Planning for Serendipity - Altruism in Human-Robot Cohabitation

Tathagata Chakraborti[1] and **Gordon Briggs**[2] and **Kartik Talamadupula**[3]
Matthias Scheutz[2] and **David Smith**[4] and **Subbarao Kambhampati**[1]

Department of Computer Science[1]
Arizona State University
Tempe, AZ 85281, USA
{tchakra2,rao}@asu.edu

HRI Laboratory[2]
Tufts University
Medford, MA 02155, USA
{gbriggs,mscheutz}@cs.tufts.edu

Cognitive Learning Department[3]
IBM T. J. Watson Research Center
Yorktown Heights, NY 10598, USA
krtalamad@us.ibm.com

Intelligent Systems Division[4]
NASA Ames Research Center
Moffett Field, CA 94035, USA
david.smith@nasa.gov

Abstract

Recently there has been a lot of focus on human robot co-habitation issues that are usually orthogonal to 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 planning paradigm. 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.

Automated planners are increasingly being used to endow robots with autonomous planning capabilities in joint human-robot task scenarios (Talamadupula et al. 2010). 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 (Talamadupula et al. 2014) 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 (Sisbot et al. 2007; Kuderer et al. 2012) and in task planning (Koeckemann, Pecora, and Karlsson 2014; Cirillo, Karlsson, and Saffiotti 2010), that aim to provide social skills to robots so as to make them produce plans con-

forming 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 (which are not necessarily aligned with just the notion of avoidance of conflicts but can involve more positive and active help) 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 exogenous events (Gerevini, Saetti, and Serina 2011), 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 the different modes of such behavior. We also provide an IP-based planner that models the ideas involved in this planning paradigm. 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. Before we go into the details of the planning framework, let us first look at the scenario and understand the ideas involved in planning for serendipity in this context.
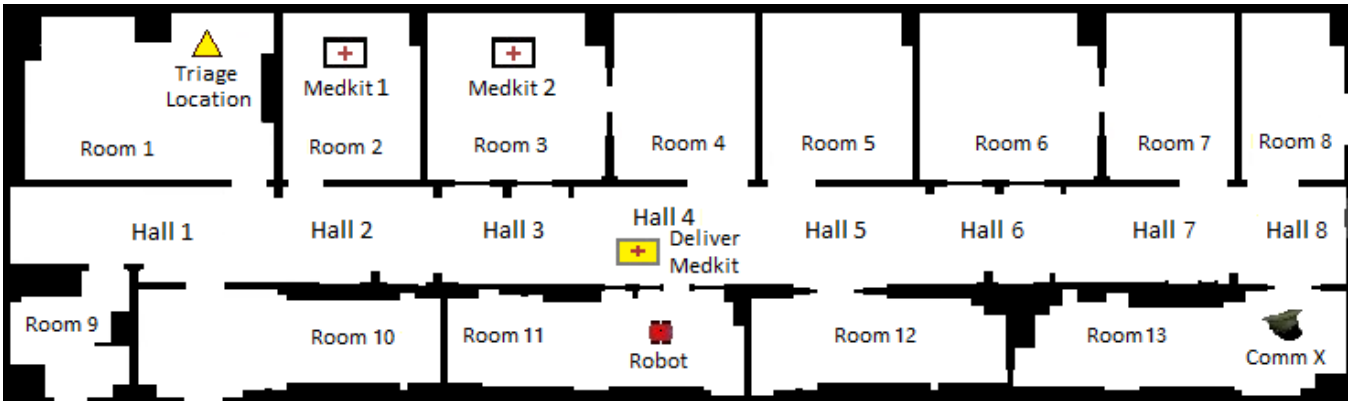
Figure 1: USAR setting involving a commander and a robot. The commanders have goal(s) to do triage at the different room locations, while the robot can help them both actively (as part of a team) or passively (planning for serendipity).

1 Overview of the USAR Scenario

Figure 1 shows a typical USAR setting, unfolding inside a building with interconnected rooms and hallways, with a human commander 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, dropoff 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, and as we discussed before, such teaming behaviors have been studied extensively in literature over the years. 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.

1.1 Agent Models

Each agent α (in the current scenario α is either H or R depending on when it is being used to refer to the human or the robot respectively) in the environment is described by a domain model $D_\alpha = \langle T_\alpha, V_\alpha, S_\alpha, A_\alpha \rangle$, where T_α is a set of object types; V_α is a set of variables that describe objects that belong to T_α; S_α is a set of named first-order logical predicates over the variables V_α that describe the world state $\mathbb{W}$; and A_α 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 preconditions 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_α 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 $\delta(\langle a_1, a_2, \ldots, a_n \rangle, s) = \delta(\langle a_2, \ldots, a_n \rangle, \delta(a_1, s))$.

Further, the robot can also maintain belief models Bel_α about the other agents in its environment as described in more detail in (Talamadupula et al. 2014). 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 while plan execution. We will relax this assumption later.

1.2 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 consisting of one or more agents in order to achieve a common goal.

Definition 1.0: An *individual plan* π_α of an agent α with the domain model D_α 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 $\pi_\alpha = \langle a_1, a_2, \ldots, a_n \rangle$, $a_i \in A_\alpha$ 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* $\pi_\mathbb{A} = \langle \mu_1, \mu_2, \ldots, \mu_n \rangle$, where $\mu = \{a_1, \ldots, a_{|\mathbb{A}|}\}$, $\mu(\alpha) = a \in A_\alpha \; \forall \mu \in \mathbb{U}$

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).

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 2.1 and 2.2), given the human's individual plan currently "planned" for execution.

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

Lemma 1.1b : A composite plan $\pi_\mathbb{A} = \langle \mu_1, \mu_2 \ldots, \mu_T \rangle$ 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 $t < T$, $[\pi_\alpha]_{execution}$ is not necessarily same as $\langle a_1, a_2, \ldots, a_T \rangle$, $a_i = \mu_i(\alpha)$, $\mu_i \in \pi_\mathbb{A}$.

The challenge then 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.

2 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, and 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 (Nilsson 1985) 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.

2.1 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 1), 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 $\pi_H = \langle a_1, a_2, \ldots, a_T \rangle$ of the human H with $\delta(\mathbb{I}_H, \pi_H) \models \mathbb{G}_H$, then any subplan $\pi_H^{ij} = \langle a_i, \ldots, a_j \rangle, 1 \leq i < j \leq |\pi_H|$ 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 $i' > i$, $\pi_\mathbb{A}(H) = (\subseteq \pi_H[1 : i-1]) \bullet \pi_\mathbb{A}(H)[i : i'] \bullet (\subseteq \pi_H[j+1 : |\pi_H|])$ and $C(\pi_\mathbb{A}(H)) < C(\pi_H)$ (here $\bullet$ 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 (described later in Section 2.3). We intend to do HRI studies in the future to see what kind of exceptions humans really respond to (Narayanan et al. 2015).

To account for normal human cognition, 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 $\pi_H = \langle a_1, a_2, \ldots, a_T \rangle$ of the human H with $\delta(\mathbb{I}_H, \pi_H) \models \mathbb{G}_H$, then any subplan $\pi_H^{ij} = \langle a_i, \ldots, a_j \rangle, 1 \leq i < j \leq |\pi_H|$ 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 $\pi_\mathbb{A}(H)[1 : i-1] \subseteq \pi_H[1 : i-1] = \langle a_1, \ldots, a_{i-1} \rangle$ and $\pi_\mathbb{A}(H)[i : |\pi_\mathbb{A}(H)|]$ is the optimal plan such that $\delta'(\delta'(\bigcup_{\alpha \in \mathbb{A}} \mathbb{I}_\alpha, \pi_\mathbb{A}[1 : i-1]), \pi_\mathbb{A}(H)[i : |\pi_\mathbb{A}(H)|]) \models \mathbb{G}_H$ 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 2.0-1 to determine opportunities to plan for such serendipitous moments.

2.2 Preservation Constraints

Let us now go back to the setting in Figure 1. 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 π_H^{ij} from the original plan π_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 π_H^{ij} from the original plan π_H of the human is a serendipitous plan iff $\delta'(\bigcup_{\alpha \in \mathbb{A}} \mathbb{I}_\alpha, \pi_\mathbb{A}[1 : j]) \models \delta(\mathbb{I}_H, \pi_H[1 : j])$.

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).

2.3 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 π_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 predicts the plan he is currently following (assuming optimality) - this forms π_H. Planning for serendipity involves 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 for 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 (Vossen et al. 1999)) to showcase these modes of behavior in our motivating scenario.

Henceforth, when we refer to the domain D_α of agent α, 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:

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

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

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

We also define two variables $\xi_1, \xi_2 \in [1,T], 1 \leq \xi_1 < \xi_2 \leq T$ to represent the subplan that gets positively removed by Definition 2.0. We also 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 (the constraints are explained after the formulation):

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

such that

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

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

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

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

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

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

$$\xi_1 \leq \sum_t (t \times x_{a,t})(1 - \sum_t \sum_{a \in \pi_{\alpha_1}} x_{a,t})$$
$$+ T(1 - \sum_t x_{a,t}) + T(\sum_t \sum_{a \in \pi_{\alpha_1}} x_{a,t})$$
$$\forall a \in A_H,\ t \in \{1,2,\ldots,T\} \tag{7a}$$

$$x_{a,t} \geq \tfrac{1}{T}(\xi_1 - t)\ \forall a \in \pi_H,\ t \in \{1,2,\ldots,T\} \tag{7b}$$

$$x_{a,t} \leq 1 + \tfrac{1}{T}(\xi_2 - t)\ \forall a \in A_R,\ t \in \{1,\ldots,T\} \tag{8}$$

$$x_{a,t} + x_{a_\phi,t} \geq \tfrac{1}{T}(t - \xi_2)\ \forall a \in \pi_H,\ t \in \{1,2,\ldots,T\} \tag{9}$$

$$\sum_{a \in A_\alpha} x_{a,t} + \sum_{a \in A_{\mathbb{A}} \setminus \bigcup_{\alpha \in \mathbb{A}} A_\alpha} x_{a,t} \leq 1$$
$$\forall \alpha \in \mathbb{A},\ t \in \{1,\ldots,T\} \tag{10}$$

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

$$\xi_1, \xi_2 \in \{1,2,\ldots,T\},\ \xi_2 \leq \xi_1 + 1 \tag{12}$$

$$y_{f,t} \in \{0,1\}\ \forall f \in S_{\mathbb{A}},\ t \in \{0,1,\ldots,T\} \tag{13}$$

$$x_{a,t} \in \{0,1\}\ \forall a \in A_{\mathbb{A}},\ t \in \{1,2,\ldots,T\} \tag{14}$$

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. $\mathbb{C}_a = 1\ \forall a \in A_{\mathbb{A}}$ 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 ξ_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 (where the constraints are written in a way such that $\beta \geq T$, so as to render such cases trivial since we already have $1 \leq \xi_1 \leq T$) -

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

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).

Going back to Figure 1, we note that the optimal plan for CommX in order to perform triage in room1 involves picking up medkit1 from room2 -

```
MOVE_COMMX_ROOM13_HALL8
MOVE_REVERSE_COMMX_HALL8_HALL7
MOVE_REVERSE_COMMX_HALL7_HALL6
MOVE_REVERSE_COMMX_HALL6_HALL5
MOVE_REVERSE_COMMX_HALL5_HALL4
MOVE_REVERSE_COMMX_HALL4_HALL3
MOVE_REVERSE_COMMX_HALL3_HALL2
MOVE_REVERSE_COMMX_HALL2_ROOM2
PICK_UP_MEDKIT_COMMX_MK1_ROOM2
MOVE_COMMX_ROOM2_HALL2
MOVE_REVERSE_COMMX_HALL2_HALL1
MOVE_REVERSE_COMMX_HALL1_ROOM1
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 it 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
8 - MOVE_REVERSE_COMMX_HALL4_HALL3
9 - MOVE_REVERSE_COMMX_HALL3_HALL2
11 - MOVE_REVERSE_COMMX_HALL2_HALL1
12 - MOVE_REVERSE_COMMX_HALL1_ROOM1
13 - CONDUCT_TRIAGE_COMMX_ROOM1
```

2.4 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 $\pi_{\mathbb{A}}(H)[1:i-1] \subseteq \pi_H[1:i-1]$ (for a positively removable subplan π_H^{ij}) on the plan prefix, constraint $7b$ may now updated to the following -

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

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 $\pi_{\mathbb{A}}$ is given by $\mathcal{C} \propto \sum \mathbb{C}_a \ \forall a \in \pi_{\mathbb{A}}(H) \wedge a \notin \pi_H$.

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 1, 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 -

```
MOVE_COMMX_ROOM13_HALL8
MOVE_REVERSE_COMMX_HALL8_HALL7
MOVE_REVERSE_COMMX_HALL7_ROOM7
PICK_UP_MEDKIT_COMMX_MK1_ROOM7
MOVE_COMMX_ROOM7_HALL7
MOVE_REVERSE_COMMX_HALL7_HALL6
MOVE_REVERSE_COMMX_HALL6_HALL5
```

```
MOVE_REVERSE_COMMX_HALL5_HALL4
MOVE_REVERSE_COMMX_HALL4_HALL3
MOVE_REVERSE_COMMX_HALL3_HALL2
MOVE_REVERSE_COMMX_HALL2_HALL1
MOVE_REVERSE_COMMX_HALL1_ROOM1
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.

3 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 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×8 processor with a 62.9GiB memory. For the simulations, we build a suite of 200 test problems on the domain shown in Figure 1, 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.

3.1 Different Flavors of Collaboration

In Table 1 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 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.

Table 2 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 3 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

Table 1: Comparison of Costs b/w w/ comm. vs w/o comm. vs composite optimal planning (as compared to average cost of 8.25 for individual plans)

Discount	w/o comm.	w/ comm.	Comp. Optimal
0%	9.82	9.72	9.70
10%	9.8115	9.6525	9.634
30%	9.7945	9.4805	9.458
50%	10.765	9.2475	9.21
70%	10.6815	8.931	8.89
90%	9.5525	8.5105	8.47

Table 2: Serendipitous Planning Opportunities w/ and w/o Comm. as a Function of the Discount Factor

Discount →	0%	10%	30%	50%	70%	90%
w/o comm.	1/200	7/200	7/200	12/200	29/200	32/200
w/ comm.	13/200	23/200	34/200	40/200	62/200	70/200

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).

3.2 Implementation on the Nao

We now illustrate the ideas discussed so far on the Nao Robot operating in a mock implementation of Figure 1 as shown in Figure 2. We reproduce the scenarios mentioned in Sections 2.3 and 2.4, 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/DvQBBOX6Qgo.

4   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.

Acknowledgments

Table 3: Runtime Performance of the Planner

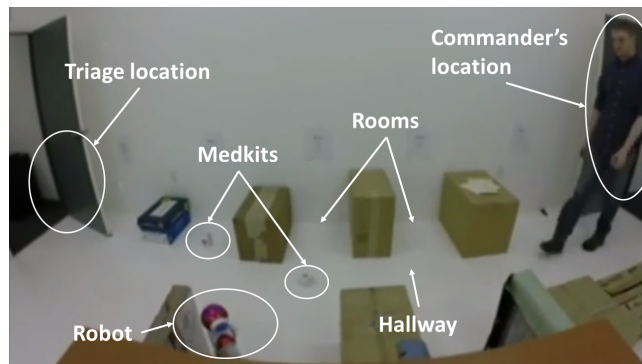|	w/o comm.	w/ comm.	Comp. Optimal
Time (in sec)	28.68	33.93	44.96

Figure 2: Demonstration of the various aspects of planning for serendipity on the Nao Robot in a mock implementation of the USAR setting from Figure 1.

References

[Cirillo, Karlsson, and Saffiotti 2010] Cirillo, M.; Karlsson, L.; and Saffiotti, A. 2010. Human-aware task planning: An application to mobile robots. *ACM Trans. Intell. Syst. Technol.* 1(2):15:1–15:26.

[Gerevini, Saetti, and Serina 2011] Gerevini, A.; Saetti, A.; and Serina, I. 2011. An approach to temporal planning and scheduling in domains with predictable exogenous events.

[Koeckemann, Pecora, and Karlsson 2014] Koeckemann, U.; Pecora, F.; and Karlsson, L. 2014. Grandpa hates robots - interaction constraints for planning in inhabited environments. In *Proc. AAAI-2010*.

[Kuderer et al. 2012] Kuderer, M.; Kretzschmar, H.; Sprunk, C.; and Burgard, W. 2012. Feature-based prediction of trajectories for socially compliant navigation. In *Proceedings of Robotics: Science and Systems*.

[Narayanan et al. 2015] Narayanan, V.; Zhang, Y.; Mendoza, N.; and Kambhampati, S. 2015. Automated planning for peer-to-peer teaming and its evaluation in remote human-robot interaction. In *Extended Abstract in ACM/IEEE International Conference on Human Robot Interaction (HRI)*.

[Nilsson 1985] Nilsson, N. 1985. Triangle tables: A proposal for a robot programming language. Technical report, Technical Note 347, AI Center, SRI International.

[Sisbot et al. 2007] Sisbot, E.; Marin-Urias, L.; Alami, R.; and Simeon, T. 2007. A human aware mobile robot motion planner. *Robotics, IEEE Transactions on* 23(5):874–883.

[Talamadupula et al. 2010] Talamadupula, K.; Benton, J.; Kambhampati, S.; Schermerhorn, P.; and Scheutz, M. 2010. Planning for human-robot teaming in open worlds. *ACM Trans. Intell. Syst. Technol.* 1(2):14:1–14:24.

[Talamadupula et al. 2014] Talamadupula, K.; Briggs, G.; Chakraborti, T.; Scheutz, M.; and Kambhampati, S. 2014. Coordination in human-robot teams using mental modeling and plan recognition. In *Intelligent Robots and Systems (IROS), IEEE/RSJ International Conference on*, 2957–2962.

[Vossen et al. 1999] Vossen, T.; Ball, M. O.; Lotem, A.; and Nau, D. S. 1999. On the use of integer programming models in ai planning. In *IJCAI*, 304–309.