

# A Mechanism of Generating Joint Plans for Self-interested Agents, and by the Agents

Wei HUANG

School of Computer Science and Technology  
University of Science and Technology of China  
Hefei, China

**Abstract**— Generating joint plans for multiple self-interested agents is one of the most challenging problems in AI, since complications arise when each agent brings into a multi-agent system its personal abilities and utilities. Some fully centralized approaches (which require agents to fully reveal their private information) have been proposed for the plan synthesis problem in the literature. However, in the real world, private information exists widely, and it is unacceptable for a self-interested agent to reveal its private information. In this paper, we define a class of multi-agent planning problems, in which self-interested agents' values are private information, and the agents are ready to cooperate with each other in order to cost efficiently achieve their individual goals. We further propose a semi-distributed mechanism to deal with this kind of problems. In this mechanism, the involved agents will bargain with each other to reach an agreement, and do not need to reveal their private information. We show that this agreement is a possible joint plan which is Pareto optimal and entails minimal concessions.

**Keywords**- joint plan; self-interested agent; private information; Pareto optimality; concession.

## I. INTRODUCTION

In the context of multi-agent planning, it is common that agents have different capabilities and they usually have to cooperate each other if they want to achieve their goals [1, 2, 3]. In many real settings, agents are self-interested, i.e., they may consider personal goals and utilities. Cooperation means finding a common plan which can increase agents' personal net benefits. Fully centralized approaches for finding such common plans in subsets of multi-agent planning problems have been proposed in the literature (see [4, 5] for instances). In these approaches, it is assumed that each agent reveals its personal goals and utilities to the other agents or to an arbitrator.

However, individualism often leads to existence of private information [6]. That is to say, in many real world multi-agent environments, self-interested agents may take advantage of knowledge about other agents' key information. Consequently, the agents don't want to reveal private information each other, and fully centralized planning approaches are not appropriate. On another hand, in order to perform an action successfully or deal with inconsistency of agents' preferences over actions, agents often need to negotiate with each other. For example, a glass bottle will be broken if there are multiple robots that are going to clasp it at the same time. In cases like this, fully

distributed planning approaches (see [2, 7] for examples) are also not appropriate.

So it is significant to deal with agents' capabilities (for cooperation and competition) and private information (for individualism) together in multi-agent planning. In this paper, we present an attempt to solve the problem. We define a rich class of planning problems in which (1) self-interested agents' values on joint plans are private information; (2) agents are ready to cooperate with each other in order to cost efficiently achieve their individual goals; (3) each agent can invite other agents to join a plan that is still beneficial to itself by providing certain amount of side payment [8]; and (4) each agent always persists in its personal goals even side payments are very attractive.

For this kind of problems, we provide a semi-distributed multi-agent planning mechanism (**MAPM**). In **MAPM**, each agent plays a role in the selection of the final joint plan. Agents will bargain with each other by giving joint plans (a joint plan includes a plan and a side payment function), and do not need to reveal their goals and utilities.

In traditional bargaining situations [9], the set of utility vectors (to the set of agents) derived from all possible proposals, is assumed to be compact and convex. In a planning domain, this assumption may be problematic, since the set of joint plans interesting to an agent is usually finite. In addition, each agent  $\phi$ 's utility function on joint plans, often takes integer values, since  $\phi$  cannot value a joint plan more closely than to the nearest penny. So in the bargaining situation discussed in this paper, only applicable joint plans and side payment functions taking integer values are considered. We show that **MAPM** always terminates; and the agreement reached by the agents through **MAPM**, is a joint plan which is Pareto optimal and entails minimal concessions.

The paper is organized as follows. Firstly, we give some preliminaries for representing planning problems. Secondly, we introduce the bargaining situation considered in **MAPM**. Thirdly, we define **MAPM** and prove its properties. Finally we discuss related work and future research directions.

## II. PLANNING DOMAINS AND PROBLEMS

The system we are interested to plan on is a multi-agent dynamic domain, modeling the potential evolutions of world.

**Definition 1:** A multi-agent planning domain is a tuple  $\mathcal{D} = \langle S, s_0, N, A, \rho \rangle$ , where  $S$  is the finite set of domain states,

$s_0 \in \mathcal{S}$  is the initial state,  $N = \{1, 2, \dots, n\}$  denotes the set of agents,  $A$  is the finite set of domain actions,  $\rho \subseteq \mathcal{S} \times N \times A \times \mathcal{S}$  is the domain transition relation.

At each time point,  $\mathcal{D}$  is in one of its states; initially,  $s_0$ . Agent  $\phi$  can perform action  $a$  on state  $s$  if  $\langle s, \phi, a, s' \rangle \in \rho$  for some  $s'$ . In this paper, we assume that all state transitions are deterministic, i.e.,  $(\forall \langle s, \phi, a \rangle \in \mathcal{S} \times N \times A) \{ | \{ s' \mid \langle s, \phi, a, s' \rangle \in \rho \} | \leq 1$  holds. All actions are assumed to be asynchronous. That is to say, at each time point, there is at most one agent that is going to perform an action. Therefore a plan can be formalized as a sequence of agent-action pairs.

**Definition 2:** A plan  $\pi$  over domain  $\mathcal{D} = \langle \mathcal{S}, s_0, N, A, \rho \rangle$  is a finite sequence in the form  $\langle \phi_1, a_1 \rangle, \langle \phi_2, a_2 \rangle, \dots, \langle \phi_m, a_m \rangle$  such that each  $\langle \phi_i, a_i \rangle \in N \times A$ .  $\varepsilon$  denotes the empty sequence.  $\pi$  is applicable if there exist  $s_1, s_2, \dots, s_m \in \mathcal{S}$  such that  $\langle s_{i-1}, \phi_i, a_i, s_i \rangle \in \rho$  for  $0 < i \leq m$ .  $m$  and  $s_0; s_1; \dots; s_m$  are called the length and the path of  $\pi$ , respectively.

In many realistic settings, agents are self-interested, have private goals and costs, and are motivated to act to increase their private net benefit. To capture such settings, for each agent  $\phi$ , we associate a set of goal states, a cost function on actions, and the reward  $\phi$  associates with the set of goal states. The set of goal states, cost function, and reward are assumed to be  $\phi$ 's private information (i.e., they are only observable to  $\phi$  itself). Formally:

**Definition 3:** A planning problem is a tuple  $\mathcal{P} = \langle \mathcal{D}, g, c, r, \delta \rangle$ , where (1)  $\mathcal{D} = \langle \mathcal{S}, s_0, N, A, \rho \rangle$  is a multi-agent planning domain; (2)  $g: N \rightarrow 2^{\mathcal{S}} \setminus \emptyset$  is a goal function that assigns each agent a set of goal states; (3)  $c: N \times A \rightarrow Z^+$  is a cost function that specifies the cost of action execution for each agent; (4)  $r: N \rightarrow Z^+$  is a function capturing the reward each agent associates with its own goal states; and (5)  $\delta \in Z^+$ , and only plans bounded in length by  $\delta$  are taken into consideration.

$g(\phi)$ ,  $c_\phi$  (It is required that  $c_\phi(a) = c(\phi, a)$  for all  $a \in A$ ), and  $r(\phi)$  are agent  $\phi$ 's private information that can not be revealed to other agents. So we freely interchange notations  $\phi.g$  and  $g(\phi)$ ,  $\phi.c$  and  $c_\phi$ ,  $\phi.r$  and  $r(\phi)$ .

Given a planning problem  $\mathcal{P}$ , let  $\Omega(\mathcal{P})$  denote the set of all applicable plans considered in  $\mathcal{P}$ . Suppose  $\phi$  is an agent and  $\pi = \langle \phi_1, a_1 \rangle, \langle \phi_2, a_2 \rangle, \dots, \langle \phi_m, a_m \rangle \in \Omega(\mathcal{P})$  such that  $s_0; s_1; \dots; s_m$  is the path of  $\pi$ . The utility of  $\pi$  to  $\phi$  is

$$u_\phi(\pi) = r' - \sum_{i=1}^m c_i \quad (1)$$

where  $m \leq \delta$ ;  $r' = \phi.r$  if  $s_m \in \phi.g$ , 0 otherwise; and  $c_i = \phi.c(a_i)$  if  $\phi = \phi_i$ ; 0 otherwise.

**Example 1:** A small planning domain  $\mathcal{D} = \langle \mathcal{S}, s_0, N, A, \rho \rangle$  is depicted in Fig. 1, where  $\mathcal{S} = \{s_0, s_1, s_2, s_3\}$ ,  $N = \{1, 2, 3\}$ ,  $A = \{a, b\}$ , and  $\langle s_0, 2, a, s_1 \rangle, \langle s_0, 3, a, s_1 \rangle, \dots \in \rho$ . It is easy to find that  $\pi = \langle 3, b \rangle; \langle 2, a \rangle$  is an applicable plan over  $\mathcal{D}$ . Let  $\mathcal{P} = \langle \mathcal{D}, g, c, r, 3 \rangle$  be a planning problem, where  $g, c, r$  are given in Table 1. So  $\pi \in \Omega(\mathcal{P})$ .

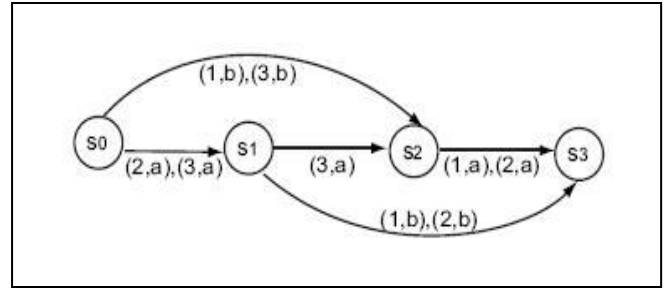


Figure 1. A small multi-agent planning domain  $\mathcal{D}$ .

TABLE I. Agents' Goals, Costs, and Rewards

$\phi \in N$	$g(\phi)$	$c(\phi, a)$	$c(\phi, b)$	$r(\phi)$
1	$\{s_1, s_3\}$	4	6	12
2	$\{s_2, s_3\}$	2	2	3
3	$\{s_3\}$	5	3	6

$$\Omega(\mathcal{P}), u_1(\pi) = 12, u_2(\pi) = 1, \text{ and } u_3(\pi) = 3.$$

### III. BARGAINING SITUATION

Let  $\mathcal{P}$  be a planning problem and  $N = \{1, 2, \dots, n\}$  be the set of agents in  $\mathcal{P}$ . Suppose  $\phi \in N$ ;  $\pi, \pi' \in \Omega(\mathcal{P})$ ; and  $u_\phi(\pi) > u_\phi(\pi')$ . Then agent  $\phi$  would prefer  $\pi$  to  $\pi'$ . If there is some  $\phi' \in N$  preferring  $\pi'$  to  $\pi$ , then  $\phi$  can propose a side payment such the amount is not greater than  $u_\phi(\pi) - u_\phi(\pi')$ . If this proposal does not work, then  $\phi$  must abandon  $\pi$  and consider  $\pi'$  instead. A joint plan can be seen as a structured contract, and consists of a plan and a side payment function.

**Definition 4:** A joint plan in  $\mathcal{P}$  is a pair  $p = \langle \pi, \zeta \rangle$ , where  $\pi \in \Omega(\mathcal{P})$  and  $\zeta: N \rightarrow Z$  is a side payment function satisfies  $\sum_{\phi \in N} \zeta(\phi) = 0$ . The utility of  $p$  to agent  $\phi$  is

$$u_\phi(p) = u_\phi(\pi) + \zeta(\phi) \quad (2)$$

In order to reach an agreement (i.e., a joint plan accepted by all the agents in  $N$ ), the agents can bargain with each other by proposing joint plans. Once an agreement  $p = \langle \pi, \zeta \rangle$  is reached, all the agents will cooperate to perform  $\pi$ , and the gross utility, i.e.,  $\sum_{\phi \in N} u_\phi(\pi)$  will be redistributed among  $N$  such that each agent  $\phi$ 's real income is  $u_\phi(p)$ . If the agents fail to reach agreement, then an agent  $\phi \in N$  would be selected randomly to control the evolution of the domain. In this case, we assume that the other agents take the "null action", i.e., do not interfere with the evolution.

Suppose  $d_\phi$  is the maximal value of utility  $\phi$  can achieve without other agent's involvement, i.e.,

$$d_\phi = \max_{\pi \in \Omega(\mathcal{P})} \{ u_\phi(\pi) \mid \text{Agents}(\pi) \subseteq \{ \phi \} \} \quad (3)$$

Where  $\text{Agents}(\pi)$  denotes the set of agents appear in  $\pi$ . Then  $\lfloor d_\phi / |N| \rfloor$  ( $\lfloor \cdot \rfloor$  and  $\lceil \cdot \rceil$  denote the ceil and the floor function on real numbers.) acts as  $\phi$ 's utility on the disagreement event,

denoted by  $u_\varphi^d$ . In other words,  $\varphi$  is not willing to cooperate with other agents if the cooperation can not bring to  $\varphi$  a utility value which is strictly greater than  $u_\varphi^d$  (individual rationality). It is not difficult to find that, without access to other agents' private information, each  $\varphi \in N$  can compute (eg., by a backward breadth-first search)  $\Pi_\varphi = \{\pi \in \Omega(\mathcal{P}) \mid u_\varphi(\pi) > u_\varphi^d\}$ , i.e., the set of plans interesting for  $\varphi$ . The set  $\Omega^+(P)$  of individual rational plans is defined as

$$\Omega^+(P) = \{\pi \in \Omega(P) \mid (\forall \varphi \in N) u_\varphi(\pi) > u_\varphi^d\} \quad (4)$$

It is easy to find that  $\Omega^+(P) = \bigcap_{\varphi \in N} \Pi_\varphi$ . We use  $u_\varphi^\perp$  and  $u_\varphi^T$  to denote the minimal utility and the maximal utility  $\varphi$  can gain in the situation where all agents are individual rational, respectively:

$$u_\varphi^\perp = \min_{\pi \in \Omega^+(P)} u_\varphi(\pi); u_\varphi^T = \max_{\pi \in \Omega^+(P)} u_\varphi(\pi) \quad (5)$$

Indeed  $u_\varphi^\perp$  is  $\varphi$ 's bottom line for bargaining, and  $u_\varphi^T$  is the ideal outcome of  $\varphi$ .  $\Lambda(P)$  denotes the set of all possible joint plans. Joint plan  $p = (\pi, \zeta) \in \Lambda(P)$  if and only if (1)  $\pi \in \Omega^+(P)$ , and (2)  $u_\varphi(p) \geq u_\varphi^\perp$  for each  $\varphi \in N$ .

$(u_1^T, \dots, u_n^T)$  is called the ideal point. For a joint plan  $p$ ,

$$\Delta(p) = \sum_{\varphi \in N} (u_\varphi^T - u_\varphi(p))^2 \quad (6)$$

describes the distance between the ideal point and the utility vector derived from  $p$ . In other words,  $\Delta(p)$  describes the concessions made by the agents to achieve  $p$ . This leads to the notion of solution which characterizes the Pareto optimal joint plans which entail minimal concessions.

**Definition 5:** Joint plan  $p$  is a solution to  $\mathcal{P}$  if it satisfies: (1)  $p \in \Lambda(P)$ , (2) there is no joint plan  $p' \in \Lambda(P)$  such that  $u_\varphi(p') > u_\varphi(p)$  for all  $\varphi \in N$ , and (3)  $\Delta(p) = \min_{p' \in \Lambda(P)} \Delta(p')$ .

This definition states that all self-interested agents should be individual rational at first (i.e., each  $\varphi \in N$  will not commit to a joint plan  $p$  if  $u_\varphi(p) < u_\varphi^\perp$ ). Second, a solution should be a Pareto optimal joint plan, in which no agent can increase its utility without decreasing other agents' utility. Finally, among the possible joint plans, a solution should entail minimal concessions.

**Example 2:** See Example 1. We can find that  $u_1^d = u_2^d = u_3^d = 0$ . So  $\pi \in \Omega^+(P)$ . Let  $p = (\pi, \zeta)$  be a joint plan such that  $\zeta(1) = -2$  and  $\zeta(2) = \zeta(3) = 1$ . Then  $u_1(p) = 10$ ,  $u_2(p) = 2$ , and  $u_3(p) = 4$ . In fact,  $u_1^T = 12$ ,  $u_2^T = 3$ ,  $u_3^T = 6$ ,  $\Delta(p) = 2^2 + 1^2 + 2^2 = 9$ , and  $p$  is a solution to  $\mathcal{P}$ .

Each  $\varphi \in N$  can compute  $\Pi_\varphi$ . And  $\Omega^+(P) = \bigcap_{\varphi \in N} \Pi_\varphi = \{\pi_i \mid 1 \leq i \leq 8\}$ , where  $\pi_1 = \langle 1, b \rangle; \langle 1, a \rangle$ ,  $\pi_2 = \langle 1, b \rangle; \langle 2, a \rangle$ ,  $\pi_3 = \langle 3, b \rangle; \langle 1, a \rangle$ ,  $\pi_4 = \langle 3, b \rangle; \langle 2, a \rangle$ ,  $\pi_5 = \langle 3, a \rangle; \langle 1, b \rangle$ ,  $\pi_6 = \langle 3, a \rangle; \langle 2, b \rangle$ ,  $\pi_7 = \langle 2, a \rangle; \langle 1, b \rangle$ ,  $\pi_8 = \langle 2, a \rangle; \langle 3, a \rangle; \langle 1, a \rangle$ . Agents' values are shown in Table 2, where  $u_\varphi^i$  denotes  $u_\varphi(\pi_i)$ .

TABLE II. Agents' Values on Each Plan

$\varphi$	$u_\varphi^1$	$u_\varphi^2$	$u_\varphi^3$	$u_\varphi^4$	$u_\varphi^5$	$u_\varphi^6$	$u_\varphi^7$	$u_\varphi^8$
1	2	6	8	12	6	12	6	8
2	3	1	3	1	3	1	1	1
3	6	6	3	3	1	1	6	1

#### IV. MECHANISM FOR GENERATING JOINT PLANS

In this section, we present **MAMP** a semi-distributed mechanism of generating joint plans for multiple self-interested agents, in which all the involved agents will bargain on the possible joint plans, and each agent's utility function and goal keep secret to other agents. **MAMP** is defined as follows.

Step1 Each  $\varphi \in N$  calculates  $\Pi_\varphi = \{\pi \mid u_\varphi(\pi) > u_\varphi^d\}$ , and sends  $\Pi_\varphi$  to an arbitrator  $\varphi^* \notin N$ .

Step2  $\varphi^*$  calculates  $\Omega^+(P) = \bigcap_{\varphi \in N} \Pi_\varphi$ . If  $\Omega^+(P) = \emptyset$ , then the result is **failure**, and the process stops. Otherwise,  $\varphi^*$  sets  $N' \leftarrow \emptyset$ ,  $\Pi \leftarrow \Omega^+(P)$ ,  $\mathbf{d}[\pi][\varphi] \leftarrow -\infty$ , and  $\mathbf{c}[\varphi] \leftarrow 0$  for all  $\varphi \in N$  and  $\pi \in \Omega^+(P)$ , and sends  $\Omega^+(P)$  to each  $\varphi \in N$ .

Step3 Each  $\varphi \in N$  puts the plans of  $\Omega^+(P)$  in a sequence  $\mathbf{sq}_\varphi = \pi_1; \pi_2; \dots$  such that  $u_\varphi(\pi_i) \geq u_\varphi(\pi_{i+1})$  for  $|\Omega^+(P)| > i \geq 1$ , and sets  $\mathbf{con}_\varphi \leftarrow 0$ . Let  $t \leftarrow 0$ .

Step4  $\varphi^*$  sends "Next proposal!" to each  $\varphi \in N - N'$ .

Step5 Let  $t \leftarrow t + 1$ . For each  $\varphi \in N - N'$ :

- If  $\mathbf{con}_\varphi = 0$  then:  $\varphi$  sets  $p_\varphi^t \leftarrow \text{Head}(\mathbf{sq}_\varphi)$ ;  $\mathbf{sq}_\varphi \leftarrow \text{Tail}(\mathbf{sq}_\varphi)$ ; if  $\mathbf{sq}_\varphi \neq \varepsilon$  then  $\varphi$  sets  $\mathbf{con}_\varphi \leftarrow u_\varphi(p_\varphi^t) - u_\varphi(\text{Head}(\mathbf{sq}_\varphi))$ <sup>1</sup>.
- Otherwise  $\varphi$  sets  $p_\varphi^t \leftarrow \text{hold}$ , and  $\mathbf{con}_\varphi \leftarrow \mathbf{con}_\varphi - 1$ .

Step6 Each  $\varphi \in N - N'$  sends  $p_\varphi^t$  to  $\varphi^*$ . If  $p_\varphi^t = \text{hold}$  then  $\varphi^*$  sets  $\mathbf{c}[\varphi] \leftarrow \mathbf{c}[\varphi] + 1$ . Otherwise  $\varphi^*$  sets  $\mathbf{d}[p_\varphi^t][\varphi] \leftarrow \mathbf{c}[\varphi]$ . Let  $\mathbf{ps}_\varphi \leftarrow \{\pi \mid \mathbf{d}[\pi][\varphi] \neq \infty\}$ .  $\varphi^*$  sets  $N' \leftarrow \{\varphi \mid \mathbf{ps}_\varphi = \Omega^+(P)\}$ ,  $\omega \leftarrow \bigcap_{\varphi \in N} \mathbf{ps}_\varphi$ . If  $\omega = \emptyset$ , then goto step 4.

Step7  $\varphi^*$  selects a plan  $\pi$  from  $\omega$  such that  $\sum_{\varphi \in N} \mathbf{d}[\pi][\varphi] \leq \sum_{\varphi \in N} \mathbf{d}[\pi'][\varphi]$  for all  $\pi' \in \omega$ , sets  $\pi^* \leftarrow \pi$ ,  $\theta \leftarrow \sum_{\varphi \in N} \mathbf{d}[\pi^*][\varphi]$ , and  $\Pi \leftarrow \{\pi \mid \sum_{\varphi \in N} \min(\mathbf{d}[\pi][\varphi], \mathbf{c}[\varphi]) < \theta\}$ . If  $\Pi \neq \emptyset$  then goto step 4.  $\varphi^*$  sets  $M \leftarrow N$ .

Step8 Let  $N'' \leftarrow \{\varphi \in M \cap N' \mid \mathbf{c}[\varphi] < \lceil \theta / |M| \rceil\}$ .  $\varphi^*$  sets  $M \leftarrow M - N''$ ,  $\theta \leftarrow \theta - \sum_{\varphi \in N''} \mathbf{c}[\varphi]$ ,  $M' \leftarrow \text{Stop}(M, \theta, \mathbf{c})$  (see Fig. 2), and  $\zeta[\varphi] \leftarrow \mathbf{d}[\pi^*][\varphi] - \mathbf{c}[\varphi]$  for each  $\varphi \in N''$ . If  $M \subseteq M' \cup N'$  then goto step 12.

<sup>1</sup> Given a nonempty sequence  $\mathbf{sq}$ , **Head** returns the first item of  $\mathbf{sq}$ , and **Tail** returns a sequence  $\mathbf{sq}'$  such that  $\mathbf{sq} = \text{Head}(\mathbf{sq}) \bullet \mathbf{sq}'$ . For example, let  $\mathbf{sq} = q_1; q_2; q_3$ , then  $\text{Head}(\mathbf{sq}) = q_1$ , and  $\text{Tail}(\mathbf{sq}) = q_2; q_3$ .

Step9  $\varphi^*$  sends "Next proposal!" to each  $\varphi \in M - (M' \cup N')$ .

Step10 Let  $t \leftarrow t+1$ . For each  $\varphi \in M - (M' \cup N')$ :

- If  $\text{con}_\varphi = 0$  then:  $\varphi$  sets  $p_\varphi^t \leftarrow \text{Head}(\mathbf{sq}_\varphi)$ ;  $\mathbf{sq}_\varphi \leftarrow \text{Tail}(\mathbf{sq}_\varphi)$ ; if  $\mathbf{sq}_\varphi \neq \varepsilon$  then  $\varphi$  sets  $\text{con}_\varphi \leftarrow u_\varphi(p_\varphi^t) - u_\varphi(\text{Head}(\mathbf{sq}_\varphi))$ .
- Otherwise  $\varphi$  sets  $p_\varphi^t \leftarrow \text{hold}$ , and  $\text{con}_\varphi \leftarrow \text{con}_\varphi - 1$ .

Step11 Each  $\varphi \in M - (M' \cup N')$  sends  $p_\varphi^t$  to  $\varphi^*$ . If  $p_\varphi^t = \text{hold}$  then  $\varphi^*$  sets  $\mathbf{c}[\varphi] \leftarrow \mathbf{c}[\varphi] + 1$ . Otherwise  $\varphi^*$  sets  $\mathbf{d}[p_\varphi^t][\varphi] \leftarrow \mathbf{c}[\varphi]$ . Let  $\text{ps}_\varphi \leftarrow \{\pi | \mathbf{d}[\pi][\varphi] \neq \infty\}$ .  $\varphi^*$  sets  $N' \leftarrow \{\varphi | \text{ps}_\varphi = \Omega^\perp(\mathcal{P})\}$ . Goto step 8.

Step12  $\varphi^*$  sets  $\theta \leftarrow \theta - \sum_{\varphi \in M - M'} \mathbf{c}[\varphi]$ ,  $\zeta[\varphi] \leftarrow \mathbf{d}[\pi^*][\varphi] - \mathbf{c}[\varphi]$  for each  $\varphi \in M - M'$ , and  $\zeta[\varphi] \leftarrow \mathbf{d}[\pi^*][\varphi] - \lfloor \theta / |M'| \rfloor$  for each  $\varphi \in M'$ . Let  $k \leftarrow \theta - |M'| \lfloor \theta / |M'| \rfloor$ .  $\varphi^*$  selects  $k$  agents (denoted as  $\varphi'_1, \dots, \varphi'_k$ ) from  $M'$  randomly, and sets  $\zeta[\varphi'_i] := \zeta[\varphi'_i] - 1$  for  $1 \leq i \leq k$ .  $\varphi^*$  announces the result of the procedure is  $\langle \pi^*, \zeta \rangle$ , and the process stops.

Observe this mechanism. We can find that, for all  $\varphi \in N$ : (1)  $\varphi$  does not communicate with other agents in  $N$  directly; (2)  $\varphi$  gets no information about other agents' proposals from  $\varphi^*$  during the course of bargaining; and (3)  $\varphi^*$  only knows  $u_\varphi(\pi) - u_\varphi(\pi')$  if  $\pi$  and  $\pi'$  have been proposed by  $\varphi$ . So in the course of bargaining, no agent  $\varphi \in N$  can extract information that would allow it to infer something about other agents' private information. In addition, for all  $\varphi \in N$  and  $\pi, \pi' \in \Omega^\perp(\mathcal{P})$ , the arbitrator  $\varphi^*$  does not know: (1)  $u_\varphi(\pi)$  (and of course, also  $\varphi.g$ ,  $\varphi.c$ , and  $\varphi.r$ ), and (2)  $u_\varphi(\pi) - u_\varphi(\pi')$  if  $\pi$  or  $\pi'$  has not been proposed by  $\varphi$ .

Consider the planning problem  $\mathcal{P}$  depicted in Example 1. We apply **MAPM** to  $\mathcal{P}$ . Firstly, each  $\varphi \in N$  calculates  $\Pi_\varphi$ , and  $\varphi^*$  finds that  $\Omega^\perp(\mathcal{P}) = \{\pi_i | 8 \geq i \geq 1\}$ , where each  $\pi_i$  is given in Example 2 (agents' values on these plans are shown in TABLE II). And then, agents in  $N$  begin to bargain. The relevant data generated by **MAPM** is illustrated in TABLE III, where  $\Pi = \{\pi_1, \pi_2, \pi_4, \pi_6, \pi_7\}$ ,  $\mathbf{h}$  and  $\Omega$  denote hold and  $\Omega^\perp(\mathcal{P})$ , respectively. Lastly, the procedure stops at time  $t=9$ , and  $\varphi^*$  announces the result is  $p = \langle \pi_4, \zeta \rangle$ , where  $\langle \zeta(1), \zeta(2), \zeta(3) \rangle = \langle -1, 0, 1 \rangle$  ( $\langle -2, 1, 1 \rangle$ , or  $\langle -2, 0, 2 \rangle$ ). So  $\langle u_1(p), u_2(p), u_3(p) \rangle = \langle 11, 1, 4 \rangle$  ( $\langle 10, 2, 4 \rangle$ , or  $\langle 10, 1, 5 \rangle$ ).

In fact, agent  $\varphi \in N$  proposes a side payment (i.e., makes a concession) such the amount is 1 based on  $p_\varphi^t$ , if it sends  $p_\varphi^{t+1} = \text{hold}$  to  $\varphi^*$ . Suppose  $\pi, \pi' \in \Omega^\perp(\mathcal{P})$  such that  $u_\varphi(\pi) > u_\varphi(\pi')$ , and  $\varphi$  sends  $\pi$  and  $\pi'$  to  $\varphi^*$  at time  $t$  and  $t'$ , respectively. Let  $C = |\{t' > t \mid p_\varphi^{t'} = \text{hold}\}|$ . In **MAPM**, it is required that  $t < t'$  and  $C = u_\varphi(\pi) - u_\varphi(\pi')$ . Please note that we do not aim at dominant-strategy truthful mechanisms in which private information does not influence the final result. In this paper, it is assumed that private information has value to agents, and to keep it secret is each agent's responsibility.

Consequently, we aim at mechanisms in which any agent is not sure that lying is better than truth telling if she does not know other agents' private information. Now let us see what would happen if  $\varphi$  deviates from **MAPM**:

- If  $t > t'$  then it is possible that  $\pi$  will be discarded and  $\varphi$  will suffer loss. For example, see TABLE II, TABLE III and suppose that agent 1 proposes  $\pi_3$  instead of  $\pi_4$  at  $t=1$ . Then  $\varphi^*$  will announce that  $p' = \langle \pi_3, \zeta' \rangle$  is the result such that  $\zeta'(1) = -1$ . It is easy to find that  $u_1(p') = 7 < u_1(p)$ .
- If  $C < u_\varphi(\pi) - u_\varphi(\pi')$  then it is possible that  $\pi'$  will be chosen and the concession made by  $\varphi$  will be underestimated (see  $\mathbf{d}[\pi^*][\varphi]$  at step 8 and step 12). For example, see TABLE II, TABLE III and suppose that agent 3 proposes  $\pi_3$  instead of hold at  $t=4$ . Then  $\varphi^*$  will announce that  $p' = \langle \pi_3, \zeta' \rangle$  is the result such that  $\zeta'(3) = -1$  or  $-2$ . It is easy to find that  $u_3(p') = 2$  or  $1 < u_3(p)$ .
- If  $C > u_\varphi(\pi) - u_\varphi(\pi')$  then it is possible that  $\pi$  will be chosen and the benefit gained by  $\varphi$  will be overestimated (see  $\mathbf{c}[\varphi]$  and  $\lfloor \theta / |N| \rfloor$  at step 8 and step 12). For example, see TABLE II, TABLE III and suppose that agent 3 insists on  $\pi_1$  (i.e.,  $p_3^1 = \pi_1$  and  $p_3^t = \text{hold}$  for all  $t > 1$ ). Then  $\varphi^*$  will announce that  $p' = \langle \pi_1, \zeta' \rangle$  is the result such that  $\zeta'(3) = -4$ . It is easy to find that  $u_3(p') = -2 < u_3(p)$ .

Consequently, the agents in  $N$  will follow **MAPM**, even though they are self-interested. We now show some properties of **MAPM**. The first key result states that **MAPM** always terminates.

```

1. subroutine Stop(M,  $\theta$ ,  $\mathbf{c}$ )
2.    $M' \leftarrow M$ ;
3.   while  $M' \neq \emptyset$ 
4.      $mc \leftarrow \min_{\varphi \in M'} \mathbf{c}[\varphi]$ ;
5.     if  $mc * |M'| \geq \theta$  then break;
6.      $M' \leftarrow \{\varphi \in M' \mid \mathbf{c}[\varphi] > mc\}$ ;
7.   return  $M'$ ;

```

Figure 2. Stop subroutine.

TABLE III. Data Generated by **MAPM** for the Example

t	$p_1^t$	$p_2^t$	$p_3^t$	$\omega$	$\pi^*$	$\Pi$	M	M'	$\theta$
1	$\pi_4$	$\pi_1$	$\pi_1$	$\emptyset$		$\Omega$			
2	$\pi_6$	$\pi_3$	$\pi_2$	$\emptyset$		$\Omega$			
3	h	$\pi_5$	$\pi_7$	$\emptyset$		$\Omega$			
4	h	h	h	$\emptyset$		$\Omega$			
5	h	h	h	$\emptyset$		$\Omega$			
6	h	$\pi_2$	h	$\emptyset$		$\Omega$			
7	$\pi_3$	$\pi_4$	$\pi_3$	$\{\pi_3\}$	$\pi_3$	$\Pi'$			
8	$\pi_8$	$\pi_6$	$\pi_4$	$\{\pi_3, \pi_4\}$	$\pi_4$	$\{\pi_1\}$			
9	h	$\pi_7$	h	$\{\pi_3, \pi_4\}$	$\pi_4$	$\emptyset$	N	N	5

**Theorem 1: MAPM** is guaranteed to terminate at some time  $T \leq |\Omega^\perp(\mathcal{P})| + \max_{\varphi \in N} (u_\varphi^T - u_\varphi^\perp)$ .

**Proof.** Observe **MAPM**. It is easy to find that computation of every step of **MAPM** always terminates. Pick any  $\varphi \in N$ ,  $\pi \in \Omega^\perp(\mathcal{P})$ , and  $t \geq 1$ . Then:

1. if **MAPM** does not stop at time  $t$ , then there exists  $\varphi' \in N$  sending a joint plan to  $\varphi^*$  at  $t$ ;
2. if  $\{\pi' \in \Omega^\perp(\mathcal{P}) \mid \varphi \text{ has sent } \pi' \text{ before } t\} = \Omega^\perp(\mathcal{P})$  then  $\varphi$  will not send any joint plan to  $\varphi^*$  at any  $t' > t$ ;
3. if  $t \geq \{ \pi' \in \Omega^\perp(\mathcal{P}) \mid u_\varphi(\pi') \geq u_\varphi(\pi) \} + u_\varphi^T - u_\varphi(\pi)$  then there exists  $1 \leq t' \leq t$  such that  $\varphi$  sends  $\pi$  at  $t'$ .

According to items 1-3, we can find that **MAPM** stops at some  $T \leq |\Omega^\perp(\mathcal{P})| + \max_{\varphi \in N} (u_\varphi^T - u_\varphi^\perp)$ .  $\square$

The second property states that if there is a solution for  $\mathcal{P}$ , then **MAPM** will not fail.

**Proposition 1: failure** is the result of **MAPM** if and only if there is no solution to  $\mathcal{P}$ .

**Proof.** Observe **MAPM** and we can find that **failure** is the result of **MAPM** if and only if  $\Omega^\perp(\mathcal{P}) = \emptyset$ . Now we show that  $\Omega^\perp(\mathcal{P}) = \emptyset$  if and only if there is no solution to  $\mathcal{P}$ .

Obviously, if  $\Omega^\perp(\mathcal{P}) = \emptyset$  then there is no solution to  $\mathcal{P}$ .

Suppose  $\Omega^\perp(\mathcal{P}) \neq \emptyset$ . Then there exists  $\pi \in \Omega^\perp(\mathcal{P})$  such that  $(\forall \pi' \in \Omega^\perp(\mathcal{P})) \sum_{\varphi \in N} u_\varphi(\pi') \leq \sum_{\varphi \in N} u_\varphi(\pi)$ . Let  $P = \{ \langle \pi', \zeta \rangle \in \Lambda(\mathcal{P}) \mid \pi' = \pi \}$ . It is easy to find that  $P \neq \emptyset$ . So there exists  $p \in P$  such that  $(\forall p' \in P) \Delta(p) \leq \Delta(p')$ . Pick any  $p' \in \Lambda(\mathcal{P})$  (suppose  $p' = \langle \pi', \zeta' \rangle$ ). We have  $\sum_{\varphi \in N} u_\varphi(p') = \sum_{\varphi \in N} u_\varphi(\pi') \leq \sum_{\varphi \in N} u_\varphi(\pi) = \sum_{\varphi \in N} u_\varphi(p)$ , i.e., there is no  $p' \in \Lambda(\mathcal{P})$  such that  $u_\varphi(p') > u_\varphi(p)$  for all  $\varphi \in N$ . Obviously,  $(\forall p' \in \Lambda(\mathcal{P})) \sum_{\varphi \in N} u_\varphi(p) \leq \sum_{\varphi \in N} u_\varphi^T$ . So there exists  $p'' \in P$  such that for each  $\varphi \in N$ :

$$\begin{cases} u_\varphi(p'') = u_\varphi(p) & \text{if } u_\varphi(p') \geq u_\varphi^T \\ u_\varphi^T \geq u_\varphi(p'') \geq u_\varphi(p) & \text{if } u_\varphi(p') < u_\varphi^T \end{cases} \quad (7)$$

Then  $\Delta(p') \geq \Delta(p'') \geq \Delta(p)$ . Consequently,  $\Delta(p) = \min_{p' \in \Lambda(\mathcal{P})} \Delta(p')$  and  $p$  is a solution to  $\mathcal{P}$ . That is to say, if there is no solution to  $\mathcal{P}$ , then  $\Omega^\perp(\mathcal{P}) = \emptyset$ .  $\square$

In the following discussion, we suppose that, **MAPM** goes out of the loop consisting of step 4, 5, and 6 at time  $T''$ , goes out of the loop consisting of step 4, 5, 6, and 7 at time  $T'$ , and terminates at time  $T$ .  $\omega_t, \pi_t^*, \Pi_t, M_t$ , and  $\theta_t$  denote the values of  $\omega, \pi^*, \Pi, M$ , and  $\theta$  at time  $1 \leq t \leq T$ .

Let us now consider the quality of the result. As shown by the following proposition, the plan given by **MAPM** maximizes the gross utility.

**Proposition 2:** If  $p = \langle \pi, \zeta \rangle$  is the result of **MAPM**, then  $\sum_{\varphi \in N} u_\varphi(\pi) \leq \sum_{\varphi \in N} u_\varphi(\pi')$  for each  $\pi' \in \Omega^\perp(\mathcal{P})$ .

**Proof.** It is easy to find that, if  $p = \langle \pi, \zeta \rangle$  is the result then  $\pi = \pi^*_{T'}$  (see step 7 of **MAPM**). Now we show that  $(\forall \pi' \in \Omega^\perp(\mathcal{P})) \sum_{\varphi \in N} u_\varphi(\pi') \leq \sum_{\varphi \in N} u_\varphi(\pi^*_{T'})$ .

According to the steps from 4 to 7 in **MAPM**, we have (for any  $T'' \leq t \leq T'$  and  $T'' \leq t' < T'$ ):

1.  $\sum_{\varphi \in N} u_\varphi(\pi^*_t) = \max_{\pi \in \omega_t} \sum_{\varphi \in N} u_\varphi(\pi)$ ;
2.  $\pi^*_t \in \omega_t, \omega_t \subseteq \omega_{t'+1}$ ;
3.  $\Pi_{t-1} \supseteq \Pi_t, \Pi_{t-1} - \Pi_t = \{ \pi \in \Pi_{t-1} \mid \sum_{\varphi \in N} \min(\mathbf{d}[\pi][\varphi], \mathbf{c}[\varphi]) \geq \sum_{\varphi \in N} \mathbf{d}[\pi^*_t][\varphi] \}$ ;
4.  $\omega_{T''} \subseteq \Pi_{T''-1} = \Omega^\perp(\mathcal{P})$ , and  $\Pi_{T''} = \emptyset$ .

Pick any  $T'' \leq t \leq T'$ . According to item 3, we have

$$(\forall \pi \in \Pi_{t-1} - \Pi_t) \sum_{\varphi \in N} u_\varphi(\pi) \leq \sum_{\varphi \in N} u_\varphi(\pi^*_t) \quad (8)$$

According to items 1 and 2, we have

$$\sum_{\varphi \in N} u_\varphi(\pi^*_t) \leq \sum_{\varphi \in N} u_\varphi(\pi^*_{T'}) \quad (9)$$

According to formula (8), (9), and item 4, we have  $(\forall \pi \in \Omega^\perp(\mathcal{P})) \sum_{\varphi \in N} u_\varphi(\pi) \leq \sum_{\varphi \in N} u_\varphi(\pi^*_{T'})$ .  $\square$

Let us now give our final result which characterizes the quality of the result of **MAPM**. The following theorem shows that the resulting joint plan is a solution to  $\mathcal{P}$ .

**Theorem 2:** If  $p \neq \text{failure}$  is the result of **MAPM**, then  $p$  is a solution to  $\mathcal{P}$ .

**Proof.** Observe **MAPM**. We can find that  $\mathbf{d}[\pi^*_{T'}][\varphi] = u_\varphi^T - u_\varphi(\pi^*_{T'})$  at any  $t \geq T'$  for all  $\varphi \in N$ . So (see step 8, 9, 10, 11, and 12) the result of **MAPM** should be  $p = \langle \pi^*_{T'}, \zeta \rangle$ , such that:

- $u_\varphi(p) = u_\varphi^\perp$  for all  $\varphi \in N - M'_T$ ,
- $\theta_T = \sum_{\varphi \in N} (u_\varphi^T - u_\varphi(\pi^*_{T'})) - \sum_{\varphi \in N - M'_T} \mathbf{c}[\varphi]$ ,
- $M'' \subseteq M'_T$  and  $|M''| = \theta_T - |M'_T| \lfloor \theta_T / |M'_T| \rfloor$ ,
- $u_\varphi(p) = u_\varphi^T - \lfloor \theta_T / |M'_T| \rfloor - 1$  for all  $\varphi \in M''$ , and
- $u_\varphi(p) = u_\varphi^T - \lfloor \theta_T / |M'_T| \rfloor$  for all  $\varphi \in M'_T - M''$ .

Now we show that  $p$  is a solution to  $\mathcal{P}$ .

It is easy to find that  $\pi^*_{T'} \in \Omega^\perp(\mathcal{P})$ . According to the Stop subroutine (see Fig. 2),  $u_\varphi^T - u_\varphi^\perp \geq \lfloor \theta_T / |M'_T| \rfloor$  for all  $\varphi \in M'_T$ . So  $u_\varphi(p) \geq u_\varphi^\perp$  for all  $\varphi \in N$ . Consequently,  $p \in \Lambda(\mathcal{P})$ . Pick any  $p' = \langle \pi', \zeta' \rangle \in \Lambda(\mathcal{P})$ . According to Proposition 2, we have:

$$\sum_{\varphi \in N} u_\varphi(p') = \sum_{\varphi \in N} u_\varphi(\pi') \leq \sum_{\varphi \in N} u_\varphi(\pi^*_{T'}) = \sum_{\varphi \in N} u_\varphi(p) \quad (10)$$

So there is no  $p' \in \Lambda(\mathcal{P})$  such that  $u_\varphi(p') > u_\varphi(p)$  for all  $\varphi \in N$ .

Let  $P = \{ \langle \pi^*_{T'}, \zeta' \rangle \in \Lambda(\mathcal{P}) \mid (\forall \varphi \in N) \zeta'(\varphi) \leq u_\varphi^T - u_\varphi(\pi^*_{T'}) \}$  and  $P' = \{ p'' \in P \mid (\forall \varphi \in N - M'_T) u_\varphi(p'') = u_\varphi^\perp \}$ . It is easy to find that:

1. there exists  $p'' \in P$  such that  $\Delta(p'') \leq \Delta(p')$ , and
2.  $\Delta(p) = \min_{p'' \in P'} \Delta(p'')$ .

According to steps 8, 10, 11, and 12,  $(\forall \varphi \in M'_T) (\forall \varphi' \in N - M'_T) 0 \leq u_\varphi^T - u_\varphi^\perp - \lfloor \theta_T / |M'_T| \rfloor \leq u_\varphi^T - u_\varphi^\perp$  and  $\theta_T + \sum_{\varphi \in N - M'_T} (u_\varphi^T - u_\varphi^\perp)$

$u_{\varphi}^{\perp} = \sum_{\varphi \in N} (u_{\varphi}^T - u_{\varphi}(\pi_{\varphi}^{*T}))$ . So there exists  $p^{\wedge} \in P'$  such that  $\Delta(p^{\wedge}) \leq \Delta(p)$ . According to items 1 and 2,  $\Delta(p) \leq \Delta(p^{\wedge}) \leq \Delta(p) \leq \Delta(p')$ , i.e.,  $\Delta(p) = \min_{p' \in \Lambda(\varphi)} \Delta(p')$ .  $\square$

#### V. RELATED WORK, AND FUTURE WORK

Unlike the previous work that consider planning for self-interested agents in fully adversarial settings (e.g., [10, 11]), [4, 5] propose the notion of planning games in which self-interested agents are ready to cooperate with each other to increase their personal net benefits. Fully centralized planning approaches have been proposed for subsets of planning games (see [4, 5] for examples). All these approaches require agents to fully disclose their private information. However, this requirement is not acceptable in this paper. Following the tradition of non-cooperative game theory, [7] presents a distributed multi-agent planning and plan improvement method, which is guaranteed to converge to stable solutions for congestion planning problems [12, 13]. But multi-agent planning problems discussed in this paper are different from congestion planning problems, in which each agent can plan individually to reach its goal from its initial state and no other agent can contribute to achieving its goal. On another hand, probably because of the loose interaction between agents in congestion planning problems, private information is not a focus of the work. [14] investigates algorithms for solving a restricted class of "safe" coalition-planning games, in which all the joint plans constructed by combining local solutions to the smaller planning games over disjoint subsets of agents are valid. It is easy to find that the multi-agent planning problems discussed in this paper are not "safe".

Different from the mechanisms [15, 16, 17] of single agent's plan execution, a mechanism of multiple self-interested agents' joint plan execution must deal with values' realization. As future work, we plan to investigate agents' power [8] and values' realization in joint plans' execution. Planning domains can be nondeterministic. So we will redefine the concept of solutions and multi-agent planning algorithms in strong [18] or probabilistic style. On another hand, we will design a more general bargaining mechanism for generating joint plans, which can deal with changing goals [19], incomplete information [20], and concurrent actions.

#### ACKNOWLEDGMENT

The work is sponsored by the Fundamental Research Funds for the Central Universities of China (Issue Number WK0110000026), and the National Natural Science Foundation of China under grant No.61105039.

#### REFERENCES

- [1] Ronen I. Brafman, Carmel Domshlak: From one to many: Planning for loosely coupled multi-agent systems. In: ICAPS (2008)
- [2] Raz Nissim, Ronen I. Brafman, Carmel Domshlak: A general, fully distributed multi-agent planning algorithm. In: AAMAS, pp. 1323-1330. (2010)
- [3] Alexandros Belesiotis, Michael Rovatsos, Iyad Rahwan: Agreeing on plans through iterated disputes. In: AAMAS, pp. 765-772. (2010)
- [4] Ronen I. Brafman, Carmel Domshlak, Yagil Engel, Moshe Tennenholtz: Planning games. In: IJCAI, pp. 73-78. (2009)
- [5] Ronen I. Brafman, Carmel Domshlak, Yagil Engel, Moshe Tennenholtz: Transferable utility planning games. In: AAI, pp. 709-714. (2010)
- [6] Ning Sun, Zaifu Yang: A double-track adjustment process for discrete markets with substitutes and complements. *Econometrica*. 77(3), 933-952 (2009)
- [7] Anders Jonsson, Michael Rovatsos: Scaling up multiagent planning: A best-response approach. In: ICAPS, pp. 114-121. (2011)
- [8] Sviatoslav Brainov, Tuomas Sandholm: Power, dependence and stability in multiagent Plans. In: AAI, pp. 11-16. (1999)
- [9] John Nash: The bargaining problem. *Econometrica*. 18(2), 155-162 (1950)
- [10] M. H. Bowling, R. M. Jensen, M. M. Veloso: A formalization of equilibria for multiagent planning. In: IJCAI, pp. 1460-1462. (2003)
- [11] R. Ben Larbi, S. Konieczny, P. Marquis: Extending classical planning to the multi-agent case: A Game-theoretic Approach. In: ECSQARU, pp. 731-742. (2007)
- [12] Rosenthal, R. W: A class of games possessing pure-strategy Nash equilibria. *International Journal of Game Theory*. 2, 65-67 (1973)
- [13] Monderer, D., Shapley, L. S.: Potential games. *Games and Economic Behaviour*. 14(1), 124-143 (1996)
- [14] Matt Crosby, Michael Rovatsos: Heuristic multiagent planning with self-interested Agents. In: AAMAS, pp. 1213-1214. (2011)
- [15] Piergiorgio Bertoli, Alessandro Cimatti, Marco Pistore, Paolo Traverso: A framework for planning with extended goals under partial observability. In: ICAPS, pp. 215--225. (2003)
- [16] Sebastian Sardina, Giuseppe De Giacomo, Yves Lesperance, Hector J. Levesque: On the limits of planning over belief states under strict uncertainty. In: KR, pp. 463-471. (2006)
- [17] Wei Huang, Zhonghua Wen, Yunfei Jiang, Hong Peng: Structured plans and observation reduction for plans with contexts. In: IJCAI, pp. 1721-1727. (2009)
- [18] Alessandro Cimatti, Marco Pistore, Marco Roveri, Paolo Traverso: Weak, strong, and strong cyclic planning via symbolic model checking. *Artificial Intelligence*. 147, 35-84 (2003)
- [19] Tran Cao Son, Chiaki Sakama: Negotiation using logic programming with consistency restoring rules. In: IJCAI, pp. 930-935. (2009)
- [20] Piergiorgio Bertoli, Alessandro Cimatti, Marco Roveri, Paolo Traverso: Strong planning under partial observability. *Artificial Intelligence*. 170, 337-384 (2006)