutz, and S. Kambhampati, “Coordination in human-robot teams using mental modeling and plan recognition,” in *Intelligent Robots and Systems (IROS), IEEE/RSJ International Conference on*, 2014, pp. 2957–2962.
- [3] E. Sisbot, L. Marin-Urias, R. Alami, and T. Simeon, “A human aware mobile robot motion planner,” *Robotics, IEEE Transactions on*, vol. 23, no. 5, pp. 874–883, Oct 2007.
- [4] M. Kuderer, H. Kretzschmar, C. Sprunk, and W. Burgard, “Feature-based prediction of trajectories for socially compliant navigation,” in *Proceedings of Robotics: Science and Systems*, 2012.
- [5] U. Koeckemann, F. Pecora, and L. Karlsson, “Grandpa hates robots - interaction constraints for planning in inhabited environments,” in *Proc. AAAI-2010*, 2014.
- [6] M. Cirillo, L. Karlsson, and A. Saffiotti, “Human-aware task planning: An application to mobile robots,” *ACM Trans. Intell. Syst. Technol.*, vol. 1, no. 2, pp. 15:1–15:26, Dec. 2010.
- [7] A. Gerevini, A. Saetti, and I. Serina, “An approach to temporal planning and scheduling in domains with predictable exogenous events,” *JAIR*, vol. 25, pp. 187–231, 2011.
- [8] N. Nilsson, “Triangle tables: A proposal for a robot programming language,” Technical Note 347, AI Center, SRI International, 1985.
- [9] Y. Zhang, V. Narayanan, T. Chakraborti, and S. Kambhampati, “A human factors analysis of proactive support in human-robot teaming,” in *Intelligent Robots and Systems (IROS), IEEE/RSJ International Conference on*, 2015.
- [10] T. Chakraborti, Y. Zhang, D. Smith, and S. Kambhampati, “Planning with stochastic resource profiles: An application to human-robot co-habitation,” in *ICAPS Workshop on Planning and Robotics*, 2015.
- [11] T. Vossen, M. O. Ball, A. Lotem, and D. S. Nau, “On the use of integer programming models in ai planning,” in *IJCAI*, 1999, pp. 304–309.