

# Lagrangian Relaxation for Large-Scale Multi-Agent Planning\*

## (Extended Abstract)

Geoffrey J. Gordon<sup>†</sup>, Pradeep Varakantham<sup>‡</sup>, William Yeoh<sup>‡</sup>, Hoong Chuin Lau<sup>‡</sup>,  
Ajay S. Aravamudhan<sup>‡</sup> and Shih-Fen Cheng<sup>‡</sup>

<sup>†</sup>Machine Learning Department, Carnegie Mellon University, Pittsburgh, PA 15213, USA

<sup>‡</sup>School of Information Systems, Singapore Management University, Singapore 178902

<sup>†</sup>ggordon@cs.cmu.edu    <sup>‡</sup>{pradeepv,williamyeoh,hclau,ajaysa,sfcheng}@smu.edu.sg

### ABSTRACT

Multi-agent planning is a well-studied problem with applications in various areas. Due to computational constraints, existing research typically focuses either on unstructured domains with many agents, where we are content with heuristic solutions, or domains with small numbers of agents or special structure, where we can find provably near-optimal solutions. In contrast, here we focus on provably near-optimal solutions in domains with many agents, by exploiting *influence limits*. To that end, we make two key contributions: (a) an algorithm, based on Lagrangian relaxation and randomized rounding, for solving multi-agent planning problems represented as large mixed-integer programs; (b) a proof of convergence of our algorithm to a near-optimal solution.

### Categories and Subject Descriptors

I.2.11 [Artificial Intelligence]: Distributed AI

### General Terms

Algorithms, Experimentation

### Keywords

Multi-agent Planning, Lagrangian Relaxation

## 1. INTRODUCTION

Rapid progress in ubiquitous computing has enabled real-time delivery of contextualized information via devices (such as mobile phones and car navigation devices) over wide areas. As a result, a new kind of information service for mass user support is beginning to emerge. Examples include services that coordinate movements of first responders during a disaster rescue [1], movements of taxis in a fleet [3] and movements of visitors in leisure destinations (such as theme parks or world expositions). In these services, users are

\*This research is supported by the Singapore National Research Foundation under its International Research Centre @ Singapore Funding Initiative and administered by the IDM Programme Office.

Appears in: *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2012)*, Conitzer, Winikoff, Padgham, and van der Hoek (eds.), 4-8 June 2012, Valencia, Spain.

Copyright © 2012, International Foundation for Autonomous Agents and Multiagent Systems ([www.ifaamas.org](http://www.ifaamas.org)). All rights reserved.

typically represented by computational agents that perform real-time planning and adaptation. Designing coordination mechanisms that can govern these services in ways that meet global criteria such as fairness, revenue maximization, stability/convergence, and efficient resource utilization is a research challenge. Motivated by this challenge, we present an algorithm, based on Lagrangian relaxation and randomized rounding, for large-scale multi-agent planning problems. We prove convergence to an optimal solution as the number of agents increases; in fact, the quality of the solution actually *improves* as the problem size increases.

## 2. ILLUSTRATIVE DOMAIN

We motivate our work with a theme park crowd management problem, represented with a tuple  $\langle A, P, \{A(p_i)\}_1^n, \{d_{a_i}\}_1^k, \{U_i\}_1^n, H \rangle$ , where  $A = \{a_i\}_1^k$  is the set of attractions in the theme park;  $P = \{p_i\}_1^n$  is the set of patrons in the theme park;  $A(p_i) \subseteq A$  is the subset of attractions that patron  $p_i$  prefers to visit;  $d_{a_i}$  is the service rate of attraction  $a_i$ , that is, the number of patrons it can serve per time step;  $U_i$  is the utility function of patron  $p_i$ ; and  $H$  is the time horizon. The goal is to find the route  $\pi_i$  for each patron  $p_i$  such that the sum of utilities  $U_i(\pi_i)$  over all patrons is maximized.

## 3. MULTI-AGENT PLANNING PROBLEM

We represent the multi-agent planning problems as a large-scale mixed-integer program with special structure. This representation is very general, subsuming for example Markov decision processes, network flows, and graphical models such as influence diagrams, via reductions based on sampling scenarios [2]. Our chief assumptions are *factored structure*, the existence of *local planning subroutines*, and an *influence limit* for each agent. The efficiency of our algorithm will depend on the factored structure and the number and difficulty of local planning problems; our solution quality bounds will improve with more agents and tighter influence limits.

In more detail, we suppose that agent  $i$ 's plan is represented by a set of decision variables  $x_i \in \mathbb{R}^{n_i}$ , subject to *local* constraints  $A_i x_i = b_i$ ,  $x_i \in X_i$  and *local* costs  $c_i^\top x_i$ . The agents interact through *coupling* constraints  $\min f_j(y_j)$ , where  $y_j = \sum_{i=1}^n \ell_{ij}^\top x_i$  is resource consumption and  $f_j : \mathbb{R} \rightarrow \mathbb{R} \cup \{\infty\}$  is a closed proper convex function representing resource cost. The *global* planning problem is

Inputs: $c_i, \mathcal{C}_i, \ell_{ij}, f_j, \eta, T, \epsilon_j, \alpha_j^{\max}, \alpha_j^{\min}$	Outputs: $\bar{x}_i, \bar{\lambda}_j$
$\lambda_{ij} \leftarrow 0$	$j = 1 \dots m$
for $t \leftarrow 1, 2, \dots, T$	
$x_{it} \leftarrow \arg \max_x [c_i^\top x - \sum_{j=1}^m \lambda_{ij} \ell_{ij}^\top x]$	$i = 1 \dots n$
s.t. $x \in X_i, A_i x = b_i$	
$y_j \leftarrow \epsilon_j + \sum_{i=1}^n \ell_{ij}^\top x_{it}$	$j = 1 \dots m$
$z_j \leftarrow \arg \max_z [\lambda_j z - f_j(z)]$	$j = 1 \dots m$
$\lambda_{jt} \leftarrow \lambda_{j,t-1} + \frac{\eta}{\sqrt{t}} (y_j - z_j)$	$j = 1 \dots m$
$\lambda_{jt} \leftarrow \max(\alpha_j^{\min}, \min(\alpha_j^{\max}, \lambda_{jt}))$	$j = 1 \dots m$
$\bar{x}_i \leftarrow \frac{1}{t} \sum_{k=1}^t x_{ik}$	$i = 1 \dots n$
$\bar{\lambda}_j \leftarrow \frac{1}{t} \sum_{k=1}^t \lambda_{jk}$	$j = 1 \dots m$
round $\bar{x}_i$ to $x_i$ as described in text	$i = 1 \dots m$

**Figure 1: SLR Pseudocode**

therefore:

$$\begin{aligned} \max_x V_p(x) \text{ s.t. } A_i x_i = b_i, x_i \in X_i \quad & \forall i \\ V_p(x) = \sum_{i=1}^n c_i^\top x_i - \sum_{j=1}^m f_j \left( \sum_{i=1}^n \ell_{ij}^\top x_i \right) \end{aligned} \quad (1)$$

This problem is NP-hard and inapproximable; but, we can take advantage of a limit on the largest *influence* of any agent to solve it efficiently. More formally, we assume, first, that no agent controls a disproportionate share of the utility or resources: there is a constant  $U > 0$  such that

$$-\frac{U}{n} |V_p^*| \leq c_i^\top x_i \leq \frac{U}{n} |V_p^*| \quad (2)$$

$$-\frac{U}{n} |y_j^*| \leq \ell_{ij}^\top x_i \leq \frac{U}{n} |y_j^*| \quad (3)$$

for all  $i, j$ , and  $x_i \in X_i$ . Here  $V_p^*$  is the optimal value in Eq. 1 and  $y_j^*$  is the usage of resource  $j$  in some optimal solution. Second, we suppose that the optimization problem as a whole is well conditioned: suppose we redefine the consumption cost in Eq. 1 to be

$$f_j(\epsilon_j + \sum_{i=1}^n \ell_{ij}^\top x_i) \quad (4)$$

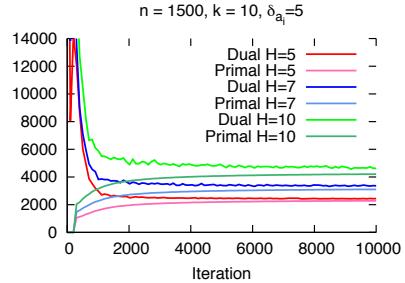
for some small  $\epsilon_j \geq 0$ . Let  $V_\epsilon^*$  be the optimal value of Eq. 1 in this case. Then, we assume that there exists an  $\epsilon_{\max} > 0$ , a  $\kappa > 0$ , and a  $\Delta > 0$  such that, whenever  $0 \leq \epsilon_j \leq \epsilon_{\max}$ ,

$$V_\epsilon^* \geq V_p^* - \kappa \sum_j \epsilon_j - \Delta/n \quad (5)$$

## 4. SLR ALGORITHM

Fig. 1 shows the *Subgradient Lagrangian Relaxation* (SLR) algorithm. Inputs are the problem parameters  $c_i, A_i, b_i, \ell_{ij}, f_j$  (as in Eq. 1); learning rate  $\eta > 0$  and number of iterations  $T$ ; bounds  $\alpha_j^{\max}$  and  $\alpha_j^{\min}$  on the slope of  $f_j$ ; and target margins  $\epsilon_j$ . Outputs are expected plans  $\bar{x}_i$  for each agent, as well as prices  $\bar{\lambda}_j$  for each resource; the latter can be used to check convergence.

We cannot directly execute the final aggregated policy  $\bar{x}$ : since the domains  $X_i$  are typically non-convex, averaging feasible solutions does not typically yield a feasible solution. To remedy this problem, we use randomized rounding: each agent independently picks a random locally-feasible policy  $x_i$  according to a distribution which makes  $\mathbb{E}(x_i) = \bar{x}_i$ . (One



**Figure 2: Experimental Results of SLR**

such distribution is the uniform distribution over  $x_{it}$  for  $t = 1 \dots T$ .) To ensure feasibility, we set the margin  $\epsilon_j$  to trade off total predicted utility against the possibility of violating resource constraints.

The following theorem shows that we can set the parameters of SLR to guarantee that rounding yields a high-quality plan with high probability, and that, with these parameters, the expected runtime of SLR will be a low-order polynomial in the problem size. In particular, we can pick any desired failure probability, say  $\delta = 0.01$ , and a decreasing convergence tolerance, say  $\gamma = 1/\sqrt{n}$ . Then, we can set  $\epsilon = \Theta(\frac{\ln 1/\delta}{\sqrt{n}})$ ,  $\eta = \Theta(1)$ , and  $T = \Theta(\gamma^{-2}) = \Theta(n)$  to achieve low error, polynomial runtime, and high success probability. (And, we can make the success probability arbitrarily close to 1 by repeating the rounding step.)

**THEOREM 1.** Suppose influence limits are guaranteed by Eqs. 2–5. Fix  $\epsilon \leq \epsilon_{\max}$ , set  $\epsilon_j = \epsilon$  for all  $j$ , and run SLR (Fig. 1) to some tolerance  $\gamma$ . Let each agent randomize independently with  $\mathbb{E}(x_i) = \bar{x}_i$ . Set

$$\delta = e^{-n\epsilon^2/2U^2|V_p^*|^2} + me^{-n\epsilon^2/2U^2|y_j^*|^2}$$

Then, with probability at least  $1 - \delta$ ,

$$V_p(x) \geq V_p^* - \Delta/n - (\kappa m + 1)\epsilon - \gamma$$

## 5. EXPERIMENTAL RESULTS

For our experiments, using the notation from above, we set  $|A| = 10$ ,  $\delta_{a_i} = 5$  for all attractions,  $n = 1500$ ,  $k = 10$ , and vary  $H$  from 5 to 10. Fig. 2 shows a set of representative results, where we plot the primal and dual values (from Eq. 1 and its dual) across iterations. The primal and dual values increase with  $H$ , since higher  $H$  lets some patrons visit attractions that they would otherwise have skipped. Convergence is fast for all problems, and the duality gap is small, indicating that we have reached a near-optimal solution in this large-scale problem instance.

## 6. REFERENCES

- [1] S. Koenig, P. Keskinocak, and C. Tovey. Progress on agent coordination with cooperative auctions [Senior Member Paper]. In *Proceedings of AAAI*, pages 1713–1717, 2010.
- [2] A. Ng and M. Jordan. PEGASUS: A policy search method for large MDPs and POMDPs. In *Proceedings of UAI*, pages 406–415, 2000.
- [3] P. Varakantham, S.-F. Cheng, and T. D. Nguyen. Decentralized decision support for an agent population in dynamic and uncertain domains. In *Proceedings of AAMAS*, pages 1147–1148, 2011.