Multi-Target Observation: An Application of Multi-Robot Learning

Presenter: Rachel Yeargin

Mentor: Lynne Parker

RAM at ORNL, CSMD

University of Missouri Kansas City

Abstract

In recent years, interest in the area of the performance of multi-robot teams in cooperative tasks has significantly increased. As a result, scientists have delved into a new realm of research and experimentation with multi-robot learning. This paper examines various types of robot Learning and the benefits or challenges incorporated with each type. The Cooperative Multi-Robot Observation of Multiple Moving Targets (CMOMMT) application is presented and viewed as a valuable testing domain in the field of multi-robot teams and cooperative learning. The hand-generated approach to this application (A-CMOMMT) will be used as a control model in our research into generating learning technique that can improve upon this approach's results. The goal for this Particular project is to improve the performance of a previous approach that uses lazy Q-learning and self-organizing maps. This project is designed to generate a learning algorithm that reaches or exceeds the performance of the hand-generated approach. We do this by introducing a component that allows the robots to not only be aware of the nearby targets, but also of the nearby robots and their actions. The ultimate And considerably broader goal of this research is to develop learning techniques that allow for a more generalized application of cooperative robotics to numerous real world problems.

What is Robot Learning?

- Robot learning is the automatic modification of the robot behavior to improve its performance.
- Robot learning is the selection of an efficient behavior from the set of potential behaviors.
- So a learning robot is one that can improve its behavior as a result of interaction with an environment.

Advantages of Multi-Robot Teams

- Distributed Action: Many robots can be in many places at the same time
- Inherent Parallelism: Many robots can do many different things at the same time
- Divide and Conquer: Certain problems are better suited for decomposition and allocation among many robots
- Simpler is Better: Solutions in a team of robots can be simpler than a comprehensive solution in one robot.

Complications of Multi-Robot Learning

- Adds much large search space
- Requires awareness of other team members
- Demands synthesis of individual behaviors with respect to the tasks given to the entire group
- Actions of one member of the group depend on the actions of the other team members (Inherently cooperative tasks)

Types of Learning

- Reinforcement Learning involves an agent which is the learner or decision maker, the environment which is everything surrounding the agent, and the actions which are things that the agent can do. Each action is associated with a reward. The goal is for the agents to choose actions in a given environment in such a way that it maximizes the reward.
- Lazy learning samples the situation-action space, storing the succession of events in memory, and probes the associative memory for the best move. This exploration is done only once and is then stored and used by all future experiments.
- Q-Learning uses a state-action table which contains the gain that the agent obtains by executing a action from a state. The table represents a function Q which tells the agent which action to execute in order to obtain a maximum gain.
- Lazy Q-learning is becoming the most widely used type.

Self-Organizing Maps (SOM)

- A clustering technique that adds a neighborhood property between clusters.
- The number of clusters is fixed or predefined.
- The positions of the clusters in the input space are defined by the input point distribution.
- Size is directly related to distribution.
- The greater the density of the clusters the smaller the field of attraction.

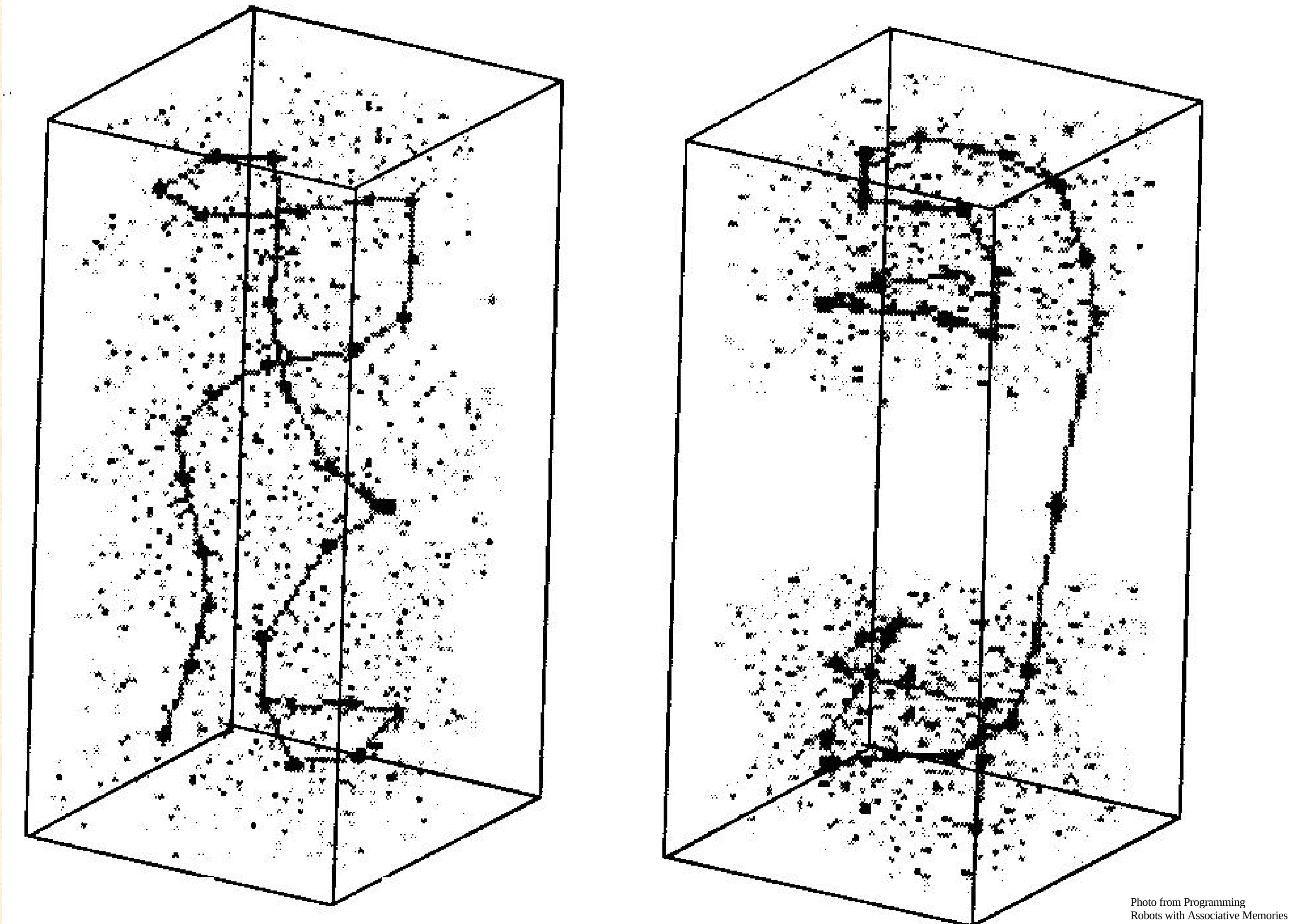

Photo from Programming
Robots with Associative Memories
By Claude Touzet

Cooperative Multi-Robot Observation of Multiple Moving Targets

A team of robots with 360° field of view sensors of limited range has to maximize the observation time of a set of targets moving randomly in a bounded area.

A robot is said to be monitoring the target when the target is within the robot's field of view.

The object is to maximize the time during which targets are being monitored by at least one robot.

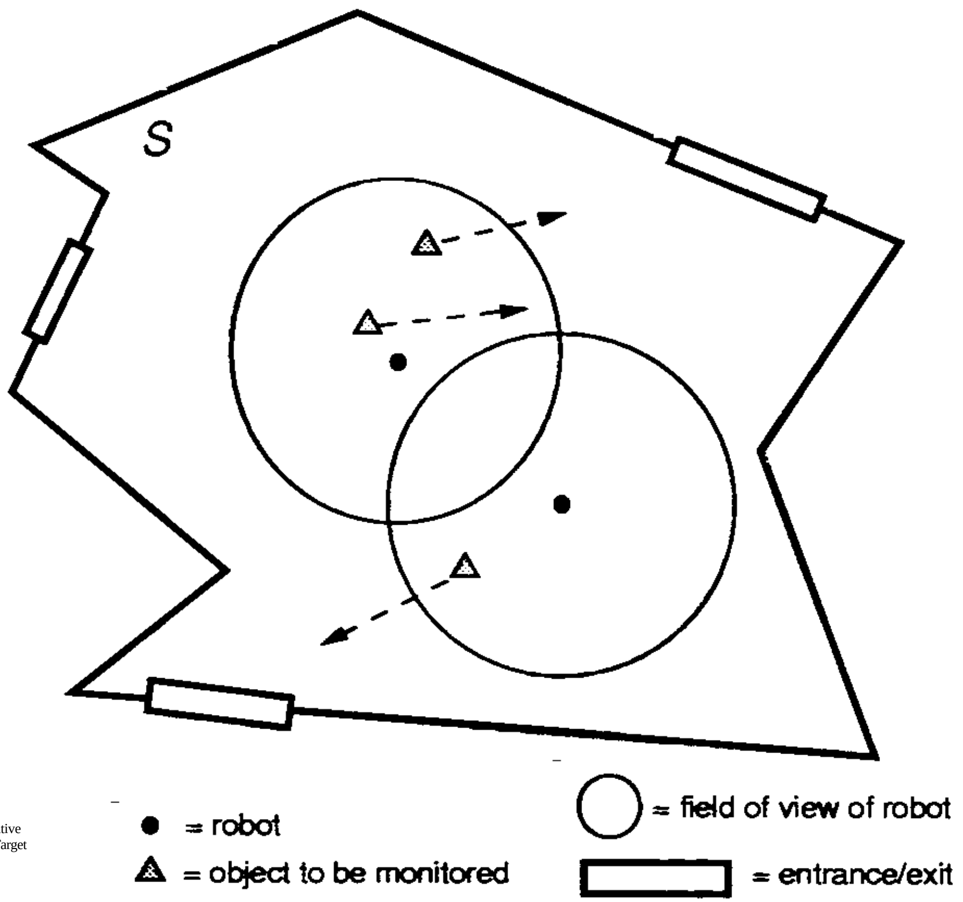

Photo from Cooperative Robotics for Multi-Target Observation By Lynne E. Parker

Goals and Applications

- The main goal of this application is to, maximize the average number of targets in the bounded region that are being observed by at least one robot throughout the mission.
- There are numerous applications of the ability to monitor a number of targets in a bounded area. Most of these are present in surveillance and security tasks.

Hand Generated Solution

- Called the A-CMOMMT
- Robots use weighted local force vectors that attract them to nearby targets and repel them from nearby robots.
- Weights are computed in real time and are based on relative locations.
- These weights create an improved collective behavior across all the robots.

Existing Approach

- ❑ Lazy Q-learning and self organizing maps can be used in the CMOMMT problem to provide much needed flexibility.
- ❑ We use two self-organizing maps. The first builds a representation of the situation space. It is used to find the path towards the goal situation. The second map generates the action that will change the sensory inputs from the current to the intermediate situation.

Existing Approach

- Each robot behavior is learned by lazy reinforcement learning.
- The lazy memory is created through the initial exploration of the targets and robots through random selection policy. Which allows a memory of situation-action pairs to be built.
- We use the following reinforcement function:

  +1 if the number of acquired targets has increased in comparison to the previous situation.

  -1 if the number of acquired targets has decreased in comparison to the previous situation.

  0 otherwise

Graphical Comparisons

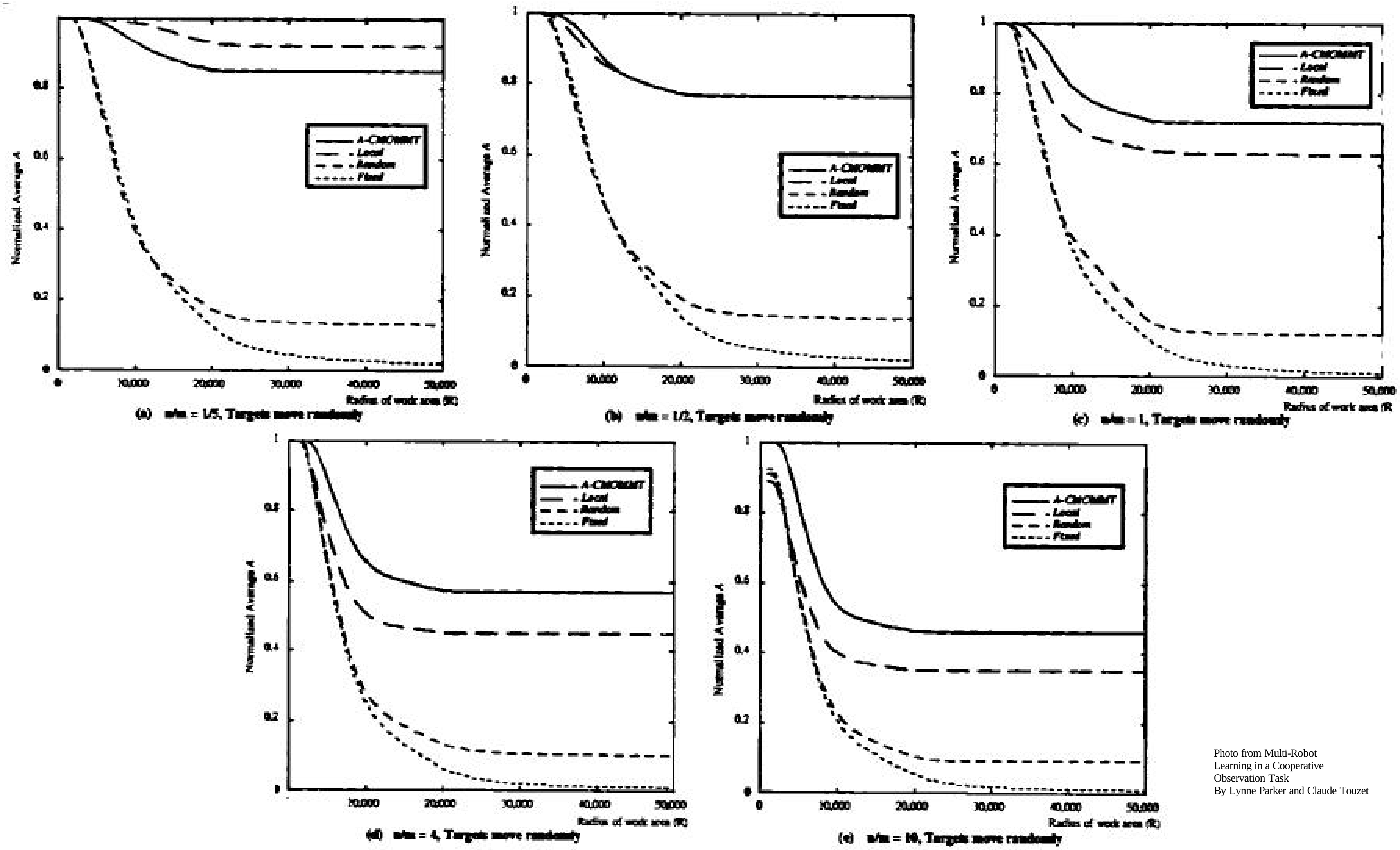

Photo from Multi-Robot
Learning in a Cooperative
Observation Task
By Lynne Parker and Claude Touzet

Observations

- This learning has not yet taken into account the positions and velocities of other robots. In other words the robots are not aware of one another and this has a definite impact on the results.
- For particularly small observation areas robot awareness has a significantly positive impact because it allows robots to select untracked targets and maintain minimal distance between team members.
- In larger areas, the impact is not as great because in the larger spaces robots and targets do not come close to one another.

Conclusions

- **In conclusion, it can be determined that despite the already extremely successful approach to the CMOMMT problem the A-CMOMMT, this application can be improved by introducing the option of inter-robot awareness.**
- **Since robot team members already determine the positions and velocities of targets within their own field, they should now be able to determine these same things about the robots in their field of vision.**
- **This new approach, based on the already existent lazy learning approach, will minimize the occurrence of robots converging upon the same robot and leaving another untracked.**

Acknowledgements

Special thanks to Lynne Parker for all of your patience and guidance throughout this project. To the Department of Energy and the Computer Science and Mathematics Division thanks for making this internship a possibility, it was a wonderful learning opportunity.

This research was performed under the Research Alliance for Minorities Program administered through the Computer Science and Mathematics Division, Oak Ridge National Laboratory. This Program is sponsored by the Mathematical, Information, and Computational Sciences Division; Office of Advanced Scientific Computing Research; U.S. Department of Energy. Oak Ridge National Laboratory is managed by UT-Battelle, LLC, for the U.S. Department of Energy under contract DE-AC05-00OR22725.