A System for Multi-Agent Coordination in Uncertain Environments

Lucian Vlad Lita, Jamieson Schulte, Sebastian Thrun
$\{llita, jschulte, thrun\}$@cs.cmu.edu
Carnegie Mellon University

1. INTRODUCTION

This paper present a multi-agent architecture for coordinating large numbers of mobile agents (e.g. robots) cooperating in uncertain environments. In particular, the Canadian Traveler Problem (CTP) is the problem of finding a shortest path to a goal location in a graph, where individual edges of the graph might or might not be traversable[1]. The agent has an initial probabilistic knowledge about the states of the edges. Whether or not an edge is traversable can only be found out by moving there. Hence, an optimal solution to a CTP is a contingency plan, which offers alternative routes if edges are not available. Finding an optimal contingency plan is known to be NP-hard. We focus on the multi-agent CTP, which involves multiple agents attempting to reach multiple target locations. Finding an optimal solution is even harder, since the space of actions at each point in time is exponential in the number of agents.

Our multi-agent architecture approaches the above mentioned set of intractable problems in an efficient, real-time manner. The architecture supports a large number of mobile, goal-driven information agents that strive to maximize their reward for reaching goals. These agents are coordinated at a higher level by dispatcher agents whose purpose is to maximize the total reward accumulated over time.

Extensive experimental results have been obtained in the context of natural disaster relief. Our experiments have been carried out in a realistic simulation of Honduras after Hurricane Mitch destroyed most of the country's infrastructure.

2. INFORMATION AGENTS

Information sharing, goal driven agents operate according to plans limited by their computational power and response (action) time. Each agent generates a set of contingency plans based on its own observed history as well as observations shared by its peers [3]. These observations inform agents about the true state of the world which is otherwise defined only probabilistically. From the set of constructed plans the agent selects the best plan and then acts according to it, sharing its discoveries with other agents. During execution, an agent reevaluates its plan and re-plans if useful new information has been discovered.

Traditionally, planning problems in deterministic spaces involve finding a path from an agent to its goal. A commonly used algorithm is A* which uses an optimistic measure of goal distance to efficiently find the optimal path. However, under uncertainty agents have to plan for contingencies and cannot ignore the uncertainty inherent in the world [2]. Belief A* (BA*), is an algorithm based on A* which handles planning under uncertainty. BA* searches in the space of all possible plans and produces the best contingency plan over the belief space found under specified time and computation restrictions.

The BA* algorithm uses a future discounted reward (FDR) heuristic in order to direct the search toward states with higher expected reward by prioritizing the expansion of belief states during planning The estimated utility of each belief state s given the observed history h is measured by the expected FDR:

$$\widehat{R}(s,h) = P(s|h) \cdot e^{-\alpha(t_o(s)+t_u(s))} \tag{1}$$

where alpha is a decay parameter, $t_o(s)$ time the agent would take to reach belief state s following some trajectory in h, $t_u(s)$ is a lower bound on the time the agent will take to reach the goal from s, and $P(s|h)$ is the probability of reaching state s following the trajectory in h. Response time and computational resources dictate how much BA* expands belief states during planning.

Each information agent evaluates the quality of a potential plan. It computes the expected reward upon reaching the goal over the set of potential future histories leading to the goal. In order to choose the best available plan, the information agent maximizes the expected future discounted reward over the spectrum of generated plans. The advantage of the BA* algorithm is that it tends to generate relevant belief states towards reaching the goal.

3. DISPATCHER AGENTS

Under real-time requirements, full communication of agents' plans is impractical. On the other hand, agent independence is very restrictive. We introduce *dispatcher agents* - a dynamic, fast and scalable mechanism that coordinates the assignment of agents to goals. Dispatcher agents attempt to maximize the total reward attained by the information agents, using only an estimate of individual agent reward expectations

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

AGENTS'01, May 28-June 1, 2001, Montréal, Quebec, Canada.

Copyright 2001 ACM 1-58113-326-X/01/0005 ...$5.00.

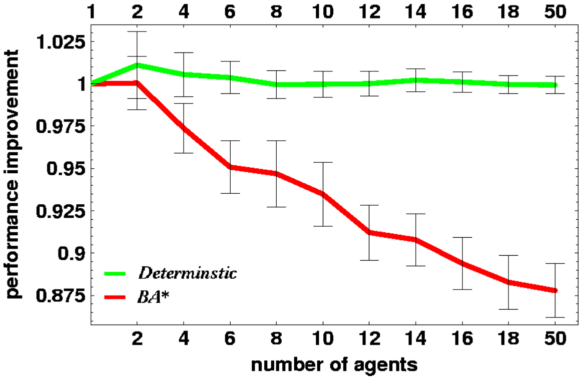

Figure 1: Hurricane Mitch scenario: BA* based Information Agents outperform Deterministic Agents.

Optimal dispatching is a combinatorial problem, intractable for large numbers of agents. Instead we adopt a greedy dispatcher that allows agents to communicate. The dispatcher avoids the redundant computation of expected reward. Instead, the dispatcher uses an expected FDR measure computed in parallel by information agents. Based on the available FDR the dispatcher assigns agents to goals.

An architecture requirement is to minimize on-line agent communication so that the architecture remains scalable. However, the architecture must also allow enough communication so that agents benefit from knowledge sharing. The greedy expected FDR dispatcher allows these restrictions to be practical. The FDR dispatcher absorbs planning information using little communication while benefiting from the parallelism and coordination inherent in the multi-agent architecture.

4. EXPERIMENTS

Experiments show that our multi-agent architecture supports a large number of agents that plan and act efficiently, in real-time in realistic uncertain environments.

We analyzed the performance of the system using different disaster scenarios - different state transition probabilities. We varied the world instance as a sampling of random variables representing more than 300 road segments and having the distributions dictated by each specific scenario. For these world instances we varied the number of agents and goals as well as their physical locations. The simulations were performed on up to 100 information agents and we found that a single dispatcher managed the information agents well.

In most situations the BA* based algorithm took advantage of state transition probability information in order to weigh the impact of low cost, but high risk actions. Using realistic scenarios, we compared information agents using a BA* heuristic versus information agents using a deterministic, shortest path approach (Figure 1). *Performance improvement* is the ratio of *naive* agents' average reward to that of the agents under evaluation. The *naive* agents use simple shortest-path planning, do not communicate, and use a random dispatcher.

We limited planning time through the BA*'s belief state expansion so that a tradeoff would exist between response time and accuracy. The information agents develop high-reward plans in our large scale Honduras environment in under 5 seconds on a 500MHz PC.

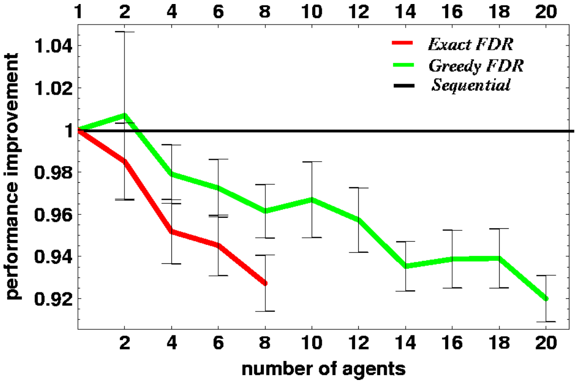

Figure 2: Greedy Dispatcher Agents outperform Sequential Dispatchers. The exact (exhaustive) FDR is also shown for comparison purposes

In a large-scale experiment, we averaged the performance of 100 information agents over a set of instances of a scenario with multiple random sampling the physical locations of goals and agents. In a Hurricane Mitch scenario BA* based agents show a 20% performance improvement over deterministic agents.

In our dispatching simulations, the sequential dispatcher matches the agents and goals as they are appear in queue. The greedy expected FDR dispatcher requires the information agents to plan for every goal, and then selects a set of pairings corresponding to the best joint set of plans. The results shown in Figure 2 allow agents to communicate. We found that dispatching overhead is insignificant compared to the amount of planning information agents are required to perform.

5. DISCUSSION

Experiments show that our multi-agent architecture, when applied to large number of coordinated information agents, is both practical and efficient. Significant improvement is seen in simulated realistic disaster scenarios, while simulations of environments with little uncertainty yield results very close to those of deterministic planners.

Future work will focus on improving dispatcher agents by reducing the joint probability of failure in information agents. We will also address re-dispatching and additive dispatching as well as modeling an ever-changing environment where observations do not preserve their information content over time

6. REFERENCES

[1] A. Bar-Noy and B. Schieber. The Canadian traveler problem. In *Proc. of the 2nd ACM-SIAM Symposium on Discrete Algorithms* pages 261-270, 1991.

[2] D.M. Blai and L.P. Kaelbling. *Shortest paths in a dynamic uncertain domain.* In *IJCAI Workshop on Adaptive Spatial Representations of Dynamic Environments*, Menlo Park, CA, 1999.

[3] B. Bonet and H. Geffner. *Planning with incomplete information as heuristic search in belief space. Proc. of the 5th Intl. Conf. on AIPS*, April 2000.