Bargain over Joint Plans*

Wei Huang[1], Dongmo Zhang[1], Yan Zhang[1], and Laurent Perrussel[2]

[1] Intelligent Systems Laboratory, University of Western Sydney, Australia
[2] IRIT, Université de Toulouse, France

Abstract. This paper studies the problem of multi-agent planning in the environment where agents may need to cooperate in order to achieve their individual goals but they do so only if the cooperation is beneficial to each of them. We assume that each agent has a reward function and a cost function that determines the agent's utility over all possible plans. The agents negotiate to form a joint plan through a procedure of alternating offers of joint plans and side-payments. We propose an algorithm that generates an agreement for any given planning problem and show that this agreement maximizes the gross utility and minimizes the distance to the ideal utility point.

Key words: multi-agent planning, joint plan, side-payment, bargaining

1 Introduction

Multiagent planning has been an emerging research topic in recent years in the area of Artificial Intelligence [1–5, 7]. Most existing studies on multiagent planning involve planning for common goals, plan coordinating, plan merging and synchronized planning. Most of the existing frameworks on multiagent planning are based either on the assumption that all agents have common goals therefore will be fully cooperative for a joint plan or on the assumption that all agents must reveal their private information, such as goals, rewards, costs and/or utilities, to other agents or arbitrators. In many real-world situations, none of the assumptions satisfies. It is a great challenge to find a joint plan for a multiagent system in which all agents are self-interested with individual goals and private information.

In this paper, we propose a solution to multiagent planning based on the following scenario:

- Each agent in the system has its own goals, reward of goal achievement and costs of actions.
- All agents are self-interested but profit-driven. An agent only concerns about its own goals. However, to attract other agents to join its plan, an agent may offer the other agents some payment (named side-payment) if the other agents agree on the joint plan.

* This research was supported by the Australian Research Council through Linkage Project LP0777015.

– An agent can make a proposal of a plan with actions from the other agents or its own (therefore a joint plan) and a side-payment scheme. An agent can accept other agents' proposal if the net profit it receives from this plan (possible reward minus costs plus side-payment) surpasses any of its own plans, reject the proposal by making a counter proposal.

Based on the about scenario, we propose a planning procedure, named *Planning Procedure based on Bargaining* (PPB). The procedure is based on an alternating-offer model of bargaining for two-agent bargaining situations [8]. The planning procedure proceeds in several rounds. In each round, only one agent can make a proposal, which consists of a plan and a side payment scheme. If the other agent accepts the proposal, the procedure terminates and the current proposal becomes the final agreement; otherwise, it is the other agent's turn to make a proposal. We show that PPB is correct, complete, and terminating.

This paper is structured as follows. Firstly, we introduce some formal preliminaries to represent the planning problems. Secondly, we define the concept of plan proposals and bargaining mechanism. Thirdly, we propose a planning procedure based on the bargaining mechanism and show its properties. Finally, we discuss related work and future research directions.

2 Planning domains and problems

In this section we present a model of dynamic systems based on which the planning problems that will be dealt with in this paper is described.

A *multi-agent planning domain* is a tuple $\Sigma = \langle \mathcal{S}, s_0, \Phi, \mathcal{A}, \mathcal{T} \rangle$, where $\mathcal{S}$ is a set of states, $s_0 \in \mathcal{S}$ is the initial state, Φ is a non-empty set of agents, $\mathcal{A}$ is a set of actions, and $\mathcal{T} \subseteq \mathcal{S} \times \Phi \times \mathcal{A} \times \mathcal{S}$ represents the state transition relation. $\langle s, \varphi, a, s' \rangle \in \mathcal{T}$ means that φ can perform action a at state s and bring about s' as one of the possible result states.

For simplicity, we assume in this paper that $|\{s' \in \mathcal{S} : \langle s, \varphi, a, s' \rangle \in \mathcal{T}\}| \leq 1$ for each $\langle s, \varphi, a \rangle$ in $\mathcal{S} \times \Phi \times \mathcal{A}$, i.e., we only consider deterministic state transitions. All actions are assumed to be asynchronous, that is to say, at most one agent performs an action at each state.

Definition 1. *Given a planning domain Σ, a* plan *π for Σ is a finite sequence in the form $\langle \varphi_1, a_1 \rangle; \langle \varphi_2, a_2 \rangle; \cdots; \langle \varphi_n, a_n \rangle$, where $\varphi_i \in \Phi$ and $a_i \in \mathcal{A}$. The plan π is called to be* applicable *to Σ if there exist $s_1, s_2, \ldots, s_n \in \mathcal{S}$ such that $\langle s_{i-1}, \varphi_i, a_i, s_i \rangle \in \mathcal{T}$ for all $0 < i \leq n$. s_n and n are referred to as the* last state *and the* length *of the plan, denoted by* $\textsc{Lstate}(\pi)$ *and* $\textsc{Length}(\pi)$*, respectively.* $\textsc{Agts}(\pi)$ *denotes the set of agents that are involved in π, i.e.,* $\textsc{Agts}(\pi) = \{\varphi \in \Phi : \varphi$ *appears in* $\pi\}$.

Give a planning domain, assume that each agent has its own goals, rewards if the goals are achieved and costs of actions. A multi-agent planning problem is to find a joint plan that can achieve the goals of all the agents meanwhile maximize their rewards and minimize their costs of actions.

Definition 2. *A* planning problem *is a tuple* $\mathcal{P} = \langle \Sigma, \mathcal{G}, r, c \rangle$*, where*

- $\Sigma = \langle \mathcal{S}, s_0, \Phi, \mathcal{A}, \mathcal{T} \rangle$ *is a planning domain.*
- $\mathcal{G} : \Phi \to 2^{\mathcal{S}}$ *is a goal function that specifies each agent's goal states.*
- $r : \Phi \to \mathbb{Z}_+$ *is a reward function that assigns to each agent a non-negative integer, representing the reward an agent can received if its goals are achieved.*
- $c : \Phi \times \mathcal{A} \to \mathbb{Z}_+$ *is a cost function that specifies the cost of each action to each agent.*

Note that for every agent φ, $\mathcal{G}(\varphi)$, $r(\varphi)$, and $c_\varphi = c(\varphi, a)$ are φ's private information. Therefore we write $\varphi.\mathcal{G}, \varphi.r$, and $\varphi.c$ instead of $\mathcal{G}(\varphi), r(\varphi)$, and c_φ, respectively, to indicate that these functions are implemented in agent φ.

Given a planning problem $\mathcal{P}$, let $\Omega(\mathcal{P})$ denote the set of all the applicable plans for the planning domain of $\mathcal{P}$. For each agent $\varphi \in \Phi$ and $\pi = \langle \varphi_1, a_1 \rangle; \langle \varphi_2, a_2 \rangle; \ldots \in \Omega(\mathcal{P})$, we define φ's *utility* of π as follows:

$$u_\varphi(\pi) = \textsc{Rew}_\varphi(\pi) - \sum_{i=1}^{\textsc{Length}(\pi)} \textsc{Cost}_\varphi(\varphi_i, a_i)$$

where $\textsc{Rew}_\varphi(\pi) = \varphi.r$ if $\textsc{Lstate}(\pi) \in \varphi.\mathcal{G}$; 0 otherwise and $\textsc{Cost}_\varphi(\varphi_i, a_i) = \varphi.c(a_i)$ if $\varphi = \varphi_i$; 0 otherwise.

We use $u_\varphi^\perp$ to denote the maximal value of utility that φ can achieve without other agent's involvement, i.e., $u_\varphi^\perp = \max_{\pi \in \Omega(\mathcal{P})} \{u_\varphi(\pi) | \textsc{Agts}(\pi) \subseteq \{\varphi\}\}$. $u_\varphi^\perp$ acts as φ's bottom line for bargaining. In other words, φ is willing to cooperate with other agents only if the cooperation can bring to φ a utility value which is strictly greater than $u_\varphi^\perp$ (individual rationality). Let $\Omega^\perp(\mathcal{P})$ be the set of plans which are individual rational, i.e., $\Omega^\perp(\mathcal{P}) = \{\pi \in \Omega(\mathcal{P}) | (\forall \varphi \in \Phi) u_\varphi(\pi) > u_\varphi^\perp\}$.

Similarly, we use $u_\varphi^\top$ to denote the maximal utility the agent φ can gain with respect to the current planning situation provided all other agents are individual rational, i.e., $u_\varphi^\top = \max_{\pi \in \Omega^\perp(\mathcal{P})} u_\varphi(\pi)$. Indeed $u_\varphi^\top$ is the ideal outcome of φ.

3 Bargaining Situation

To simplify the presentation of our approach, we will focus on two-agent planning problems, i.e., $\Phi = \{\varphi_{-1}, \varphi_1\}$. We call utility pair $(u_{\varphi_{-1}}^\top, u_{\varphi_1}^\top)$ the *ideal point*, denoted by $\textsc{IdP}(\mathcal{P})$. For any $j \in \{-1, 1\}$ and $\{\pi', \pi\} \subseteq \Omega^\perp(\mathcal{P})$, if $u_{\varphi_j}(\pi') > u_{\varphi_j}(\pi)$, then agent φ_j will prefer π' to π. If φ_{-j} does not agree to perform π', then φ_j can propose a side payment such that the amount proposed to φ_{-j} is not greater than $u_{\varphi_j}(\pi') - u_{\varphi_j}(\pi) - 1$. If this proposal does not work, then φ_j must abandon π' and consider π instead.

Definition 3. *A proposal to* $\mathcal{P}$ *is a pair* $p = \langle \pi, \xi \rangle$ *such that* π *is a plan for the planning domain of* $\mathcal{P}$*,* $\xi : \Phi \to \mathbb{Z}$ *is a side payment function which satisfies* $\sum_{\varphi \in \Phi} \xi(\varphi) = 0$*. For any* $k \in \mathbb{Z}$*,* ξ_k *denotes the side payment function that assigns* k *to* φ_1*, and* $-k$ *to* φ_{-1}*. For each* $\varphi \in \Phi$*, the utility of* p *to* φ *is defined as:* $u_\varphi(p) = u_\varphi(\pi) + \xi(\varphi)$*.*

$\textsc{Pro}(\mathcal{P})$ denotes the set of possible proposals. Proposal $p = \langle \pi, \xi_k \rangle \in \textsc{Pro}(\mathcal{P})$ if and only if: (1) $\pi \in \Omega^{\perp}(\mathcal{P})$ and, (2) $u_{\varphi_{-1}}(p) > u^{\perp}_{\varphi_{-1}}$ and $u_{\varphi_1}(p) > u^{\perp}_{\varphi_1}$.

In order to reach an agreement (i.e., a proposal accepted by the two agents), the agents can bargain with each other by proposing proposals one by one. Once an agreement $p = \langle \pi, \xi \rangle$ is reached, all the agents in $\textsc{Agts}(\pi)$ will cooperate to perform π, and the *gross utility*, i.e., $\sum_{\varphi \in \Phi} u_{\varphi}(\pi)$ will be redistributed among Φ such that each agent φ's real income is $u_{\varphi}(p)$. For a proposal p to $\mathcal{P}$, we use $\textsc{Dis}(p) = \sqrt{(u^{\top}_{\varphi_{-1}} - u_{\varphi_{-1}}(p))^2 + (u^{\top}_{\varphi_1} - u_{\varphi_1}(p))^2}$ to denote the distance between $\textsc{IdP}(\mathcal{P})$ and the utility pair derived from p. In other words, $\textsc{Dis}(p)$ describes the concessions made by the two agents to achieve p. This leads to the notion of *solution* which characterizes the Pareto optimal proposals which entail minimal concessions.

Definition 4. *Proposal p is a* solution *to $\mathcal{P}$ if it satisfies the following three conditions:*

Individual rationality: $p \in \textsc{Pro}(\mathcal{P})$;
Pareto optimality: *there is no proposal $p' \in \textsc{Pro}(\mathcal{P})$ such that $u_{\varphi}(p') > u_{\varphi}(p)$ for all $\varphi \in \Phi$;*
Minimal concession: $\textsc{Dis}(p) = \textsc{Min}\{\textsc{Dis}(p') | p' \in \textsc{Pro}(\mathcal{P})\}$.

4 The bargaining mechanism

In this section, we present a planning procedure based on bargaining, and show its properties. The procedure is used for two-agent planning settings, in which all utility functions and goals are private information and cannot be revealed.

The planning procedure based on bargaining (PPB) is defined as follows.

step 1: Each agent $\varphi \in \Phi$ calculates the set of plans $bups_{\varphi} = \{\pi | (u_{\varphi}(\pi) > u^{\perp}_{\varphi}) \wedge (\textsc{Length}(\pi) \leq \delta)\}$[3], and sends $bups_{\varphi}$ to an arbitrator φ^*.
step 2: φ^* calculates $\Omega^{\perp}(\mathcal{P}) = bups_{\varphi_{-1}} \cap bups_{\varphi_1}$. If $\Omega^{\perp}(\mathcal{P}) = \emptyset$, then φ^* announces the result of the procedure is failure, and the procedure stops. Otherwise, φ^* sets the set of plans to be considered $ps(0) := \Omega^{\perp}(\mathcal{P})$, $i := \textsc{Rand}(\{-1, 1\})$[4], sends $ps(0)$ and i to each $\varphi \in \Phi$.
step 3: Each $\varphi_j \in \Phi$ sets its proposal being considered $p_{\varphi_j}(0) := \langle \textsc{Rand}(pls_{\varphi_j}), \xi_0 \rangle$, where

$$pls_{\varphi_j} = \arg\max_{\pi \in ps(0)} u_{\varphi_j}(\pi),$$

and sends $p_{\varphi_j}(0)$ to φ_{-j}. Let $t := 0$, $\theta_{-1} := 0$, and $\theta_1 := 0$.
step 4: If $u_{\varphi_i}(p_{\varphi_{-i}}(t)) \geq u_{\varphi_i}(p_{\varphi_i}(t))$, then φ_i sends done to φ_{-i}, goto **step 7**. Otherwise φ_i sets $ps(t+1) := \{\pi \in ps(t) | u_{\varphi_i}(\pi) > u_{\varphi_i}(p_{\varphi_{-i}}(t))\}$, and φ_{-i} sets $p_{\varphi_{-i}}(t+1) := p_{\varphi_{-i}}(t)$.

[3] We adopt and, for ease of presentation further strengthen the *simple agents* assumption, requiring each plan to be bounded in length by a fixed δ.
[4] Given a set, $\textsc{Rand}$ returns an element of the set randomly.

step 5: Suppose $p_{\varphi_i}(t) = \langle \pi, \xi \rangle$. If $ps(t+1) = \emptyset$ or $u_{\varphi_i}(p_{\varphi_i}(t)) > \textsc{Max}\{u_{\varphi_i}(\pi) \mid \pi \in ps(t+1)\}$ then $\theta_i := \theta_i + 1$ and φ_i sets $p_{\varphi_i}(t+1) := \langle \pi, \xi' \rangle$ such that $\xi'(\varphi_i) = \xi(\varphi_i) - 1$ and $\xi'(\varphi_{-i}) = \xi(\varphi_{-i}) + 1$. Otherwise φ_i sets $p_{\varphi_i}(t+1) := \langle \textsc{Rand}(pls'_{\varphi_i}), \xi_0 \rangle$, where

$$pls'_{\varphi_i} = \arg \max_{\pi' \in ps(t+1)} u_{\varphi_i}(\pi').$$

step 6: φ_i sends $p_{\varphi_i}(t+1)$ to φ_{-i}. Let $t := t+1$ and $i := -i$. Return to step 4.

step 7: Suppose $p_{\varphi_{-i}}(t) = \langle \pi^*, \xi^* \rangle$. Then φ^* sets $j := \textsc{Rand}(\{-1, 1\})$, and announces $p = \langle \pi^*, \xi' \rangle$ is the result of the procedure, where $\xi'(\varphi_j) = \xi^*(\varphi_j) + \theta_{\varphi_j} - \llcorner 0.5 * w \lrcorner$[5], $\xi'(\varphi_{-j}) = \xi^*(\varphi_{-j}) + \theta_{\varphi_{-j}} - \ulcorner 0.5 * w \urcorner$, and $w = \theta_{\varphi_{-1}} + \theta_{\varphi_1}$.

If we observe this procedure, we remark that, for all $j \in \{-1, 1\}$, φ_j only sends proposals to φ_{-j} and φ^*. So for all $\pi \in \Omega^{\perp}(\mathcal{P})$, φ_{-j} and φ^* can not know $u_{\varphi_j}(\pi)$ (and of course, also $\varphi_j.\mathcal{G}$, $\varphi_j.r$, and $\varphi_j.c$) during the procedure.

We now show the properties of PPB. The first key result states that PPB always terminates in polynomial time.

Theorem 1. *Under the* simple agents *assumption, PPB is guaranteed to terminate, and it is polynomial in* $\min\{u^*_{\varphi_{-1}}, u^*_{\varphi_1}\}$, *where* $u^*_{\varphi_i} = u^{\top}_{\varphi_i} - min\{u_{\varphi_i}(\pi) | \pi \in \Omega^{\perp}(\mathcal{P})$ *and* $u_{\varphi_{-i}}(\pi) = u^{\top}_{\varphi_{-i}}\}$.

The second property states that if there is a solution for the planning problem, then the proposed procedure will not fail.

Theorem 2. *failure is the result of PPB if and only if there is no solution to* $\mathcal{P}$.

The following theorem shows that the resulting proposal is a *solution* to $\mathcal{P}$.

Theorem 3. *If PPB returns* $p \neq$ *failure, then* p *is a solution to* $\mathcal{P}$.

5 Conclusion and the related work

In this paper, we have proposed a model of multi-agent planning problems based on a bargaining mechanism. We have considered a class of planning situations in which each agent has its own goals, reward function and cost function. Agents bargain over joint plans with possible side payments. We have proposed a planning procedure which possesses the following properties: (1) the procedure always terminates in polynomial time; (2) for any given planning problem, if the set of individual rational plans is non-empty, the procedure can generate a joint plan at its termination; (3) the side payment associated with the resulting plan leads to a bargaining solution that is individual rational and Pareto optimal with minimal distance to the ideal point.

Most of the early work on multiagent planning is built up on fully cooperative multi-agent systems, such as the multi-entity model [7] and MA-STRIPS

[5] $\ulcorner \urcorner$ and $\llcorner \lrcorner$ denote the ceil and floor function on real numbers, respectively.

planning [4]. Recently, game-theoretic approaches were applied to the problem of mutliagent planning so that common plans or joint plans can be generated among self-interested agents [1, 2]. In particular, Brafman *et al.* formalized a mulitagent planning problem into a planning game which captures a rich class of planning scenarios [3]. However, these existing works on multiagent planning are based on either the assumption that all agents have common goals or the assumption that all agents must reveal their private information, such as goals, rewards, costs and/or utilities, to other agents or arbitrators. In contrast, our approach to multiagent planning is based on a bargaining mechanism, which assumes that goals, rewards and costs are private information and will not be revealed to any other agents or arbitrators. In fact, these pieces of information determine the bargaining power of an agent.

As future work, we will extend the present planning model to n-agent systems ($n > 2$). The main challenge of the extension is how to offer side-payment to each other agent in the situation of unknowing other agents' demands (obviously equal distribution does not work). Secondly, it is interesting to extend the current work to nondeterministic cases. This requires to redefine the solution concept and the CoAchieve algorithm in strong [6] or probabilistic style. Finally, more general mechanisms can be designed for multi-agent planning to deal with changing goals, incomplete information [9, 10], and reasoning agents [11, 12].

References

1. Ben Larbi, R., Konieczny, S., Marquis, P.: Extending Classical Planning to the Multi-agent Case: A Game-theoretic Approach. In: *ECSQARU-07*, 731–742 (2007)
2. Bowling, M., Jensen, R., Veloso, M.: A Formalization of Equilibria for Multiagent Planning. In: *IJCAI-03*, 1460–1462, 2003.
3. Brafman, I. R., Domshlak, C., Engel, Y., Tennenholtz, M.: Planning Games. In: *AAAI-09*, 73–78, 2009.
4. Brafman, I. R., Domshlak, C.: From One to Many: Planning for Loosely Coupled Multi-agent Systems. In: *ICAPS-08*, pp.28–35, 2008.
5. Brainov, S., Sandholm, T.: Power, Dependence and Stability in Multiagent Plans. In: *AAAI-99*, 11–16, 1999.
6. Cimatti, A., Pistore, M., Roveri, M., Traverso, P.: Weak, Strong, and Strong Cyclic Planning via Symbolic Model Checking. *Artificial Intelligence*. 147, 35–84, 2003.
7. Moses, Y., Tennenholtz, M.: Multi-entity Models. *Machine Intelligence*, 14, 63–88, 1995.
8. Muthoo, A., *Bargaining Theory with Applications*. Cambridge University Press, 1999.
9. Wei, H., Zhonghua, W., Yunfei, J., Lihua, W.: Observation Reduction for Strong Plans. In: *IJCAI-07*, 1930–1935, 2007.
10. Wei, H., Zhonghua, W., Yunfei, J., Hong, P.: Structured Plans and Observation Reduction for Plans with Contexts. In: *IJCAI-09*, 1721–1727, 2009.
11. Tran, S., Sakama, C.: Negotiation Using Logic Programming with Consistency Restoring Rules. In: *AAAI-09*, 930–935, 2009.
12. Zhang, D.: Reasoning about bargaining situations. In *AAAI-07*, 154–159, 2007.