Utilizing Internal State in Multi-Robot Coordination Tasks

Chris Jones and **Maja J Matarić**
Computer Science Department, University of Southern California
941 West 37th Place, Los Angeles, CA 90089-0781 USA
http://robotics.usc.edu/interaction/research/projects/chris-project.php
{cvjones|mataric}@usc.edu

Introduction and Related Work

The success of a task-achieving multi-robot system (MRS) depends on effective coordination mechanisms to mediate the robots' interactions in such a way that a given task is achieved. In the MRS community, many elegant coordination mechanisms have been empirically demonstrated. However, there is a lack of systematic procedures for the synthesis of coordinated MRS. In this paper, we address this issue by presenting a principled framework suitable for describing and reasoning about the intertwined entities involved in any task-achieving MRS – the task environment, task definition, and the capabilities of the robots themselves. Using this framework, we present a systematic procedure by which to synthesize controllers for robots in a MRS such that a given sequential task is correctly executed. The MRS is composed of homogeneous robots capable of maintaining internal state but not capable of direct inter-robot communication. This systematic approach to the synthesis of coordinated MRS permits formal identification of the benefits and limitations of MRS composed of robots maintaining internal state and when other types of controllers may become necessary, such as those using communication.

The most relevant related work includes work on finding optimal policies in partially observable stochastic domains (Cassandra, Kaelbling, & Littman 1994). In MRS, the study of information invariants (Donald 1995) addresses the problem of determining information requirements for performing robot tasks. The work presented here builds upon our previous results on automated synthesis of controllers using internal state (Jones & Matarić 2003) by addressing the issue of uncertainty in sensing and action. Our work presented in (Jones & Matarić 2004) presents results on the synthesis of stateless controllers which are communicative but stateless. In the multi-agent systems community, the issues addressed in (Pynadath & Tambe 2002) and (Stone & Veloso 1999) are related to ours in that they are systematically studying the role different control characteristics play in coordinated systems.

Definitions and Notation

The *world* is the domain in which the MRS performs a given task. We assume the world is Markovian and the state is an element of the finite set S of all possible states. An action a performed in the world by a single robot r is drawn from the finite set A of all possible actions. An *observation* x made by robot r, drawn from the finite set of all observations X, consists of accessible information external to the robot and formally represents a subset of the world state. Given a world state s at time t, a robot r making observation x and executing action a, and a world state s' at time $t+1$, we define a probabilistic world state transition function as $P(s, x, a, s') = Pr(S^{t+1} = s'|S^t = s, X_r^t = x, A_r^t = a)$. We note that the world state transition function involves an observation because the tasks we consider are spatial in nature and the physical location where an action is performed is just as important as the action itself. We define a *task*, assumed to be Markovian, as a set of n ordered world states $T_s = \{s_0, s_1, ..., s_n\}$ which must be progressed through in sequence. We assume the initial state of the world is s_0. We define *correct task execution* to be the case where for all task states $s_i \in T_s$, $i < n$ the only actions executed by any robot are those that transition the world state to s_{i+1}. Once the world state is $s_n \in T_s$ the task is terminated. Therefore, we define an observation and action pair for a robot, x and a, to be correct for task state s_i if $P(s_i, x, a, s_{i+1}) > 0$. We assume that an observation x and action a cannot be correct for more than one task state. The probabilistic *observation function* $O(s, x) = Pr(X_r^t = x|S^t = s)$ gives the probability observation x will be made in state s by a robot r. Furthermore, we assume that an observation x may only be made at one physical location in the world in a state s. A robot's internal state value m at any time is a member of the finite set $M = \{m_0, m_1, ..., m_p\}$. Two probabilistic functions define a robot r's behavior in the world, known collectively as the robot's *controller*. The controller is comprised of an *action function* $A(x, m, a) = Pr(A_r^t = a|X_r^t = x, M_r^t = m)$ and an *internal state transition function* $L(m, x, m') = Pr(M_r^{t+1} = m'|X_r^t = x, M_r^t = m)$. Although the controller is modeled with probabilistic functions for generality, in this paper these functions will always be either 0 or 1.

Principled Controller Synthesis

Leveraging the formalism presented above, Figure 1 presents a controller synthesis procedure consisting of four phases. The first phase (lines 2-3) synthesizes an action function for a baseline stateless controller which contains a rule for each action in the task, but there is no guarantee the task will be performed in the correct sequence.

The second phase (lines 4-6) initializes some relevant variables. The variable $X_a(s_i)$ contains the set of all observations x for which there exists an action a such that x and a are correct for state s_i, for all $s_i \in T_s$. The set $V_a(s_i)$ will contain the index of the internal state value (i.e., 0 represents m_0) that a robot will need to have in order to execute an action in state s_i, for all $s_i \in T_s$. Initially, all values in V_a are assigned the same internal state value m_0 – this is equivalent to not using any internal state at all. Lastly, the set $O_a(s_i)$ will contain the observation, if any, that will be used to transition the internal state value in state s_i, for all $s_i \in T_s$. Initially, all values in O_a are $NULL$ as no internal state transitions are defined at this point.

The third phase (lines 7-12) identifies situations in which internal state can be used to improve coordination and assigns appropriate values to the sets V_a and O_a. The basis for determining when internal state can be used to improve coordination is in identifying task states where an observation x in $X_a(s_j)$, where x and some action a are correct for s_j, can also be made in some earlier task state s_i for which x and a are not correct. We note that our synthesis method does not deal with the situation where there exists a task state s_p which occurs *later* than s_j and $O(s_p, x) > 0$, where x and a are not correct for s_p.

The fourth phase (lines 13-18) synthesizes the final controller by augmenting the stateless controller synthesized in the first phase. This is accomplished by adding the internal state transition function and appropriately modifying the action function such that an action is not executed unless the robot's current internal state value is appropriate. The internal state transition function is constructed (lines 14-15) by mapping the internal state value $m_{V_a(s_i)-1}$ and observation $O_a(s_i)$ to the next internal state value of $m_{V_a(s_i)}$, for all $s_i \in T_s$. The action function is modified (lines 16-17) such that for each rule of the action function $A(x, m_0, a) = 1$ where x and a are correct for a state s_i is modified to become $A(x, m_{V_a(s_i)}, a) = 1$, where $m_{V_a(s_i)}$ is the required internal state value for task state s_i as determined in the third phase. All probabilities not explicitly declared are 0.

(1) procedure Synthesize_Controller()
(2) for all $s_i \in T_x, a \in A, x \in X(O(s_i, x) > 0 \wedge P(s_i, x, a, s_{i+1}) > 0)$ do
(3) $A(x, m_0, a) = 1$

(4) for all $s_i \in T_s$ do
(5) $X_a(s_i) = \{x_0, x_1, \ldots, x_n\}$ s.t. $\forall x \in X_a(s_i) \exists a(O(s_i, x) > 0$ $\wedge A(x, m_0, a) = 1)$
(6) $V_a(s_i) = 0; O_a(s_i) = NULL$

(7) for all $s_i, s_j \in T_s (i < j)$ do
(8) if $\exists x \in X_a(s_j)(O(s_i, x) > 0)$ then
(9) for $s_k = s_j$ downto s_{i+1} do
(10) if $\exists z \nexists s_u (O(s_k, z) > 0 \wedge u < (i+1) \wedge O(s_u, z) > 0)$ then
(11) $O_a(s_k) = z$
(12) $\forall s_w (w >= k) \rightarrow V_a(s_w) = V_a(s_w) + 1$

(13) for all $s_i \in T_s$ do
(14) if $O_a(s_i) \neq NULL$ then
(15) $L(m_{V_a(s_i)-1}, O_a(s_i), m_{V_a(s_i)}) = 1$
(16) for all $x \in X_a(s_i), a \in A(A(x, m_0, a) = 1)$ do
(17) $A(x, m_0, a) = 0; A(x, m_{V_a(s_i)}, a) = 1$
(18) end procedure Synthesize_Controller

Figure 1: Controller synthesis procedure.

Experimental Validation

We experimentally validate our controller synthesis procedure in a multi-robot construction domain, through extensive physically-realistic simulations and on a limited set of real-robot experiments. The construction task requires the sequential placement of a series of cubic colored bricks into a planar structure. Each robot's sensing capabilities limit it to sensing only a limited profile of the current construction at any given time. Additional details on the multi-robot construction domain can be found in (Jones & Matarić 2004).

We performed 300 experimental trials in simulation for a specific construction task made up of 7 bricks. The synthesized controller correctly executed the construction task in 31.5% of the trials. We note that there exists significant uncertainty in sensing and imperfect actions, and in the absence of such uncertainties, the synthesized controller would be gauranteed to correctly execute the task. For comparison, the stateless controller from the first phase of the synthesis procedure resulted in only 0.9% of the trials resulting in correct task execution. A limited number of successful real-robot experiments were performed to verify that any assumptions we made are reasonable and realistic. The experiments were also performed to show that our formalism and synthesis method are not merely abstract concepts but have been successfully implemented on real systems and capture the difficult issues involved in real-world embodied MRS.

References

Cassandra, A.; Kaelbling, L. P.; and Littman, M. 1994. Acting optimally in partially observable stochastic domains. In *Proceedings of the Twelfth National Conference on Artificial Intelligence*, 1023–1028.

Donald, B. R. 1995. Information invariants in robotics. *Artifical Intelligence* 72(1–2):217–304.

Jones, C., and Matarić, M. 2003. Towards a multi-robot coordinaton formalism. In *Proc. of the Workshop on the Mathematics and Algorithms of Social Insects*, 60–67.

Jones, C., and Matarić, M. 2004. Communication in multi-robot coordination. Technical report, University of Southern California Center for Robotics and Embedded Systems, CRES-04-001.

Pynadath, D., and Tambe, M. 2002. Multiagent teamwork: Analyzing the optimality and complexity of key theories and models. In *International Joint Conference on Autonomous Agents and Multi-Agent Systems*, 873–880.

Stone, P., and Veloso, M. 1999. Task decomposition, dynamic role assignment, and low-bandwidth communication for real-time strategic teamwork. *Artificial Intelligence* 110(2):241–273.