

# A case study for life-long learning and adaptation in cooperative robot teams

Lynne E. Parker

Center for Engineering Science Advanced Research  
Computer Science and Mathematics Division  
Oak Ridge National Laboratory, P.O. Box 2008, Oak Ridge, TN 37831-6355 USA

## ABSTRACT

While considerable progress has been made in recent years toward the development of multi-robot teams, much work remains to be done before these teams are used widely in real-world applications. Two particular needs toward this end are the development of mechanisms that enable robot teams to generate cooperative behaviors on their own, and the development of techniques that allow these teams to autonomously adapt their behavior over time as the environment or the robot team changes. This paper proposes the use of the Cooperative Multi-Robot Observation of Multiple Moving Targets (CMOMMT) application as a rich domain for studying the issues of multi-robot learning and adaptation. After discussing the need for learning and adaptation in multi-robot teams, this paper describes the CMOMMT application and its relevance to multi-robot learning. We discuss the results of the previously-developed, hand-generated algorithm for CMOMMT and the potential for learning that was discovered from the hand-generated approach. We then describe the early work that has been done (by us and others) to generate multi-robot learning techniques for the CMOMMT application, as well as our ongoing research to develop approaches that give performance as good, or better, than the hand-generated approach. The ultimate goal of this research is to develop techniques for multi-robot learning and adaptation in the CMOMMT application domain that will generalize to cooperative robot applications in other domains, thus making the practical use of multi-robot teams in a wide variety of real-world applications much closer to reality.

**Keywords:** Multi-robot learning, adaptation, multi-robot cooperation, robot teams

## 1. INTRODUCTION

Research in multi-robot cooperation has grown significantly in recent years. While the growth of this topic area is due in part to a pure scientific interest in teams of autonomous robots, much of the growth is due to the increasing realization by the user community that teams of robots may provide solutions to difficult problems that previously were untenable. Certainly, it has been shown (e.g., in<sup>1</sup> and elsewhere) that multi-robot teams can increase the reliability, flexibility, robustness, and efficiency of automated solutions by taking advantage of the redundancy and parallelism of multiple team members.

However, before multi-robot teams will ever become widely used in practice, we believe that advances must be made in two important areas: (1) the development of automated mechanisms that enable the cooperative robots to be initially programmed much more easily than is currently possible, and (2) the development of automated techniques that enable robot team members to automatically adapt their actions over time in response to changes in their environment or in the robot team itself. With the current state of the art, the implementation of cooperative behaviors on physical robot teams requires expert behavior programming and experimentation, followed by extensive tuning and revision of the cooperative control algorithms. It is unlikely that a significant real-world impact of cooperative robot teams will occur as long as the current level of effort is required to implement these systems.

Researchers have recognized that an approach with more potential for the development of cooperative control mechanisms is autonomous learning. Hence, much current work is ongoing in the field of multi-agent learning

---

Other author information: Email: parkerle@ornl.gov; Telephone: (423) 241-4959; Fax: (423) 574-0405. This research is sponsored by the Engineering Research Program, Office of Science, of the U.S. Dept. of Energy. This article has been authored by a contractor of the U.S. government under Contract DE-AC05-96OR22464. Accordingly, the U.S. Government retains a nonexclusive, royalty-free license to publish or reproduce the published form of this contribution, or allow others to do so, for U.S. Government purposes.

(e.g.,<sup>2</sup>). The types of applications that are typically studied vary considerably in their characteristics. Some of the applications include air fleet control,<sup>3</sup> predator/prey,<sup>4-6</sup> and multi-robot soccer.<sup>7,8</sup> Of these previous application domains, however, only the multi-robot soccer domain addresses the real-world complexities of embodied robotics, such as noisy and inaccurate sensors and effectors in a dynamic environment that is poorly modeled.

In this paper, we propose the use of a second, new application domain for embodied multi-robot learning and adaptation – the Cooperative Multi-robot Observation of Multiple Moving Targets (CMOMMT) application. Like the multi-robot soccer domain, the CMOMMT application has a number of characteristics that make it suitable for study as a real-world application of multi-agent systems.

The generally-accepted motivations<sup>9</sup> for incorporating learning and adaptation into autonomous robots are to ease the difficulty of the initial robot programming and to enable the robots to change their performance over time as

usually distinguished from multi-robot adaptation by the degree to which new behaviors and processes are generated. Typically, in multi-robot learning, new behaviors or behavior sequences are generated or functions are learned, thus giving the robot team radically new capabilities. Frequently, the learning takes place in an initial phase, where performance during learning is not of importance. In multi-robot adaptation, the robot team exercises a control policy that gives reasonable results for the initial situation, and which the team is able to gradually change during the application to improve performance over time. The emphasis in multi-robot adaptation is the ability to change its control policy on-line while the team is performing its mission, in response to changes in the environment or the robot team.

An important goal in the development of autonomous robot systems is enabling robots to perform their tasks over a long period of time without human supervision. So-called “lifelong” robot systems must be capable of dealing with dynamic changes that will inevitably occur over time, such as changes in the environment or incremental variations in their own performance capabilities. Environmental changes may occur due to shifts in the weather, time of day daylight and temperature variation, the effect of the robot system itself in performing its application, or other causes external to the robot system. Incremental variations in robot performance capabilities may occur as a

performance of a particular task.

The ability to adapt to these types of dynamic changes is especially important in multi-robot applications, since

the environment drift further and further from the original state. Instead, a successful lifelong multi-robot team will adapt to changes in robot team member capabilities, robot team composition, mission requirements, and the environmental state.

One important consequence of dynamic changes in lifelong multi-robot systems is that continuing drift in individual robot capabilities creates a team of heterogeneous robots, even if the original team was designed to be homogeneous. This may actually be true at the outset of the application. Experienced roboticists are aware that several copies of the same model of robot, even direct from the manufacturing line, can vary in capabilities due to differences in sensor tuning, etc. Over time, the minor initial differences among robots will grow, since sensors and effectors may degrade or break on certain robots, or, for learning robots, some robot team members may learn faster than others. Thus, mechanisms for generating lifelong multi-robot teams must necessarily deal with the issue of heterogeneity among robot team members.

Heterogeneity in multi-robot teams presents a particular challenge to efficient autonomous control when overlap in team member capabilities occurs. Overlap in team member capabilities means that more than one robot may be able to perform a given task, but with different levels of efficiency. In these cases, the robots must continually determine which individual on the team is currently the best suited for a given task in the application. These types of decisions are usually not easy to make, especially when the multi-robot team control is distributed across all

team members. Thus, the multi-robot team control mechanism must have some effective means of distributing tasks so that an acceptable level of efficiency is achieved without sacrificing the desirable features of fault tolerance and robustness.

### 3. THE PROPOSED LEARNING APPLICATION: CMOMMT

The application domain that we propose to use as a multi-robot learning testbed is the problem we entitle *Cooperative Multi-Robot Observation of Multiple Moving Targets* (CMOMMT). This problem is defined as follows. Given:

- $\mathcal{S}$  : a two-dimensional, bounded, enclosed spatial region
- $\mathcal{V}$  : a team of  $m$  robot vehicles,  $v_i, i = 1, 2, \dots, m$ , with  $360^\circ$  field of view observation sensors that are noisy and of limited range
- $\mathcal{O}(t)$  : a set of  $n$  targets,  $o_j(t), j = 1, 2, \dots, n$ , such that target  $o_j(t)$  is located within region  $\mathcal{S}$  at time  $t$

We say that a robot,  $v_i$ , is *observing* a target when the target is within  $v_i$ 's *sensing range* (defined explicitly below).

Define an  $m \times n$  matrix  $B(t)$ , as follows:

$$B(t) = [b_{ij}(t)]_{m \times n} \text{ such that } b_{ij}(t) = \begin{cases} 1 & \text{if robot } v_i \text{ is observing target } o_j(t) \text{ in } \mathcal{S} \text{ at time } t \\ 0 & \text{otherwise} \end{cases}$$

Then, the goal is to develop an algorithm, which we call *A-CMOMMT*, that maximizes the following metric  $A$ :

$$A = \sum_{t=1}^T \sum_{j=1}^n \frac{g(B(t), j)}{T}$$

where:

$$g(B(t), j) = \begin{cases} 1 & \text{if there exists an } i \text{ such that } b_{ij}(t) = 1 \\ 0 & \text{otherwise} \end{cases}$$

In other words, the goal of the robots is to maximize the average number of targets in  $\mathcal{S}$  that are being observed by at least one robot throughout the mission that is of length  $T$  time units. Note that we do not assume that the membership of  $\mathcal{O}(t)$  is known in advance. Also note that when evaluating adaptive algorithms that improve their performance over time, this metric would be changed to sum over a recent time window, rather than the entire

In addressing this problem, we assume the following:

Define *sensor\_coverage*( $v_i$ ) as the region visible to robot  $v_i$ 's observation sensors, for  $v_i \in \mathcal{V}$ . Then we assume that, in general, the maximum region covered by the observation sensors of the robot team is much less than the total region to be observed. That is,

$$\bigcup_{v_i \in \mathcal{V}} \text{sensor\_coverage}(v_i) \ll \mathcal{S}.$$

This implies that fixed robot sensing locations or sensing paths will not be adequate in general, and instead, the robots must move dynamically as targets appear in order to maintain their target observations and to maximize the coverage.

In some situations, the observation sensor on each robot may be directional (e.g., a traditional CCD camera), and can only be used to observe targets within that sensor's field of view. However, new sensory technology, such as the recently prototyped omni-directional cameras (e.g., by SRI, CMU/Columbia, etc.), make it feasible in future implementations for a robot to be able to observe targets around its entire  $360^\circ$  field of view.

#### 4. CMOMMT AS A TESTBED FOR LEARNING AND ADAPTATION

CMOMMT offers a rich testbed for research in multi-robot cooperation, learning, and adaptation because it is an inherently cooperative task. That is, the utility of the action of one robot is dependent upon the current actions of the other team members. In addition, many variations on the dynamic, distributed sensory coverage problem are possible, making the CMOMMT problem arbitrarily more difficult. For example, the relative numbers and speeds of the robots and the targets to be tracked can vary, the availability of inter-robot communication can vary, the robots can differ in their sensing and movement capabilities, the terrain may be either enclosed or have entrances that allow objects to enter and exit the area of interest, and so forth. Many other subproblems can also be addressed, including the physical tracking of targets (e.g. using vision, sonar, IR, or laser range), prediction of target movements, multi-sensor fusion, and so forth.

One previously-identified application domain for multi-robot cooperation and learning is the multi-robot soccer domain,<sup>10</sup> which is of considerable interest in the research community. Like the multi-robot soccer domain, the CMOMMT application requires addressing many issues that are of importance in general multi-agent systems. Stone and Veloso<sup>11</sup> have conducted an extensive survey of the multi-agent learning research, and have identified the following issues that are of importance in various types of multi-agent interaction:

- Homogeneous non-communicating agents:
  - Reactive vs. deliberative agents
  - Local or global perspective
  - Modeling of other agents' states
  - How to affect others
- Heterogeneous non-communicating agents:
  - Benevolence vs. competitiveness
  - Evolving agents (arms race, credit/blame)
  - Modeling of others' goals, actions, and knowledge
  - Resource management (interdependent actions)
  - Social conventions
  - Roles
- Heterogeneous communicating agents:
  - Understanding each other
  - Planning communicative acts
  - Benevolence vs. competitiveness
  - Resource management (schedule coordination)
  - Commitment/decommitment
  - Truth in communication

Like the multi-robot soccer domain, the CMOMMT domain also encompasses these challenging issues of multi-agent cooperation. The CMOMMT application also has the benefit of being perceived as a "real-world" task, as opposed to a game (even though the issues of cooperation are quite similar), which is of importance in some research endeavors, and is potentially important in the perception of the research by the general public. Additionally, one aspect of CMOMMT that is different from multi-robot soccer is that it raises the issue of scalability much more than in multi-robot soccer. The issue of dealing with larger and larger numbers of robots and targets must be taken into account in the design of the cooperative control approach\*. Thus, we feel that these characteristics make the CMOMMT domain worthy of study for multi-robot cooperation, learning, and adaptation.

\*We assume that the game of soccer does not involve arbitrarily large numbers of players!

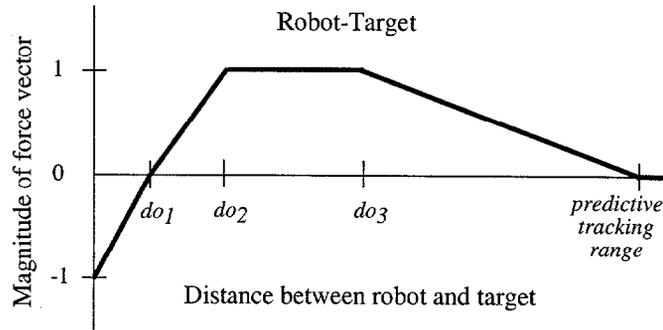


Figure 1. Function defining the magnitude of the force vectors acting on a robot due to a nearby target.

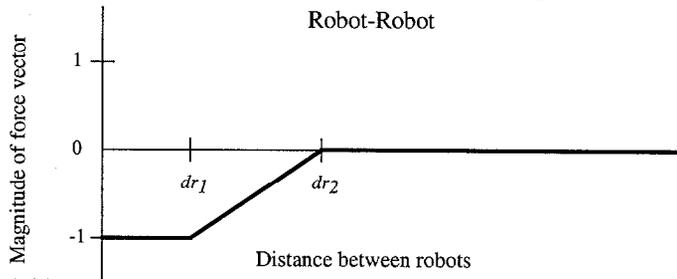


Figure 2. Function defining the magnitude of the force vector acting on a robot due to another nearby robot.

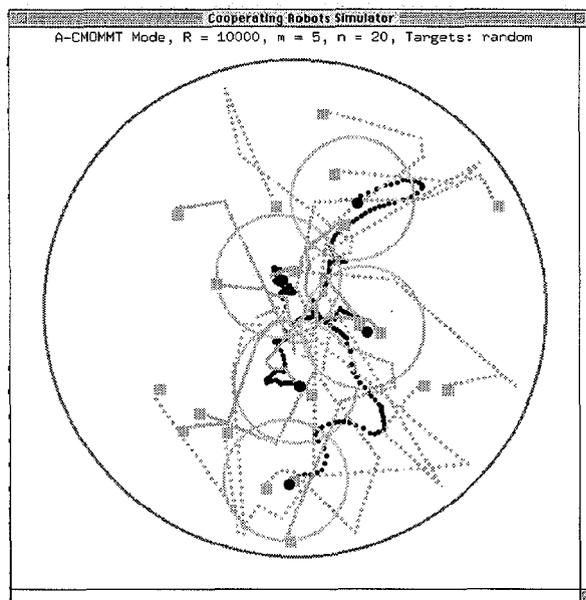
## 5. A HUMAN-DEVELOPED SOLUTION TO CMOMMT

In previous research,<sup>12,13</sup> we have developed a hand-generated solution to the CMOMMT problem that performs well when compared to naive approaches. This solution has been implemented on both physical and simulated robots to demonstrate its effectiveness. We emphasize, however, that the developed solution does not enable the robot team members to adapt their performance over time in response to changes in the environment, or changes in the team. The hand-generated solution also required considerable "tweaking" of parameters to develop a good solution.

The hand-generated solution, which we call *A-CMOMMT*, is described briefly as follows. Robots use weighted local force vectors that attract them to nearby targets and repel them from nearby robots. The weights are computed in real time by a higher-level reasoning system in each robot, and are based on the relative locations of the nearby robots and targets. The weights

The local force vectors are calculated using the functions illustrated in Figures 1 and 2. In this context, the *predictive tracking range* is a range just beyond the sensing capability of the robots, but within which robot trackers should continue to adjust their motions if targets are nearby (i.e., corresponding to targets that have just recently left the sensing range). The function in Figure 1 defines the relative magnitude of the attractive forces of a target within the predictive tracking range of a given robot. Note that to minimize the likelihood of collisions, the robot is repelled from a target if it is too close to that target ( $distance < do_1$ ). The range between  $do_2$  and  $do_3$  defines the preferred tracking range of a robot from a target. In practice, this range will be set according to the type of tracking sensor used and its range for optimal tracking. The proper setting of these thresholds is the subject of the multi-robot learning task we are proposing in this paper. The attraction to the target falls off linearly as the distance target is too far to have an effect on the robot's movements. The robot *sensing range* will lie somewhere between  $do_3$  and the predictive tracking range.

Figure 2 defines the magnitude of the repulsive forces between robots. If the robots are too close together ( $distance < dr_1$ ), they repel strongly. If the robots are far enough apart ( $distance > dr_2$ ), they have no effect upon each other in terms of the force vector calculations. The magnitude scales linearly between these values.



**Figure 3.** Simulation results of 5 robots at 20 targets, with the robots using the hand-generated solution to CMOMMT, and the targets moving randomly.

Using only local force vectors for this problem neglects higher-level information that may be used to improve the team performance. Thus, the hand-generated approach enhances the control approach by adding higher-level control to weight the contributions of each target's force field on the total computed field. This higher-level knowledge can express any information or heuristics that are known to result in more effective global control when used by each robot team member locally. The hand-generated approach expresses this higher-level knowledge in the form of a weight,  $w_{lk}$ , that reduces robot  $r_l$ 's attraction to a nearby target if that target is within the field of view of another nearby robot. Using these weights helps reduce the overlap of robot sensory areas toward the goal of minimizing the likelihood of a target escaping detection. Here, each robot assumes that its teammates have the same sensing range as its own.

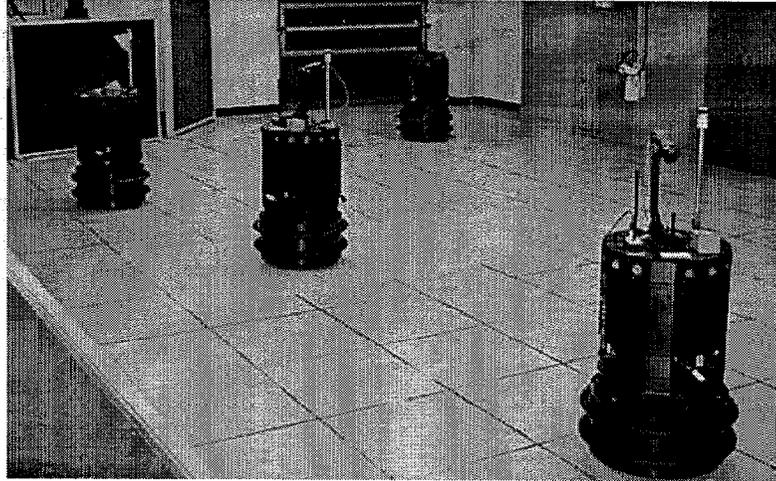
The proper setting of these weights is also appropriate for the multi-robot learning task proposed in this paper. Intuitively, it seems that if a robot  $v_l$  detects another robot  $v_j$  nearby and within sensing range of target  $o_k$ , then  $w_{lk}$  should be set to a low value. In the simplest case, since we define (in section 3) a robot  $v_l$  to be *observing* a target  $o_k$  when it is within  $v_l$ 's sensing range, we could assign  $w_{lk}$  to be zero whenever another robot is within sensing range of  $o_k$ . However, it is possible that will increase the likelihood that a target will escape detection, and thus it may give better results to set  $w_{lk}$  to some non-zero value. It remains to be seen if the results of the learning correspond to this finding.

The proper setting of  $w_{lk}$  is also dependent upon the estimated density of targets in the vicinity. If targets are sparsely located in the area, then the robot team risks losing track of a higher percentage of targets if any targets are ignored. On the other hand, if targets are densely distributed, then the risks are lower. Thus, a multi-robot learning task must address the proper computation of these weights based upon these issues.

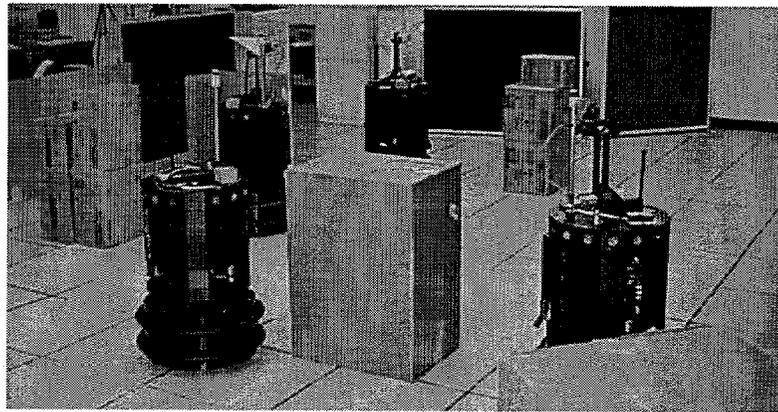
These weights have the effect of causing a robot to prefer the observation of certain targets over others. In more complex versions of the CMOMMT problem, robots could also learn about the viewing capabilities of their teammates, and discount their teammates' observations if that teammate has been unreliable in the past.

Figure 3 shows one of the simulation runs of the hand-generated algorithm (out of over 1,000,000 simulation test runs), in which 5 robots attempt to observe 20 targets. Figure 4 shows a snapshot of one of the physical robot experiments (out of over 500) in which the robots perform the task with no obstacles in the work area. Figure 5

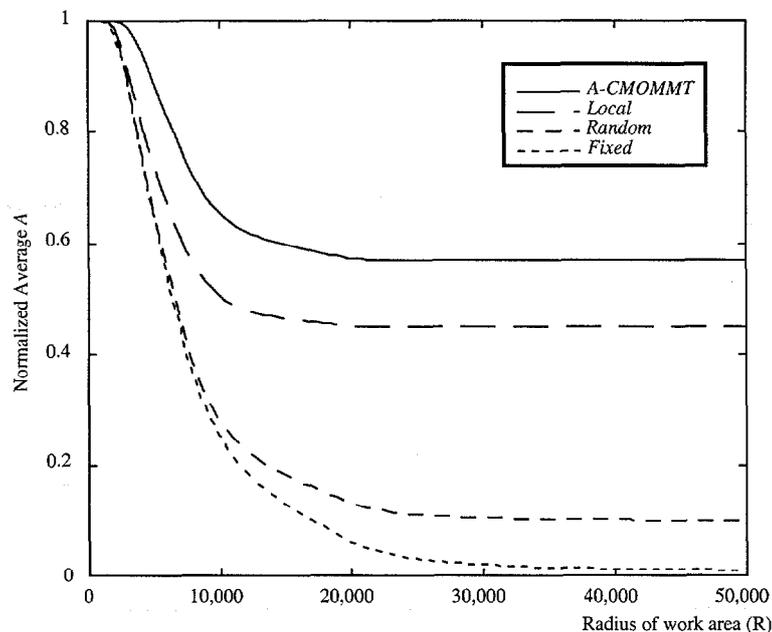
actual experimental runs).



**Figure 4.** Robot team executing hand-generated solution to CMOMMT in area with no obstacles.



**Figure 5.** Robot team executing hand-generated solution to CMOMMT in area with obstacles in the work area.



(d)  $n/m = 4$ , Targets move randomly

**Figure 6.** Comparison of *A-CMOMMT* results with non-weighted local force vector approach (*Local*) and two naive control algorithms (*Random* and *Fixed*).

The results of the hand-generated approach to CMOMMT vary depending upon a number of factors, including the relative numbers of robots and targets, the size of the work area, the motions of the targets (i.e., whether random

targets to robots greater than  $1/2$ . We compared the hand-generated *A-CMOMMT* approach with a non-weighted local force vector approach, as well as two control cases in which robots either maintained fixed positions or are moved randomly. Figure 6 gives a typical set of these comparative results. Refer to<sup>12</sup> for more details on these results.

## 6. LEARNING IN THE CMOMMT APPLICATION

### 6.1. Cooperative behavior learning without *a priori* models

Touzet has studied the CMOMMT problem<sup>14</sup> from a learning perspective without the assumption of an *a priori* model. In his approach, he combines reinforcement learning, lazy learning, and a pessimistic algorithm able to compute for each team member a lower bound on the utility of executing an action in a given situation. The challenges in this multi-robot learning problem include a very large search space, the need for communication or awareness of robot team members, and the difficulty of assigning credit in an inherently cooperative problem.

In Touzet's learning approach, lazy learning<sup>15</sup> is used to enable robot team members to build a memory of situation action pairs through random exploration of the CMOMMT problem. A reinforcement function gives the utility of a given situation. The pessimistic algorithm for each robot then uses the utility values to select the action that maximizes the lower bound on utility. The resulting algorithm is able to perform considerably better than a random action policy, although it is still significantly inferior to the hand-generated algorithm described earlier. Even with its relatively inferior performance, Touzet's approach is an important contribution, because it does not assume the existence of a model (as is the case in the Partially Observable Markov Decision Process – POMDP – domain), the existence of local indicators that help individual robots perform their tasks, nor the use of symbolic representations.

However, Touzet's approach does not address the issue of adaptation over time as the environment or robot team changes over time. In fact, the lazy learning approach assumes the environment is not changing. Thus, there is the need for new techniques that enable the robot team to adapt their performance over time.

## 6.2. Adaptation through parameter tuning

In previous research,<sup>16</sup> we developed the L-ALLIANCE approach to multi-robot adaptation that enabled teams of members and their capabilities in performing the tasks required by their mission. The L-ALLIANCE approach, which is an extension of our earlier work on ALLIANCE,<sup>17</sup> is a distributed, behavior-based architecture aimed for use in applications consisting of a collection of independent tasks. The key issue addressed in L-ALLIANCE is the determination of which tasks robots should select to perform during their mission, even when multiple robots with heterogeneous, continually changing capabilities are present on the team. In this approach, robots monitor the performance of their teammates performing common tasks, and evaluate their performance based upon the time of task completion. Robots then use this information throughout the lifetime of their mission to automatically update their control parameters. This approach has been shown to enable heterogeneous robot teams to continue to adapt to changes in the robot team member capabilities and in the environment.

However, the parameter update strategy used by L-ALLIANCE is dependent upon the assumption of independent subtasks whose performance can be evaluated based upon the time of task completion. This assumption does not hold for the CMOMMT application domain, and thus a new technique is required. In our current work, we are exploring a variety of approaches to online adaptation for the CMOMMT domain. In these approaches, we are giving the robot team members the general model of cooperation used in the hand-generated approach to CMOMMT – namely, the concept of weighted force vector functions that attract robots to nearby targets while repelling them from nearby robots. However, the parameters and thresholds of these functions are not fixed, and must be adapted by the robot team based upon factors such as the robot team size, the density of targets, the (potentially heterogeneous) sensing capabilities of the robots, the online performance of the team, the behavior of the targets, and so forth. Future articles will present the approaches and results of this current research.

## 7. CONCLUSIONS

In this paper, we have proposed that the Cooperative Multi-robot Observation of Multiple Moving Targets (CMOMMT) application domain provides a rich testbed for learning and adaptation in multi-robot cooperative teams. We have described the need for learning and adaptation in multi-robot teams, and have defined the CMOMMT application, along with the characteristics that make it an interesting testbed for learning and adaptation. We reported on a hand-generated solution to the CMOMMT problem and discussed how the results from the implementation of this solution reveal the need for learning and adaptation in this domain. We discussed some current work that uses the CMOMMT problem as a learning domain, as well as some of our ongoing efforts to generate new techniques for on-line multi-robot adaptation. The hope is that the new techniques eventually developed using the CMOMMT domain will generalize to other real-world domains, and will thus help realize the ultimate goal of enabling the widespread, practical use of multi-robot teams.

## REFERENCES

1. L. E. Parker, "On the design of behavior-based multi-robot teams," *Journal of Advanced Robotics*, 1996.
2. G. Weiss and S. Sen, eds., *Adaption and Learning in Multi-Agent Systems*, Springer, 1996.
3. R. Steeb, S. Cammarata, F. Hayes-Roth, P. Thorndyke, and R. Wesson, "Distributed intelligence for air fleet control," Technical Report R-2728-AFPA, Rand Corp., 1981.
4. M. Benda, V. Jagannathan, and R. Dodhiawalla, "On optimal cooperation of knowledge sources," Technical Report BCS-G2010-28, Boeing AI Center, August 1985.
5. R. Korf, "A simple solution to pursuit games," in *Working Papers of the 11th International Workshop on Distributed Artificial Intelligence*, pp. 183–194, 1992.
6. T. Haynes and S. Sen, "Evolving behavioral strategies in predators and prey," in *Adaptation and Learning in Multi-Agent Systems*, G. Weiss and S. Sen, eds., pp. 113–126, Springer, 1986.
7. P. Stone and M. Veloso, "A layered approach to learning client behaviors in the robocup soccer server," *Applied Artificial Intelligence* **12**, pp. 165–188, 1998.
8. S. Marsella, J. Adibi, Y. Al-Onaizan, G. Kaminka, I. Muslea, and M. Tambe, "On being a teammate: Experiences acquired in the design of robocup teams," in *Proceedings of the Third Annual Conference on Autonomous Agents*, O. Etzioni, J. Muller, and J. Bradshaw, eds., pp. 221–227, 1999.

9. J. Connell and S. Mahadevan, eds., *Robot Learning*, Kluwer Academic Publishers, 1993.
10. H. Kitano, Y. Kuniyoshi, I. Noda, M. Asada, H. Matsubara, and E. Osawa, "Robocup: A challenge problem for AI," *AI Magazine* **18**, pp. 73-85, Spring 1997.
11. P. Stone and M. Veloso, "Multiagent systems: A survey from a machine learning perspective," *submitted to IEEE Transactions on Knowledge and Data Engineering*, June 1996.
12. L. E. Parker, "Distributed algorithms for multi-robot observation of multiple moving targets," *to appear in Autonomous Robots*, 1999.
13. L. E. Parker, "Cooperative robotics for multi-target observation," *Intelligent Automation and Soft Computing* **5**(1), pp. 5-19, 1999.
14. C. Touzet, "Distributed lazy q-learning for cooperative mobile robots," *submitted to Autonomous Robots*, 1999.
15. D. Aha, ed., *Lazy Learning*, Kluwer Academic Publishers, 1997.
16. L. E. Parker, "L-ALLIANCE: Task-oriented multi-robot learning in behavior-based systems," *Journal of Advanced Robotics*, 1997.
17. L. E. Parker, "ALLIANCE: An architecture for fault-tolerant multi-robot cooperation," *IEEE Transactions on Robotics and Automation* **14**(2), pp. 220-240, 1998.