Simulating UAV Surveillance for Analyzing Impact of Commitments in Multi-Agent Systems

David C. Han, The University of Texas at Austin
Suzanne K. Barber, The University of Texas at Austin

ABSTRACT

Autonomous agents, by definition, have the freedom to make their own decisions. Rational agents execute actions that are in their “best interests” according to their desires. Action selection is complicated due to uncertainty when operating in a dynamic environment or where other agents can also influence the environment. This paper presents an action selection framework and algorithms that are rational with respect to multiple desires and responsive to changing desires. Coordination is layered on top of this framework by describing and analyzing how commitments affect the agents’ desires in their action selection models. Commitments may have a positive or a negative effect on an agent’s ability to satisfy its desires. This research uses simulation in the domain of UAV surveillance to experimentally explore the balance between under-commitment and over-commitment.

Keywords: Commitment, Coordination, Desire, Multi-Agent Systems, Planning, Unmanned Aerial Vehicle (UAV) Surveillance

INTRODUCTION

Rational agents execute actions that are in their “best interests” according to their desires. Rational action selection, already a difficult problem (in both representation and reasoning), is complicated by dynamic environments. This includes, in the case of multi-agent systems, dynamism resulting from interactions with other agents. Dynamism introduces uncertainty about the effects of actions, obfuscating the true value of each action. In the case of dynamism introduced through agent interactions, this may be somewhat mitigated if an agent can to predict future effects based on knowledge about the agent interactions that have occurred. In ad hoc agent systems, where the agents were not designed to plan together, this knowledge is gained through the announcement of agent intentions, resulting in social commitments among the agents. Commitments direct and constrain an agent’s future actions.

DOI: 10.4018/jats.2012010101
There has been much work on studying the semantics of commitment (Jennings, 1993). The objective of this research is to analyze the pragmatics of commitment in agent action and multi-agent coordination for the purposes of furthering our understanding of how to design and program coordinating agents. The approach taken is to explicitly model and reason about agent desires, specifically calculating the change in expected utility due to commitments. Commitments may have a negative impact on an agent’s ability to satisfy its desires. An accurate valuation of commitment is needed to determine if it is rational for an agent to commit to a goal or not. The domain of Unmanned Aerial Vehicle (UAV) surveillance is used as the representative domain for a class of problems addressed in this research. Simulation of this UAV domain is used to experimentally test this research.

**UAV Surveillance**

The UAV Surveillance domain consists of a set of UAVs that are tasked with surveying a number of moving targets, maintaining an accurate model of the positions of those targets, and finally “servicing” those targets. Servicing is an abstraction that can represent a number of different activities, such as applying specific sensors to the target or firing upon and destroying the target. In order to service a target, a UAV is required to visit the target’s location. By treating each target as a separate goal, the existence of numerous goals allows demonstration of goal tradeoffs and complex goal-level interactions among agents.

At its core, UAV Surveillance is a navigation domain. An agent is associated with each UAV for the purpose of controlling the movement of that UAV. It is assumed that the agents hold an accurate model of their own locations against an arbitrary, but static, global coordinate system. The agents maintain models, or beliefs, of the last known location of each of the targets. To avoid synchronization issues among the agents, the agents are given access to a global clock. Upon scanning a target, agents can record the time at which each target’s position was last checked. It is assumed that each target is uniquely identifiable to avoid recognition issues in target tracking.

Typically, navigation is modeled as a cost-to-move domain. In cost-to-move problems, each action the agent executes incurs some cost $c < 0$ as part of the reward structure. The cost represents resource usage by the UAVs as they move from one location to another. This provides incentive for the agent to reach its goal states with the minimal amount of movement actions. The desires of the agent in navigation domains are represented by the reward structure. For example, $R(s) \geq 0$ for the agent’s goal location, $s$, and $R(s) = 0$ otherwise.

Typically for planning approaches, if the agent is unable to find a path that includes all goal locations, it returns failure. In UAV Target Tracking, a partial solution visiting some subset of targets should be counted as a better result than not finding a solution at all. In the UAV domain simulation, the agents are rewarded independently for each individual target they service. This provides incentive for partial plans and partial plan execution. This type of over-subscribed problem is solved through partial satisfaction (Van Den Briel, 2004).

The UAVs are resource constrained (in time) forcing the agents to consider the costs and rewards for servicing each target when making their target selection decisions. If a target is unprofitable for an agent, the agent has the autonomy to say “no” and decide not to work on that particular target. Costs are based on distance traveled by the UAV, with closer distances incurring less cost. Hence, it is important that the agent consider not only target selection but also target ordering.

This domain is similar to both the Vehicle Routing Problem (VRP) (Vokřínek, 2010) and the Continuous Area Sweeping Task (Ahmadi, 2006). Vokřínek approaches the VRP by casting the problem as a combination of the Multiple Traveling Salesman Problem (MTSP) and the Bin Packing Problem (BPP). The main difference between this domain and both the VRP and the sweeping problem is the dynamic nature; targets are added and removed. For sweep-
Reward Shaping and Mixed Resolution Function Approximation
www.igi-global.com/chapter/reward-shaping-mixed-resolution-function/49358?camid=4v1a