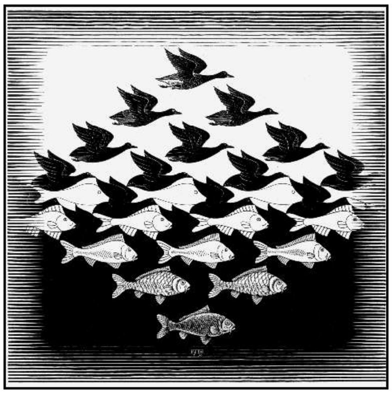
Chapter 9

Conclusion



This work is based on the idea that behavioral diversity should be evaluated as a *result* rather than an initial condition of multi-robot experiments. Previously, researchers configured robot teams as homogeneous or heterogeneous *a priori*, then compared performance of the resulting teams [FM97, GM97, Par94]. That approach does not support the study of behavioral diversity as an emergent property in multi-robot teams.

Defining behavioral diversity as an independent rather than dependent variable enables the examination of heterogeneity from an ecological point of view. How and when does diversity arise in robot teams interacting with each other and their environment? This work provides the methodology and quantitative measures necessary for this new type of investigation.

The specific contributions of this work are

- **a methodology** for building and evaluating learning behavior-based robot teams (Chapter 3),
- **a taxonomy of multi-robot reinforcement functions** (Chapter 4),
- **a classification of multi-robot tasks** (Chapter 4),
- **a quantitative measure of behavioral diversity** in multi-robot teams (Chapter 5),
- **a quantitative measure of behavioral difference** between individual robots (Chapter 5),
- **evaluation procedures** for multi-robot task performance (Chapter 3), and
- **a significant body of experimental results** illustrating the use of these tools in three multi-robot task domains (Chapters 6, 7 and 8).

This chapter reviews the key contributions of the research, provides an analysis of the experimental data and concludes with a discussion of promising directions for future work.

9.1 Methodology

Principled research in any field requires adherence to a methodological framework. Over the last decade the Mobile Robot Laboratory at Georgia Tech has evolved and refined a successful approach to behavior-based robot design and implementation. Key components of the method are the use of simulation for experimentation and behavioral prototyping along with verification of the results on mobile robots.

The framework was extended significantly in this research. First, a formal view of multi-robot task is adopted; multi-robot tasks are classified according to how performance is measured in them. This enables a principled description and exploration of the multi-robot task space. Second, a classification of reward functions is employed. This classification defines an experimental space for investigating the impact of reward on multi-robot systems. To support these experiments, motor schema control and reinforcement learning are integrated using a new object-oriented system for behavioral specification. Finally, new evaluation metrics necessary for the measurement of diversity in multi-robot teams are developed and employed in the analysis of experimental data.

The design and implementation of multi-robot systems and their experimental evaluation is carried out in the following steps:

1. **Task and performance metric specification:** This step defines performance, one of the dependent variables of experimentation.

2. **Behavioral design:** In this phase, a library of behaviors are developed for solving the task. Both hand-coded and learning systems are built using the behavioral components.

3. **Reinforcement function specification:** A goal of the research is to explore how different reinforcement functions impact performance and diversity in learning systems. In each task domain, several reinforcement functions are employed, with primary focus on the comparison of local and global rewards.

4. **Simulation:** The behaviors and learning systems are prototyped and tested in simulation.

5. **Implementation on mobile robots:** Performance of the simulated system is validated on mobile robots. If inconsistencies are discovered the simulation environment is refined to more closely approximate mobile robot performance.

6. **Data collection:** Multiple runs (thousands, usually) are conducted in simulation, and when possible, on mobile robots. The experimental space is explored by varying the independent variables (e.g. number of robots and/or the reward function).

7. **Analysis:** The data are analyzed using the performance metric the measures of diversity presented in Chapter 5.

Each step in the methodology is covered in detail in Chapter 3. The remaining contributions of the research each play an important role in the application of this methodology: quantitative measures of diversity and behavioral difference are used to evaluate the multiagent systems developed using the methodology; classifications of task and reward establish a frame of reference for analyzing the experimental results.

9.2 Classifications of task and reward

Task characterization helps answer questions regarding how the same type of reinforcement can lead to different performance and and diversity levels in different tasks. Without answering "how are soccer and foraging different?" for instance, we can't answer "why is diversity good in soccer but bad in foraging?"

A taxonomy of reward structure is also important. Experimental results (Chapters 6 and 7) show that performance and diversity in a learning team depend on the form of reinforcement used to train the robots. The results indicate that there are tradeoffs to consider in the selection of a reward function; for instance some functions provide quicker learning but slightly poorer performance. Without a characterization of the differences between reward functions, it would be impossible for a robot systems designer to consider these tradeoffs intelligently.

Prior to this research no taxonomies of task or reinforcement existed. To address this crucial gap a new system for characterizing multi-robot tasks and a taxonomy of reinforcement functions for multi-robot teams are introduced in Chapter 4. As well as aiding this investigation, these classifications are potentially useful for other researchers investigating multiagent robotic systems.

One focus of the experiments in this work is to determine whether it is better for each agent to be rewarded individually or if all agents should receive the same reinforcement. `GLOBAL` reinforcement refers to the case where a single reinforcement signal is simultaneously delivered to all agents, while `LOCAL` reinforcement rewards each agent individually. In the case of soccer, a global system would reward all team members when any one of them scores a goal. With local reinforcement only the agent that scored the goal would be rewarded. Global rewards correspond more closely with overall system performance, but they may not be appropriate for all tasks. The relative advantages of these two types of reward examined in each of the three experimental task domains.

Classifications of task and reward provide a description of the independent variables in this work. We now consider the evaluation of the key dependent variable: diversity.

9.3 New quantitative metrics

It is impossible to correlate heterogeneity with performance in multiagent robotic systems without a quantitative metric of diversity. Previously, diversity in multi-robot teams was evaluated on a bipolar scale with systems classified as either *heterogeneous* or *homogeneous*, depending on whether any of the agents differ [FM97, GM97, Par94]. Unfortunately, this labeling doesn't tell us much about the *extent* of diversity in heterogeneous teams. Heterogeneity is better viewed on a sliding scale providing for quantitative comparisons. Such a metric enables the investigation of issues like the impact of diversity on performance, and conversely, the impact of other task factors on diversity.

Social entropy, inspired by Shannon's information entropy [Sha49], is introduced as a measure of diversity in robot teams (Chapter 5). It captures important components of the meaning of diversity, including the number and size of groups in a society. In order to evaluate the diversity of a team, however, a way to categorize or differentiate the behavior of individuals is also required. To address this, a measure of *behavioral difference* that provides for agent categorization is also developed. Difference refers to disparity between two specific agents, while diversity is a measure of the entire society.

Diversity may not always be desirable. In fact, experimental results presented in Chapter 6 show that for at least one multi-robot task (multi-foraging) homogeneous robot teams perform better than diverse teams. The aim of this work is to discover when diversity is important and which conditions give rise to it in learning teams. Social entropy provides the objective quantitative metric required for a principled investigation of these issues.

This research is focused specifically on diversity in teams of mechanically similar agents that use reinforcement learning to develop behavioral policies. Evaluation of diversity in teams of mechanically similar robots is challenging because when agents differ, they differ only in their behavior. Behavior is an especially interesting dimension of diversity in learning systems since as they learn, agents effectively choose between a hetero- or homogeneous society. The metrics developed in this work will help researchers investigate the origin and benefits of diversity in these learning systems.

The next section describes how the new metrics are applied in the evaluation of learning multi-robot teams.

9.4 Experiments

The relationships between task, reward, performance and diversity are explored through experiments in three task domains: multi-foraging, soccer and cooperative movement. The motivation for experiments in a variety of domains is to investigate how differences in task impact the utility of diversity, and to see whether the usefulness of a particular reward structure depends on the task. Results are reviewed at a high level here. Details for each domain are reported in Chapters 6, 7 and 8.

Following the methodology introduced in Chapter 3, non-learning, or fixed, team strategies are developed for each task. A schema-based reactive control system is used for robot programming in each domain. In this approach, the agent is provided several pre-programmed behavioral assemblages that correspond to steps in achieving the task. Binary perceptual features (also referred to as perceptual triggers) are used to sequence the robot through steps in achieving the task. Selection of the appropriate behavior, given the situation, may be hand-coded or discovered by the robot through reinforcement learning.

While the focus of the dissertation is diversity in *learning* robot systems, human-coded non-learning teams play an important role as well. The performance of human-designed teams establishes a baseline for comparison with results from learning systems. Also, human-coded strategies provide additional data points regarding the relationship between diversity and performance in each task. Learning teams are provided the same behavioral assemblages from which to build a task-achieving strategy, but they must discover appropriate sequences of behavior through interaction with the environment.

The space of tasks and reward strategies explored experimentally is summarized in Table 9.1. Local and global reward strategies were evaluated in each task. Additional learning strategies were tested in the multi-foraging and RoboCup soccer tasks. In all three tasks, at least one of the learning strategies led to teams that perform as well as, and in some cases better than the best human-coded strategies.

The extent to which diversity is beneficial depends on the task. Table 9.2 summarizes these results. A link between diversity and performance is found in all three tasks. In the soccer and cooperative movement experiments the best performing systems are also the most diverse. In multi-foraging, however, diversity is strongly negatively correlated with performance: the best teams are homogeneous.

The utility of global versus local rewards also depends on the task. In soccer experiments global rewards are found to work best. Just the opposite is true in foraging; local rewards lead to the best-performing multi-robot teams. Finally, in the cooperative movement task local and global rewards generate systems that perform equally well. These results are also summarized in Table 9.2.

Table 9.1: Summary of the task and reward space explored in robot team experiments. Experiments for all three tasks were run in simulation. Foraging experiments were also conducted on mobile robots. The reward strategies are classified according to the taxonomy introduced in Chapter 4. In addition to the features listed, all rewards are also `INTERNAL_SOURCE`.

Task	Fixed Strategies	Learning Strategies

multi-foraging	homogeneous
	territorial
	specialize-by-color
		R_{local}
		`LOCAL PERFORMANCE DELAYED DISCRETE`
		R_{global}
		`GLOBAL PERFORMANCE DELAYED DISCRETE`
		R_{shaped}
		`LOCAL HEURISTIC IMMEDIATE CONTINUOUS`

soccer		

simplified	control team	R_{local}
		`LOCAL PERFORMANCE DELAYED DISCRETE`
		R_{global}
		`GLOBAL PERFORMANCE DELAYED DISCRETE`

RoboCup	control team
	BriSpec by McNamara
	DTeam by Johnson
	Kechze by Lyons *et al*
		R_{touch}
		`COMB_LOCALITY PERFORMANCE DELAYED CONTINUOUS`
		R_{global} (R_{touch} with $\gamma = 1$)
		`GLOBAL PERFORMANCE DELAYED CONTINUOUS`

cooperative movement	homogeneous diamond
	homogeneous line
	homogeneous column
		R_{local}
		`LOCAL PERFORMANCE IMMEDIATE CONTINUOUS`
		R_{global}
		`GLOBAL PERFORMANCE IMMEDIATE CONTINUOUS`

Table 9.2: Description of each task according to the classification introduced in Chapter 4 and a summary of the key results in each task.

	Multi-foraging	**Simplified Soccer**	**RoboCup Soccer**	**Coop. Movement**

Classification	`TIME_LIM`
	`RESOURCE_LIM`
	`OBJECT_BASED`
	
	`COMP_INT`
	`SINGLE_AGENT`
	`SENSOR_LIM`
		`TIME_UNLIM`
		
		`OBJECT_BASED`
		`COMP_EXT`
		`COMP_INT`
		`MULTI_AGENT`
		`SENSOR_COMPLETE`
			`TIME_LIM`
			
			`OBJECT_BASED`
			`COMP_EXT`
			`COMP_INT`
			`MULTI_AGENT`
			`SENSOR_COMPLETE`
				`TIME_MIN`
				
				`ROBOT_BASED`
				
				`MULTI_AGENT`
				`SENSOR_LIM`

diversity/perf. correlated?	yes, negatively	yes, positively	yes, positively	yes, positively
best reward(s)	R_{local} R_{shaped}	R_{global}	R_{global}	R_{global} R_{local}

9.5 Discussion of results

The long-range goal for the research begun here is a general model explaining the relationships between task, reward, agent diversity and performance in multi-robot teams. From a practical perspective this model will help robot system designers select appropriate learning strategies for robot teams according to the task for which they are designed. From a more philosophical point of view, the model might help explain how and why diversity arises in natural and artificial societies.

This dissertation provides the framework for the model by identifying and defining the independent (task and reward) and dependent (performance and diversity) variables. Mathematical relations between the variables will be derived as results are gathered from more points in the experimental space. At this point we can only hypothesize on the basis of results gathered thus far:

First, in all three domains examined here global rewards generate the most diverse societies. In soccer and foraging the globally rewarded teams are significantly more diverse than other systems. In the cooperative movement task global and local rewards generate about the same degree of diversity in robot teams. It makes sense that global rewards lead to greater diversity because it is more likely agents will be rewarded for different behavior under global reinforcement. We can hypothesize that, in general, global rewards lead to greater diversity and that global rewards are the better choice in domains where behavioral diversity is important.

Similarly, local rewards encourage uniformity in agent behavior. There are several advantages to the use of local rewards. First, they usually require only local sensing – the extra cost of communication is not incurred. Also, in some tasks (e.g. soccer and foraging) local rewards provide more rapid learning (policy convergence). In the foraging task, where homogeneous teams provide the best behavior, we also find that local rewards generate the best performing teams. It is likely that local rewards are the best choice in tasks where homogeneous behavior is preferred.

These observations beg the question "when is diverse behavior preferred?" The results of this work show that diverse teams perform best in soccer and cooperative movement tasks. Foraging, however, seems to be a decidedly homogeneous task. The task classifications provide a clue: soccer and cooperative movement are both classified as `MULTI_AGENT`, while foraging is a `SINGLE_AGENT` task (this distinction is based on whether an individual agent could reasonably perform the task alone). We can speculate that homogeneous teams excel in `SINGLE_AGENT` tasks.

9.6 Future directions

Important future work includes the application of these tools in new and different multi-robot task domains. The author hopes other researchers will adopt the measures of behavioral difference and robot team diversity introduced here in the evaluation of new multi-robot systems. This will provide additional data points in the multiagent task/reward space and help us derive the relations between task, reward, diversity and performance more precisely.

Another important direction for future research is the extension of these tools to a broader range of robotic systems. The behavioral difference metric, for instance, is limited to the comparison of deterministic policies. Can we compare the behavior of agents coded in FSAs or more complex representations? It may also prove useful to extend and refine the classifications of task and reward. Perhaps a broader range of tasks can be described or more subtle distinctions can be drawn between them.

Finally, can the results of this research be applied in other fields? Researchers in behavior-based robotics often draw inspiration from biology and psychology; perhaps roboticists can provide tools for the sociobiologist. It is tempting, for instance, to draw parallels between robotic tasks and rewards and their counterparts in human and animal societies. Local and global reinforcement for robots, for example, bears a strong resemblance to the capitalist and socialist economies of human society. It would be presumptuous to suppose such comparisons are valid now, but as the research and theory mature we may gain insights into the origins and benefits of diversity in natural as well as artificial social systems.